DOCUMENT RESUME

ED 376 554 EA 026 214

AUTHOR Tymms, Peter
TITLE Evaluating the Long Term Impact of Schools.
SPONS AGENCY Economic and Social Research Council, Lancaster (England).
PUB DATE Jul 94
CONTRACT R-000-23-3726-01
NOTE 14p.; Paper presented at the National Evaluation Institute of the Center for Research on Educational Accountability and Teacher Evaluation (Gatlinburg, TN, July 10-15, 1994).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Accountability; Elementary Secondary Education; Foreign Countries; Instructional Effectiveness; Regression (Statistics); *School Effectiveness; Self Esteem; Student Motivation; *Teacher Effectiveness; *Teacher Influence
IDENTIFIERS *England

ABSTRACT

Should teachers be held responsible for their long-term influence over students? Should monitoring systems be extended to cover the long-term impact of teachers, departments, and schools? This paper describes a longitudinal study that extended a British comprehensive monitoring system--the A Level Information System (ALIS)--to examine the effects of effective and ineffective departments on their pupils as they moved on to University and paid employment. A total of 2,578 students took their A-level exams in 1988 and completed questionnaires toward the end of their coursework. Of these, 1,167 students (47%) were sent followup questionnaires in 1993. It was hypothesized that effective departments could have negative long-term consequences as they pushed their students on to courses for which they were unprepared. This might adversely affect students' self-esteem and academic achievement. Regression analysis indicated that the impact of having attended effective rather than ineffective departments was slight; rather, the people and life experiences that students encountered after graduation exerted significant influence. It is recommended that monitoring and accountability systems focus on teachers' short-term impact. Nine statistical tables are included. (LMI)

Evaluating the Long Term Impact of Schools[1]

Peter Tymms

Department of Education
University of Newcastle upon Tyne
England

Paper given at Gatlinburg July 1994

Abstract

Should teachers be held responsible for their long term influence over children and should monitoring systems be extended to cover the long term impact of teachers, departments and schools? A comprehensive monitoring system within England (ALIS: The A Level Information System) was extended to look at effective and ineffective departments and their pupils as they moved on to University and paid employment. It was thought possible that effective departments could have negative long term consequences as they pushed their pupils on to courses where they found themselves out of depth. This might have an adverse impact on their self-esteem and academic achievement. This hypothesis was not supported by the data and several independent positive long term consequences of attending effective departments were noted. The effects however were slight and the results are discussed in relation to monitoring, accountability and complexity.

[1] This work was support by the ESRC: Grant number R-000-23-3726-01

Introduction

There is widespread agreement about the purposes of schooling and fairly general agreement about the accountability of teachers so far as the short term impact of their work is concerned. One would expect that teachers would enable their pupils to learn and that the pupils should finish at the end of a course having positive attitudes towards the subjects they have been taught and that the teachers should finish with students who have had a reasonable quality of life within the school (see for example Oakes 1989; Fitz-Gibbon 1985, 1992). There would, of course, be considerable debate around the specifics within these general principles but at a reasonably high level of generality there is agreement. Most educationalists would also agree that teachers should only be held accountable for outcomes over which they have control. And it is useful here to consider the two possible sources of variation in the outcomes of systems as described by Deming (1986). He suggested that outcome variation could be ascribed to **common cause** or to **special cause**. In the educational world, for example, it might be reasonable to suggest that teachers should not be held accountable for the commonly observed phenomenon that males rather than females are most involved in violent crime. On the other hand teacher might reasonable be expected to be held accountable for some of the **progress** which students make in their classrooms. The former has its root in society and biology (common cause) whereas the latter is much more specific (special and, presumably common cause). One of the challenges at the interface between educational research and classroom practice is to explore the extent to which teachers and teaching are responsible for the variation in outcomes seen between pupils, classrooms, schools and districts.

School effectiveness studies have provided us with a good basis from which to look at the accountability of teachers, schools and districts in the short term (Tymms 1993). That is to say we now know how to operate monitoring systems which will give us quantitative information relating to the performance of the students. For example, by looking at the residual gains for students, as measures of the progress which they have made, and by aggregating them up into the class level or to another appropriate unit of analysis a way forward may be found. There are still questions concerning the extent to which those residuals should be used as informational or as controlling or as linked to some other kinds of mechanisms in relation to teachers. But, we know how to do it: the technicalities are quite clear. What is not so clear is the extent to which one should be looking at the long term impact of schools. A teacher teaches his or her pupils for a certain period. What is the impact of that teaching on the pupils five years later? If there is a large impact then the accountability issue is very great. If a teacher, early in elementary school, turns a pupil off mathematics for life then that is a disastrous consequence and probably more important than any short term loss. Or if the teacher is able to motivate a child to be interested in science and this leads to a long term career, then that is probably of far greater import than a few gains in points on maths, science or reading scores. This paper addresses the issue of long term impact and it does so using research carried out in England within the A Level Information System (ALIS) which is described in another paper given at this conference (Fitz-Gibbon 1994)

Background to the research

There is a considerable body of research, originating from studies from around the world, to suggest that schools have an impact on the achievements and attitudes of their pupils (see for example Rutter 1979, Willms 1985, 1987) Several researchers have emphasised that schools may not be the best unit of analysis and it is departmental effects which should be considered.(Fitz-Gibbon, 1991a, 1992a; Smith & Tomlinson, 1989; Tymms 1993). But, is it possible that apparently effective departments can have a negative long term impact? Perhaps departments which help their pupils to get higher grades than they might have been expected to have obtained at more mediocre establishments do their pupils no favours. Could it be that those pupils move on to higher level courses in which they find themselves out of their depth? This might have a detrimental effect on their academic Self-Concept and so, in turn, on their achievement. This hypothesis, dubbed the Big Fish Little Pond effect (Davis 1966, Marsh 1987, Marsh 1991a), has some empirical support (Alwin and Otto 1977 and Marsh 1991a). Conversely, students who have not been supported in the school environment might find themselves in situations where they are more able than their peers. In this situation they might thrive and outperform those around them.

An alternative view holds that what matters is not so much how well a student succeeds as **what** he or she is taught. The subjects studied may have greater impact on the long-term, *retained* knowledge of students and may have more consequences for subsequent life chances - - - such as employment opportunities and quality of life - - - than does the particular level of attainment in any subject. 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 (Straker, 1988). Preece (1983), studying achievement in science, suggested that the curriculum offered has far more impact than variations in the effectiveness of the instruction. He referred to this as the qualitative hypothesis: "Learning is largely unaffected by differences in teaching style and approach." This view that what is taught is of more importance than how it is taught needs to be considered in the light of evidence relating to the relative difficulty of subjects.

It is well established that some examinations are "easier" than others (Kelly 1976; Fitz-Gibbon 1988, 1991b) in the sense that apparently similar pupils tend to be awarded higher grades in some subjects rather than others. Could it be that students who follow easy courses tend to achieve more in the long run? Analysis of the A level results suggested that in terms of raw grades it may be more advantageous to follow an "easy" course than to attend a "good" department.

Methods

Students who had taken their A level examinations in 1988 were followed up in 1993. The students had all taken part in the A Level Information System (ALIS) (Fitz-Gibbon 1985, 1990, 1991a, 1992) and had completed a questionnaire towards the end of their A level courses. In 1988 ALIS covered five north-eastern Local Education Authorities and 2,578 students took part in the survey. Five hundred and forty three replies were received to the follow up survey. Only those students (1,167) who had indicated, on completing the first questionnaire, that they would be happy to be contacted were actually sent questionnaires in 1993. Thus, the response rate was 47% of those who were sent a questionnaire.

The approach adopted during the statistical analysis was to estimate the impact of students' choices and the effectiveness of A level provision on outcomes in 1993 including degree classification, salary, prospects. Adjustments were made for important characteristics which were independent of and/or prior to the A level course.

Variables were available at three time points:

Time one (T1)

AVO	The mean grade achieved by students at O level. (O level was a national attainment qualification taken two years before A level.)
SES	The socio-economic status (SES) of parents.
Gender	(0=Male 1=Female)
ITDA	The International Test of Developed Abilities. (Ottobre and Turnbull 1987)

Compared with the original sample the follow-up group had very similar O level grades, came from similar home backgrounds and had comparable measured developed ability. The follow up group had a higher proportion of females (61%) than the original sample (48%). See Table A.

Time two (T2) At the end of the A level course

A level grades	
Attitudes to Subjects for	Summated ratings from 6 points scales for each of 11 subjects.
Attitude to Institution	Summated rating from a 6 point scale.
Facility	The relative leniency of the particular set of A levels course followed by an individual calculated for that individual by considering each A level in turn and his or her AVO in relation to the general pattern found in the data.

Institutional Boost	The mean "effectiveness" of the A level departments in which the student was taught - calculated via multi-level modelling (Rasbash, Prosser, and Goldstein, H. 1989; Tymms 1992, 1993)

See Table B

Time three (T3) Follow-up survey

The questionnaire responses indicated that about half of the students who took their A levels in 1988 still lived in the north-east in 1993 and about a third were living with their parents. Nearly three-quarters were in employment and a fifth were still studying. About 10% were unemployed.

Three quarters had obtained a degree. The most striking differences between those who had a degree and those that did not were that the graduates were less likely to be married, more likely to be renting accommodation, more likely to have moved away from the north-east and more likely to be unemployed.

The last observation may seem odd in view of the many studies (see for example Garner et al 1988) which have indicated the importance of educational qualifications in finding employment and it may be that if the survey had been conducted a little later that more graduates would have found jobs. Indeed the Associate of Graduate Recruiters survey (1993) concluded that "1993 will be a difficult year for graduate job hunters but employers are more optimistic about the future".

The questionnaire had many parts to it but important variables constructed from the questionnaire responses. were:

Degree	Whether a student had obtained a degree or not. (0=yes 1=no)		
Degree class	(0=fail	1=pass	2=3rd
	3=2:2	4=2:1	5=1st)

Self-description measures were collected using the instruments developed by Marsh (1991b) Three Self-Concept scales (Marsh & Shavelson 1985) were measured.

Academic	
Maths	
Language	
Offer (Heap)	The number of A level points which were reported in Heap (1987) as being the expected requirement for the course being followed.

On the grounds that effective, or ineffective A level provision may place a student on a course where he or she was inappropriately placed a variable was created to measure the relative standing of the course attended. The level of the course attended was measured by the entrance requirements of the course and the base line was taken as the AVO.

Relative Standing	The residual formed from a regression analysis in which Offer (Heap) was the dependent variable and AVO the independent variable.
Attitudes to Subjects for	Summated ratings from 6 points scales each of 11 subjects. (The same questions as at T2)
Attitude to Institution	Summated rating from a 6 point scale.(The same questions as at T2)

See Table C

Findings

Extensive analyses of the data focused on hypotheses constructed before the survey had been carried out. As noted earlier the procedure was generally to consider the relationship between T3 variables and T2 variables having controlled for important T1 variables. It was generally the case that the most important control variable was AVO (prior achievement). The results are summarised below:

- Students' A level grades appeared to be influenced in fairly equal measures by the effectiveness of the departments which taught them (Institutional Boost) and by the relative ease of the subjects taken (Facility). Those who attended effective departments tended to be advantaged by about a third of a grade per subject, whereas students taking a relatively easy set of A levels were advantaged by about a fifth of a grade. (Table B)

- Whether a student attained a degree or not was not related to the level of ease of the A level course studied (Facility). However, students who followed a relatively "hard" A level course were significantly more likely to move straight onto a successful degree course after A levels. (Table D)

- Students were more likely to obtain a degree if they had attended "effective" A level departments and this was true both for moving on to a degree course immediately following A levels and obtaining a degree in the long run. (Table E)

- There was no indication that degree classification was related to attendance at an effective A level courses.

- The Academic Self-Concept of students was higher the higher the effectiveness of the A level departments that they attended. And it was higher the more academic the HE course attended and the higher the Relative Standing. (Table F)

- Graduates reported less salary at the time of the survey than non-graduates, but they expected to earn a higher salary in five years time than non-graduates. A similar finding was reported in Winefield et al (1992) who noted that undertaking of Higher Education courses resulted in "temporary poverty" and "long term economic benefits". (Table G)

- For students who had obtained a degree the more effective their A level departments the higher their salary at the time of the survey. (Table H)

- For those who did not obtain a degree, the less effective their A level departments the higher their salary. (Table I)

It would seem that effective A level provision has a long term beneficial impact on students. Students attending effective A level departments were more likely to obtain degrees and have higher Academic Self-Concept. There was no evidence to support our original pessimistic hypothesis that students might suffer from effective A level provision.

There was confirmation of the slightly counter-intuitive hypothesis that students who failed to get onto degree courses having attended ineffective A level departments achieved higher salaries than would be expected otherwise.

However, all the relationships described above were small although significant at the 5% level. The measured long term impact was detectable but slight. (The Effect Size was generally around 0.2) It would seem that in the flow of life events with all that that means in terms of people, incidents, relationships, courses, accumulation of knowledge, etc. the impact of having attended effective rather than ineffective departments was slight.

A further part of the follow up survey involved estimating the extent to which students' views about A level provision over time were stable. Did they still have the same attitude to the subjects which they took at A level and to the Institution where they took their A levels and were they able to judge the effectiveness of the departments where they had taken their A levels?

The results indicated that:

- Students were able, to a limited extent, to estimate the effectiveness of their A level departments. The mean correlation between their ratings in 1993 and the measured effectiveness was 0.2 for departments in 1988. However, the ratings of students' estimates of their Institutions' effectiveness in getting good grades for their students only correlated 0.09 with the measured effectiveness.

- Students attitudes were fairly stable over time and the 1993 attitudes measures correlated with the 1988 measures at around the 0.6 level.

Conclusions

It would seem from the data there were long term consequences of attending effective or in-effective departments. These effects, however, were detectable rather than important. The problem is that the focus was on distal variables. After a pupil has attended a school a lot of things happen to him or her. He or she meets other people, has relationships, goes to other courses, hears new ideas, listens to the news, reads books, meets friends, relations die, and close friends get married or divorced. All of these things impinge on the life of the pupil. This implies that as the pupil moves away from the school situation the impact of that school gradually diminishes. Of course, one always hopes that acorns are planted and that is always a possibility but, in this kind of analysis, a low germination rate (or the impact on a small number of individuals) would not show up. In the research reported the search was for statistically significant, educationally important, long term impacts and they were not the found. A teacher might be expected to have an impact on a class of pupils at the time when he or she is responsible for the class, however, as time goes on that impact gradually wears off. It looks as though as we move away in time the increasingly distal variables decrease in importance. This fits in well with theories of chaos (Gleick 1988) and complexity (Lewin 1993; Waldrop 1992). We could imagine life as a river which flows downstream being buffeted by the banks and the stones and the rocks and rapids. If one were able to follow a particular drop of water through a rapid or a waterfall the immediate impact on the water on its motion and its temperature would be apparent, but as the water moved further down the long term impact of a particular rapid or waterfall would be fairly minimal. That is not to say that it would not have had some long term impact but rather that the effect would gradually wear out. Paradoxically in a chaotic situation small influences can have long term consequences such as the flap of a butterfly wing generating tornadoes. The difficulty is, of course, that one could not say in advance which single flap out of the billions is of consequence.

The implication of this for monitoring are clear. We should concentrate monitoring systems and accountability systems on the short term impact of teachers. This is fortunate since it would be expensive and difficult to monitor the long term impact. It is also fortunate to be able to report evidence for stable attitudes of students towards their subjects and institutions -- to have found otherwise would have raised serious questions about the validity of monitoring attitudes. It would be useful to be able to track the long run impact of teachers, departments and schools but that should not be a prime concern. The prime concern in monitoring should be the direct impact on the lives of the children within the school during that period. In seems likely that promises or thoughts along the lines of "I do this now because I think it will bring forth fruits in later years" are far fetched and claim too much for the teacher. This might seem to be a counsel of despair, it might seem as though there is not much we can do about things in the world and indeed we are beginning to understand that we can't predict the weather far ahead and that we don't know what the economy will look like two three five years time. Long term planning has not got a good basis in previous experience. We need to edge our way forward and that is as true for teaching and the monitoring of teaching as it is of the economy and other complex evolving systems. In Deming's terms, teachers should not be held accountable for long term effects since these are common cause outcomes. Short term impacts however, may be the result of specific cause and it makes sense to devote some effort to monitoring these effects which are, to some extent, the responsibility of teachers, departments and schools.

References

Alwin, D. F., & Otto, L. B. (1977). High School Context Effects on Aspirations. In *Sociology of Education* 50 259-273

Association of Graduate Recruiters (1993) *1992/1993 Salaries and Vacancies Survey* Cambridge: AGR

Davis, J.A. (1966) The campus is a frog pond: An application of theory of relative deprivation to career decisions for college men. *American Journal of Sociology* 72 17-31

Deming, W. E. (1986). *Out of the Crisis: Quality Productivity and Competitive Position* Cambridge University Press

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. (1988). Recalculating the standard. 26-8-88. *The Times Educational Supplement*, 15.

Fitz-Gibbon, C. T. (1990). Performance Indicators: educational considerations. In Levacic, R. (Ed) *Financial Management in Education* Milton Keynes: Open University Press. .

Fitz-Gibbon, C. T. (1991a) Multilevel modelling in an indicator system. In Raudenbush, S.W. and Willms, J.D. (Eds) *Pupils, classrooms and schools: international studies from a multilevel perspective* London and New York: Academic Press. .

Fitz-Gibbon, C. T. (1991b). *Evaluation of School Performance in Public Examinations: A report for the Scottish Office Education Department* Newcastle upon Tyne CEM Centre

Fitz-Gibbon, C. T. (1992). The design of indicator systems. In *Research Papers in Education* 7,3 271-300

Fitz-Gibbon, C.T. (1994) *Indicator Systems for School and Teacher Evaluation: Fire Fighting It Is!* Paper given at Gatlinburg July 1994

Garner, C.L., Main, G.M. and Raffe, D. (1988) Local Variations in School Leaver Unemployment within a Large City *British Journal of Education and Work* 1, 2, 66-78

Gleick, J. (1988). *Chaos: Making a New Science* . Heinemann London.

Heap, B. (1987) *The Complete Degree Course Offers: Official information on entry to British Universities Polytechnics and Colleges*. 1988 18th edition Worcester: Billing & sons

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

Lewin, R. (1993). *Complexity. Life at the Edge of Chaos*. Dent.

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self concept. In *Journal of Educational Psychology 79* 280-295

Marsh, H.W. and Shavelson, R. (1985) Self-Concept: Its multi-faceted, hierarchical structure *Educational Psychologist* 20 107-125

Marsh, H. W. (1991a). Failure of High-Ability High Schools to Deliver Academic Benefits Commensurate With Their Students' Ability Levels *American Educational Research Journal* 28, 2 445-480

Marsh, H.W. (1991b) *Self description Questionnaire (SDQ) III; A Theoretical and empirical basis for the Measurement of multiple dimensions of late adolescent self-concept: An interim test manual and a research monograph* School of Education, University of Western Sydney, Macarthur.

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

Rasbash, J., Prosser, R., & and Goldstein, H. (1989). *ML3 Software for three-level analysis* London: Inst. of Edu.

Rutter, M., Maughan, B., Mortimer, P., & Ousten, J. (1979). *Fifteen Thousand Homes.* London: Open Books.

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

Straker, N. (1988). Teacher shortage in mathematics and physics: some possible solutions. AM8HL pamphlet

Tymms, P. B. (1993). Accountability - Can it be fair? *Oxford Review of Education* 19, 3 291-300

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

Waldrop, N.N. (1992) *Complexity: The emerging science at the edge of order and chaos* Viking

Willms, J. D. (1985). Catholic School effects on academic achievement: new evidence from the High School and Beyond follow-up study. *Sociology of Education* 58, 98-114

Willms, J. D. (1987). Differences between Scottish educational authorities in their examinations attainment. *Oxford Review of Education* 13,2 211-232

Tables

Table A

Variable	Mean (1988)	Mean (1994)	SD (1988)	SD (1994)	n (1988)	n (1994)
AVO	5.50	5.57	0.64	0.65	2420	528
SES	3.87	3.85	1.08	1.11	1758	493
ITDA	38.93	39.32	8.21	7.99	2356	540
Gender	.48	.61				

Table B

Variable	Mean (1994)	SD (1994)
Facility	.05	.22
Institutional Boost	.04	.33

Table C

Variable	Mean	SD	n
Degree	0.76		544
D_class	3.15	1.20	411
Academic_SD	6.01	1.01	541
Maths_SD	5.19	1.82	541
Language_SD	6.38	1.05	541
Offer_actual	8.93	3.10	228
Rel_Stand	0	2.56	329

Tables D to I report correlation coefficients (* p<.05 ** p<.01)

Table D

Correlations	FACILITY	FACILITY control for AVO	FACILITY control for AVO & SES
Degree	-.04	.01	.04
Immediate Degree	-.05	0	-.10*

Table E

Correlations	Institutional Boost	Institutional Boost control for AVO	Institutional Boost control for AVO & SES
Degree	.09*	.08*	.08*
Immediate Degree	.11**	.10**	.09*

Table F

Correlations	Institutional Boost	Institutional Boost control for AVO	Rel. Stand
Academic Self-Concept	.10*	.08*	.15*

Table G

Correlations	Degree
How much did you earn last year?	-.10*
Salary in 5 years' time	.11*

Table H (For those with a degree)

Correlations	Institutional Boost	Institutional Boost control for AVO
How much did you earn last year?	.15*	.12*

Table I (For those without a degree)

Correlations	Institutional Boost	Institutional Boost control for AVO
How much did you earn last year?	-.18*	-.20*