CEM⊙

Curriculum, Evaluation and Management Centre

THE EFFECT OF A NATIONAL INITIATIVE ON EXAM

RESULTS

P.B. TYMMS
C.T. FITZ-GIBBON
R.D. HAZELWOOD

1990

The Effect of a National Initiative on Exam Results

P.B. Tymms, C.T. Fitz-Gibbon, R.D. Hazelwood, and J.J.C. McCabe
CEM, School of Education, The University of Newcastle upon Tyne

ABSTRACT

This paper considers the examination results of TVEI and non TVEI pupils in the database established at the University of Newcastle in 1985-87. The pupils took national exams in the summer of 1987, when they were about 16 years old. Their average grades (Exam Means) and their accumulated grades (Exam Totals) were considered separately.

The Exam Mean and Exam Total scores were generally worse for TVEI than non TVEI pupils, even when pupils with the same ability test scores and aspirational scores were compared. However, it is possible that non TVEI factors caused the discrepancy in results and it is these factors which are the subjects of this paper.

The paper considers firstly whether the differences are an artefact of regression analysis. Secondly, the data are considered using various multilevel models. Both analyses suggest that there was no TVEI effect on the Exam Mean score and a negative TVEI effect on the Exam Total. Various explanations for these findings, concerning internal school allocational and curriculum policies, are considered and discussed.

INTRODUCTION

There have been momentous changes in the organisation of state education within England and Wales during the 1980s. A national curriculum has been introduced, national testing is planned, schools are to manage their own budgets, 'City Technology Colleges' are being introduced and there are many other far reaching changes being implemented or being considered. The first major financial commitment to state educational innovation, introduced by the present government, was the Technical and Vocational and Education Initiative (TVEI), announced by the Prime Minister on 12th November 1982. The initiative, according to Margaret Thatcher, was intended to address 'growing concern about existing arrangements for technical and vocational education provision for young people'.

The scheme was coordinated through the Manpower Services Commission (MSC), now called the Training Agency, a central body which bypassed the usual intermediary, the Department of Education and Science, and worked directly with schools, colleges and Local Education Authorities (LEAs); this in itself was a controversial move.

Correspondence: P.B. Tymms, Moray House College, Holyrood Rd., Edinburgh, UK.

Generous funds were provided for the TVEI; Gleeson (1987) notes 'the 250 m pounds budget allocated to it ... in a period of falling rolls, cuts in public expenditure and school reorganisation.' Crucially, a proportion of the budget was required to be spent on local evaluation (sums of the order of 4000 pounds per LEA in 1986/87) in addition to national evaluation (200 000 pounds in 1986/87).

The initiative was introduced through pilot schemes, which came on stream in a series of rounds, the first of which started in 1983. (This paper draws its data from the local evaluation of Round 3 pilot schemes.) The pilots were concentrated in a selection of schools in certain LEAs and were directed at 14-18 year olds. The intention was that only a proportion of 14 year olds within the schools would become TVEI pupils and that they would continue in that vein until 18 years old, with the possibility of opting out at 16. The way in which pupils were selected to be part of TVEI varied from school to school, but it would not be too much of a simplification to say that it generally involved a collaborative process between pupils, parents and staff, in which the advantages of TVEI were made available through literature and talks. Individual counselling was also important. For most schools, the MSC's intention was that a fifth of 14 year olds, within a school, would be designated as TVEI pupils. This discriminatory aspect of the TVEI pilots schemes was resented in some areas of the educational profession.

It is important to note that TVEI was intended for the full ability range and was explicitly not simply for the less able. Despite this, the government had set up the assisted-places scheme (Edwards and Whitty 1989) whereby bright pupils in state schools were able to obtain places in selected private schools, at the tax payers' expense, where the TVEI scheme did not operate.

The aims of TVEI were outlined in the TVEI Operating Manual (MSC 1983) in which the first aim was: 'In conjunction with LEAs to explore and test ways of organising and managing the education of 14-18 year old young people across the ability range so that: i) more of them are attracted to seek the qualifications/skills which will be of direct value to them at work, and more of them achieve these qualifications.'

Although much has happened since the publication of that aim, it remained the first aim of the TVEI 'Extension', the next phase in the scheme, in which the TVEI was applied across whole LEAs for the entire 14-18 cohort.

But the TVEI has involved much more than simply the pursuit of qualifications/skills and it has been approached in distinct ways, both within individual schools and within particular pilots. Nevertheless, it would seem fair to refer to a 'TVEI philosophy' which has evolved and which is encapsulated by Pirie (1989) in her description of the Scottish TVEI. 'Changed methodology is a central purpose of the initiative, encouraging greater concern with problem-solving, with student-centred learning strategies, with independent and tutor-supported self-study and with the involvement of industry in curriculum development. Experience outside school, including work experience, community or enterprise projects and residential experiences are all integral to TVEI pilot schemes.'

In view of the TVEI's lavish funding and its first stated aim the paper by Fitz-Gibbon, Hazelwood, Tymms and McCabe (1988) came as a major shock when it reported that: 'In the schools for which we had 'hard data', TVEI pupils obtained worse examination results than non-TVEI pupils in the summer of 1987, and this finding remained even after corrections were made for initial differences between the abilities of TVEI and non-TVEI pupils. If the results on external examinations could be considered as outcome indicators for TVEI, then it must be said that the performance indicators were not favourable.'

The findings were reported in the national press and there was a flurry of letters in the Times Educational Supplement (15/7/88), presenting a variety of hypotheses to explain the findings and pointing to other achievements of the TVEI.

Interestingly, the MSC itself has never questioned the findings, despite its direct access to national TVEI data. Indeed, the first report on exam results by the TVEI Unit (Tenne 1989) concentrated almost exclusively on gender differences within TVEI and bland descriptive statistics. On the examination results of the non-TVEI students, data which the unit had, there was silence.

The only other published work to consider the TVEI/non-TVEI exam differences (Raffe 1988) unfortunately was unable to use any cognitive control variable, but it did indicate underachievement by the TVEI cohort, albeit significant only at the 6% level.

The purpose of this paper is to examine the TVEI/non-TVEI differences, reported in Fitz-Gibbon et al., in greater detail and to try to tease out more, from the extensive database established at Newcastle, than was possible in the first paper.

At the root of all studies of this sort is the problem of cause and effect. Although associations can be found between school products and processes and can be demonstrated to be unlikely to have arisen by chance, there is always the possibility that the associations result from confounds or artefacts rather than educational processes. So, did the TVEI initiative cause the lower achievement? Is it possible that the TVEI pupils would have had even worse results had it not been for the initiative? It is these questions which are to be tackled in this paper using essentially the same data as was used previously by Fitz-Gibbon et al. (1988).

The data was investigated in two ways. Firstly the methodology and findings of the Fitz-Gibbon et al. (1988) paper are looked at anew, by considering the possibility of an artefactual discrepancy arising between the TVEI and non-TVEI pupils when regression analysis is used to compare the two groups' exam results.

Secondly the apparent TVEI effect is related to the differential degree of implementation of the TVEI programme by schools, the hypothesis being that if the TVEI had negatively influenced the pupils' exam results then the schools which discriminated most between TVEI and non-TVEI pupils in terms of TVEI policy should also be the schools where the greatest exam discrepancies were visible.

REGRESSION TO THE MEAN

In describing the regressions use will be made of a graphical approach in which it is noted that the regression of Y on X is a line through the vertical tangents to an ellipse representing the scatter of the bivariate distribution as in Diagram 1. (Marks 1982).
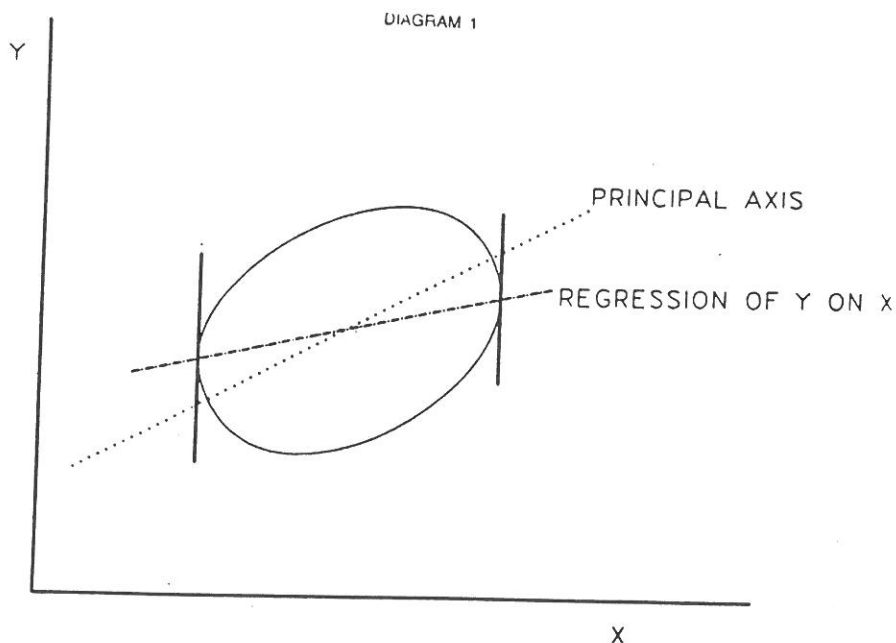
Diag. 1

Fitz-Gibbon et al. (1988) modelled the examination results using the equation:

$$Y_j = \beta_0 + \beta_1 \, Raven_j + \beta_2 LSE_j + e_j \qquad (1)$$

Where: Y_j was either the average GCE/CSE grade achieved or the total score for all exams taken for pupil j. These are referred to henceforth as the Exam Mean and the Exam Total. The two measures yield slightly different patterns.

$Raven_j$ is pupil j's score on Raven's Standard Progressive Matrices administered at the start of the fourth year. (Raven 1958).

LSE_j is a measure of pupil j's reported Likelihood of Staying in Education. It was formed from five Likert-type items each asking for responses on a five point scale. The items were part of a questionnaire administered at the start of the fourth year.

Although significant differences were found between the examination results of TVEI and non-TVEI pupils, the following point was made in the Fitz-Gibbon et al. paper:

'The principal difficulty was that the pupils in the TVEI pilot must have been different in many ways from the pupils not in the TVEI pilot and, although we made allowance for some important differences, no one could make allowance for all possible differences.'

In other words because the two groups of pupils were not equivalent the differences that were found could have arisen even if the TVEI programme had had no effect.

One way in which a difference might have arisen would have been if the TVEI group were a less academic group. This would have resulted in a pseudo-effect, because the TVEI and non-TVEI pupils would have regressed to different means. This situation is shown in diagram 2 where it is assumed that an outcome Y is regressed against a predictor X for two groups which were initially not equivalent and for which there was no actual effect in terms of outcome Y. The predictor is a proxy for the 'real' determinants and is measured with error
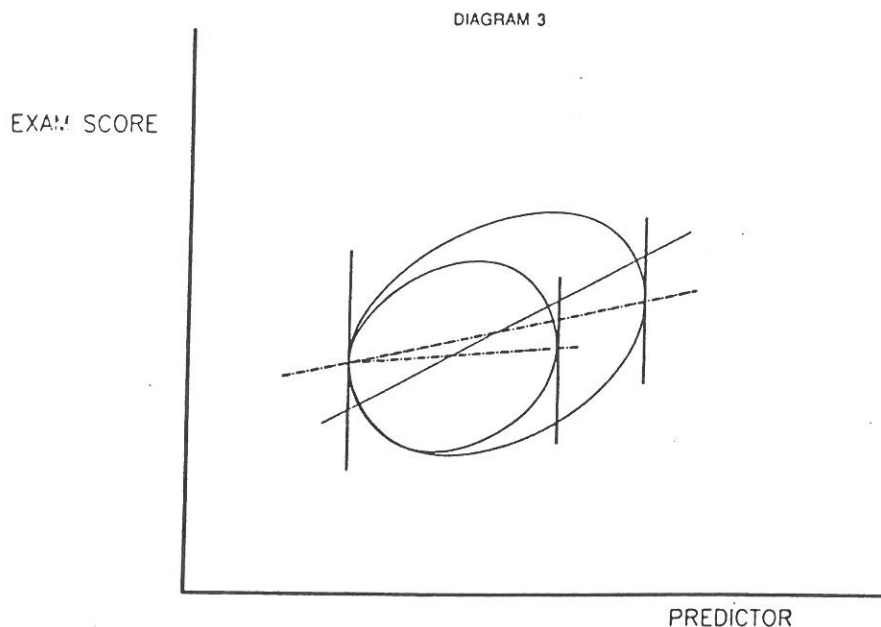

DIAGRAM 2

EXAM SCORE

PREDICTOR
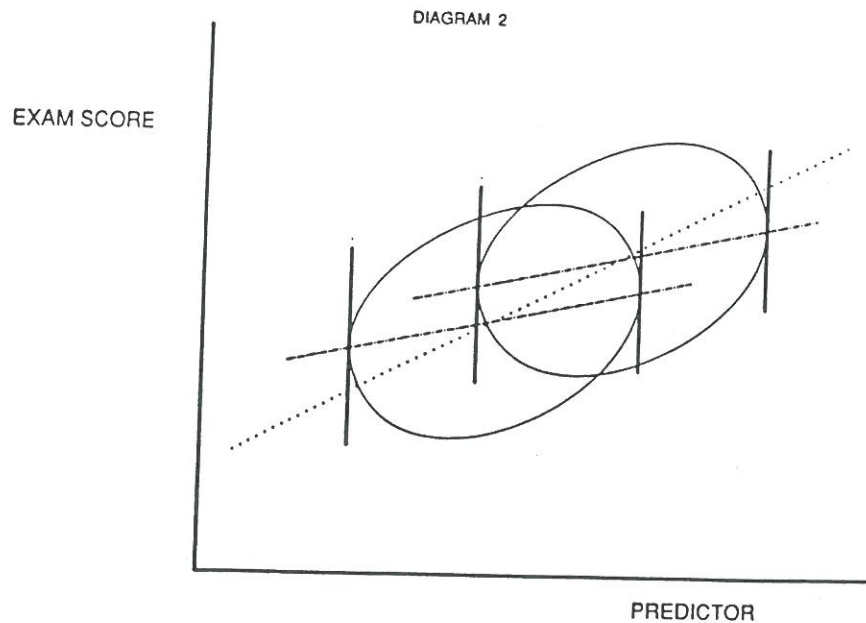
Diag. 2

In fact the mean Raven score for TVEI pupils was not significantly different from the non-TVEI pupils although the standard deviation was smaller. However the TVEI pupils had less intention of remaining in education after completing their fifth year work.

One might therefore expect that, if the difference between TVEI and non-TVEI pupils was a pseudo-effect, then it would be associated with a failure to account for factors associated with LSE rather than with the Raven score. The LSE scale was distributed approximately normally for the whole sample but it was more positively skewed for the TVEI pupils and had a smaller standard deviation (see Table 1). Furthermore, although the mean for the TVEI pupils was less than for the non-TVEI pupils the lowest scores for the two groups coincided. In diagrammatic form the situation is shown in figure 3 where the bivariate distributions for the two non-equivalent groups are again indicated by ellipses. The predictor has a lower mean for one group than the other but the minimum value for both groups coincide.

Table 1.

	Raven		LSE		Skewness
	Mean	SD	Mean	SD	
Non-TVEI	45.2	7.1	2.70	0.84	.42
TVEI	45.0	6.3*	2.54*	0.74*	.55

* Difference between non-TVEI and TVEI significant at < .01 level.

DIAGRAM 3

EXAM SCORE

PREDICTOR

Diag. 3

In order to compare this diagram with the TVEI data the following equation was used:

$$Y_j = \beta_0 + \beta_1 Raven_j + \beta_2 LSE_j + \beta_3 TVEI_j + \beta_4 TVEI_j\ Raven_j + \beta_5 TVEI_j LSE_j + e_j \quad (2)$$

$TVEI_j$ was a dummy variable coded 0 for non-TVEI pupils and 1 for TVEI pupils. The controlling variables, Raven and LSE, were centred around their respective grand means. The results of this analysis are presented in Table 2 and graphically in diagrams 4a-4d.

The similarity between diagrams 4a and 3 suggest that the discrepancy between the residual gains of TVEI and non-TVEI pupils for the Exam Mean could be due to a pseudo-effect originating in the lack of equivalence of the two groups. Although, of course, it remains possible that TVEI had the most negative effect on the pupils with the highest educational aspirations.

Table 2.

	EXAM MEAN+		EXAM TOTAL+	
	B	(SE)	B	(SE)
CONSTANT	0.049	(.020)	0.80	(.2)
RAVEN+	0.084	(.003)	0.94	(.03)
LSE+	0.65	(.03)	7.3	(.3)
TVEI	−0.25	(.04)	−4.0	(.5)
TVEI X RAVENS	−0.0007	(.007)	−0.22	(.08)
TVEI X LSE	−0.19	(.06)	−1.9	(.6)

+ Variables centred around grand mean.

Under the hypothesis that the TVEI pilot had no influence on the average exam grades the situation is as follows: The TVEI and non-TVEI pupils were of equivalent ability but the former were more inclined to leave school as soon as possible and so less inclined towards examination success. This was partially taken into account by the LSE scale but not sufficiently to explain the differences. Because the LSE scale had a floor, which TVEI pupils could not go below, their responses were positively skewed; in other words at the lower end of the scale their LSE values were being overestimated and so the pseudo-effect disappeared to zero.

Such an effect is presumably present in the Exam Total analysis, but the additional significant TVEI/Raven interaction could imply that the TVEI programme did have a negative influence on the Exam Total score perhaps by having led to a reduction in the number of subjects entered by more able TVEI pupils.
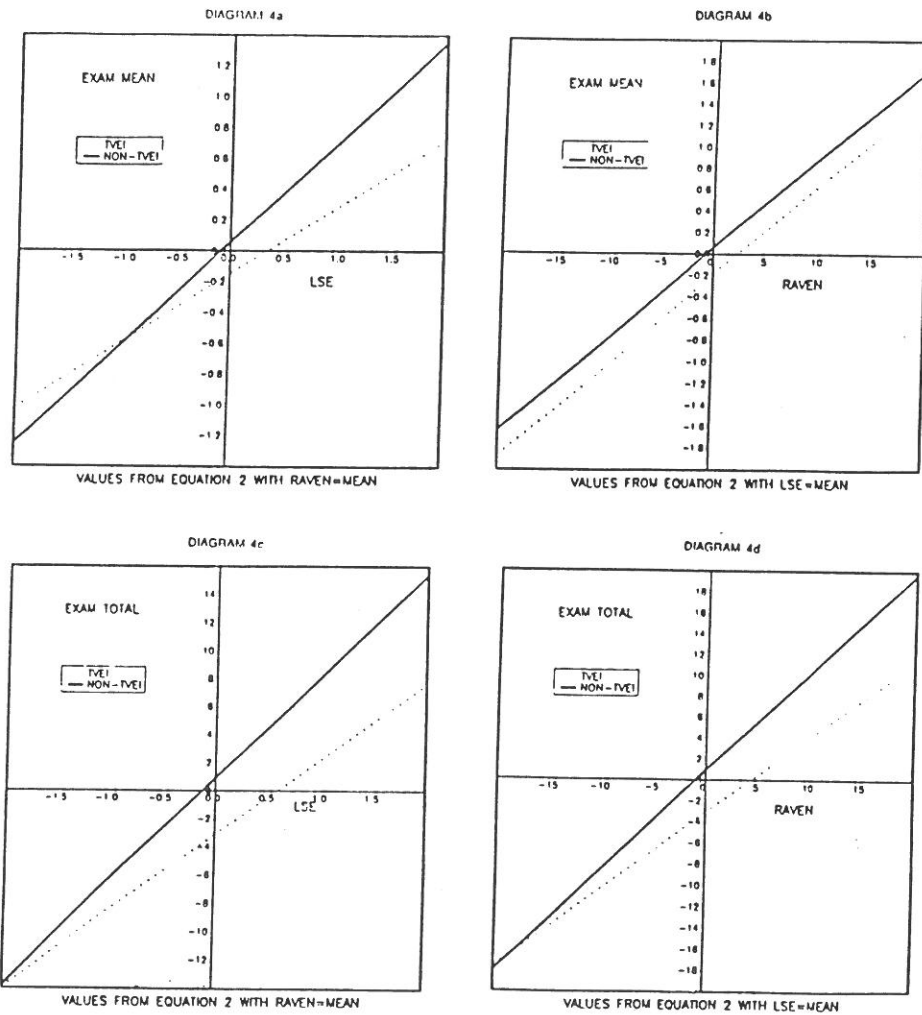
Diag. 4a, 4b, 4c and 4d.

CAUSE AND EFFECT

A clue to the re-examination of the data came from the TVEI coordinator of the project in one LEA. He said that the initiative could not have been the cause of the TVEI cohorts apparent underachievement in his area because no distinction had been made between TVEI and non-TVEI pupils. If some schools or LEAs had distinguished between TVEI and non-TVEI pupils in their treatment and some had not, then the apparent impact of the initiative could be related to the differing treatment between schools.

(The claim that no distinction had been made between TVEI and non-TVEI pupils is of interest in itself since the Manpower Services Commission had in the early years insisted on special treatment being given to the TVEI cohorts and had, on at least one occasion known to the authors, threatened to stop funding if special treatment was not given.)

So, how exactly could the data be analysed to tease out answers to the cause/ effect questions? In the first place it would not be entirely satisfactory to use Ordinary Least Squares (OLS) multiple regression for the whole sample in view of the limitations and problems associated with its use. (Aitkin and Longford, 1986; Goldstein, 1987; Cuttance, 1985; Raudenbush and Bryk, 1986), Nor would it be desirable to use separate analyses of each school using OLS multiple regression to calculate separate school effects and then use those effects since this would not provide the most efficient estimates of each school's effect. Therefore, the Hierachical Linear Models of Bryk, Raudenbush, Seltzer and Congdon (1986) as well as Rasbash, Prosser and Goldstein (1988) were used. These programs allow for multiple regression equations which vary from school to school but draw strength from the general patterns in the data and permit the coefficients within the equations to be modelled on between school differences.

The basic model employed was:

$$Y_{ij} = \beta_{0i} + \beta_{1i}Raven_{ij} + \beta_{2i}Raven^2_{ij} + \beta_{3i}LSE_{ij} + \beta_{4i}TVEI_{ij} + e_{ij} \tag{3}$$

The investigation was simplified by fixing the coefficients β_1, β_2 and β_3 but allowing β_0 and β_4 to vary between schools. All of the variables were centred about their grand means.

The between school equations were:

$$\beta_{0i} = \gamma_{00} + u_{0i} \tag{4}$$

$$\beta_{4i} = \gamma_{40} + u_{4i} \tag{5}$$

The results of the analyses of Exam Mean and Exam Total are shown in Table 3.

Table 3.

	EXAM MEAN+		EXAM TOTAL+	
	B	(SE)	B	(SE)
CONSTANT *+	−0.012	(.05)	−0.24	(.5)
RAVEN^2+	0.0018	(.0002)	0.016	(.002)
RAVEN+	−0.074	(.02)	−0.48	(.2)
LSE+	0.57	(.02)	6.4	(.2)
TVEI*+	−0.24	(.07)	−3.9	(.8)

* Coefficients allowed to vary between schools.
+ Variables centred around grand mean.

The variation in β_{4i} across schools could be largely explained by the difference between the average Raven scores of the TVEI and the non-TVEI groups in each school. Scattergrams of the TVEI residuals u_{4i} for each school from the above analyses against the differences between the mean Ravens score of TVEI and non-TVEI pupils within each school are shown in figures 5a and 5b.
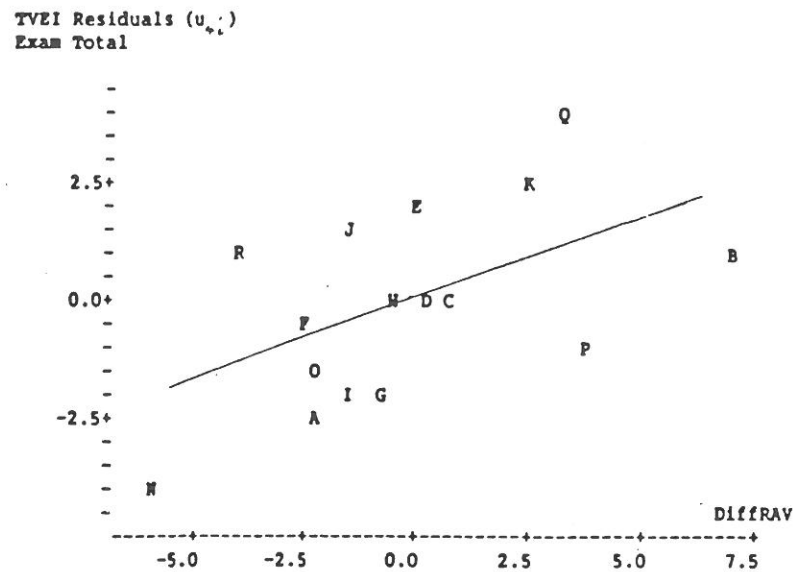
DIAGRAM 5a

TVEI Residuals (u_{4i})
Exam Mean

Diagr. 5a.
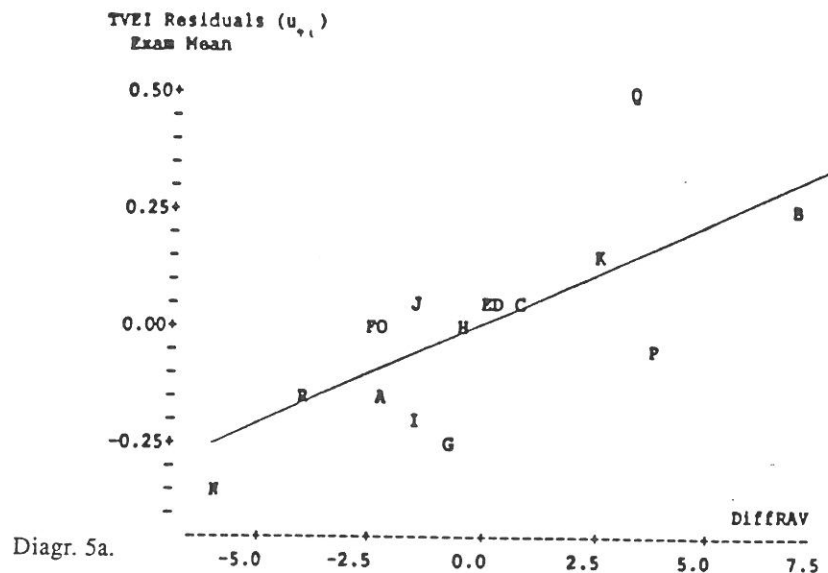
Turning now to the differential implementation of TVEI a serious limitation of the data set must be noted. It included only 21 schools, of which 4 were non-TVEI schools and one was entirely TVEI, leaving 16 institutions where there were both TVEI and non-TVEI pupils. It was these 16 schools where it was hoped to find a difference in the implementation of the initiative.

The data from the schools was collected in three stages. Firstly pupils completed a questionnaire and were administered a Raven test at the start of their fourth year. Then, during the fifth year, they were given a follow-up questionnaire and finally the schools provided the examination results after the pupils had left school. The follow-up questionnaire asked questions about the amount of work experience that the pupils had been involved in and several questions were also asked about the kind of lessons which the pupils had attended. From these ques-

DIAGRAM 5b

TVEI Residuals (u_{+i})
Exam Total

```
        -
        -                                              Q
        -
        -
   2.5+                                   K
        -                          E
        -                  J
        -          R                                            B
        -
   0.0+                        N  D C
        -          P
        -
        -                                         P
        -            O
        -                I  G
  -2.5+            A
        -
        -
        -  N
        ~                                              DiffRAV
        --------+----------+----------+----------+----------+----------+
         -5.0      -2.5       0.0       2.5       5.0       7.5
```

Diag. 5b

tions it was possible to draw up two measures of the degree of implementation of the TVEI initiative within each school. The first was simply the difference between the average amount of work experience reported by TVEI and by non-TVEI pupils. The second involved forming a scale reflecting the amount of Student Centred Learning (SCL) which the pupils reported that they had experienced (Hazelwood, Fitz-Gibbon and McCabe 1988). The difference between the average score for TVEI and non-TVEI pupils was then calculated for each school. TVEI had encouraged a more pupil centred approach and more work experience and so positive scores on the two measures of implementation should represent a greater differentiation of the TVEI philosophy for a particular cohort within each school. This data is presented in Table 4. The number of schools on which there was Work Experience and Student Centred Learning data was 14.

The correlation between the two measures of implementation was .09. This low figure could indicate that the measures did not adequately represent the degree of implementation of the policy, but it would seem more likely that it reflects the differing ways in which TVEI was interpreted within the schools.

The question to be asked now is; did those schools which implemented TVEI policy differentially for their TVEI cohort attain corresponding differential exam success?

Table 4.

SCHOOL	DiffRAV	DiffWE	DiffSCL
A	−2.28	−.07	1.48
B	7.04	.81	1.09
C	0.78	−.14	1.14
D	0.23	.55	.68
E	0.09	−.05	.69
F	−2.55	−.13	1.13
G	−0.64	.47	.89
H	−0.52	−.18	.32
I	−1.62	.04	−.28
J	−1.41	*	*
K	2.40	.25	.65
L	*	*	*
M	*	*	*
N	−6.10	*	*
O	−2.33	1.23	.88
P	3.63	−.13	1.01
Q	3.26	.24	−.21
R	−3.97	−.05	.15
S	*	*	*
T	*	*	*
U	*	*	*

* Missing data or not applicable.
DiffRAV = Mean Raven score of TVEI pupils minus non-TVEI pupils.
DiffWE = Mean reported Work Experience of TVEI pupils minus non-TVEI pupils.
DiffSCL = Mean reported amount of Student Centred Learning of TVEI pupils minus non-TVEI pupils.

This may be partially answered by modelling β_{4i} using the measures of differentiation in the equation:

$$\beta_{4i} = \gamma_{40} + \gamma_{41}DiffRav_i + \gamma_{42}DiffWE_i + \gamma_{43}DiffSCL_i + u_{4i} \qquad (6)$$

Where: DiffRAV is the difference between the mean Raven score for TVEI pupils and the mean Raven score for non-TVEI pupils.
DiffWE is the difference between the amount of work experience reported by TVEI and non-TVEI pupils.
DiffSCL is the difference between the reported amount of Student Centred Learning experienced by TVEI and non-TVEI pupils.

The results of this analysis appear in Table 5. It was not significant to add either of the measures of differentiation to the equation when Exam Mean was being modelled. But, the DiffSCL measure did contribute to the modelling of Exam Total negatively.

Table 5.

	EXAM MEAN+		EXAM TOTAL+	
	B	(SE)	B	(SE)
Pupil Level				
CONSTANT*	−0.013	(.07)	−0.061	(0.7)
RAVEN^2+	0.00072	(.0002)	0.0051	(.002)
RAVEN+	0.019	(.02)	0.43	(.17)
LSE+	0.62	(.03)	6.7	(.3)
TVEI*+	−0.22	(.1)	−2.8	(1.0)
Modelling TVEI at School Level				
DiffRav+	−0.061	(.02)	0.39	(.21)
DiffWE+	0.076	(.14)	0.044	(1.4)
DiffSCL+	−0.092	(.12)	−2.3	(1.3)

* Coefficients allowed to vary between schools.
\+ Variables centred around the grand mean.

In order to obtain a graphic representation of the analysis and to keep as many schools in mind as possible the results from Table 3 (equation 3, 4 and 5) were used to produce diagrams 5a and 5b which show the fairly strong relationship between the TVEI residuals and the ability differential. There is no clear relationship between the points on the scattergram in 5a and the measures of implementation in Table 4, however, in Diagram 5b schools N, A, I, G, P. and B clearly lie below the regression line and R, J, E, K and Q above it. Those above the line generally have a low score on DiffSCL and those below it a high score. Visually there would seem to be some evidence that TVEI implementation was associated with poorer Exam Total scores.

One final analysis was considered. If the TVEI pilot had not influenced the Exam Mean scores but had influenced the Exam Total scores, then this would show up most clearly when predicting Exam Totals from Exam Means (r=.91) as in the equations:

$$ExamTotal_{ij} = \beta_{0i} + \beta_{1i} ExamMean_{ij} + \beta_{2i} TVEI_{ij} + e_{ij} \qquad (7)$$

$$\beta_{2i} = \gamma_{20} + \gamma_{21} DiffRav_i + \gamma_{22} DiffWE_i + \gamma_{23} DiffSCL_i + u_{2i} \qquad (8)$$

The results appear in Table 6.

Table 6.

	EXAM TOTAL+	
	B	(SE)
Pupil Level		
CONSTANT*	0.033	(0.7)
EXAM MEAN*+	9.98	(.3)
TVEI*+	−0.57	(0.6)
Modelling TVEI at School Level		
DiffWE+	−0.16	(.87)
DiffSCL+	−1.50	(.77)

* Coefficients allowed to vary between schools.
+ Variables centred around the grand mean.

Once again it would appear that the emphasis on Student Centred Learning for TVEI pupils compared with non-TVEI pupils within schools was associated with poorer Exam Total scores.

(NB1 The analysis of the Exam Total score using Exam Mean as a predictor would seem to be more efficient than to look at the number of exam entries using Raven and LSE as predictors because of the very high correlation between the two exam measures.)

(NB2 Following Raffe (1988) the data were examined using dummies for the different projects (consortia) to explain β_{4i} but the results were not statistically significant, although from an evaluator's perspective there were very clear differences in almost every aspect of TVEI.)

DISCUSSION

The interpretation of the data presented above is subject to all the constrictions associated with survey data and it is particularly important to bear in mind that the data refer to a single year and that variations across time and geographical location are unknown. Nevertheless, with these provisos in mind a number of points can be made.

In the first place the TVEI initiative would seem to have had little effect on the average exam grades of pupils involved in it. Their grades were some fifth of a grade lower than expected, on the basis of regression analysis, but this could well be because those pupils who opted for TVEI tended to be less educationally oriented than those who did not. The way in which this shortfall could have

arisen was outlined in the first section of this paper, but the most important evidence comes from the second section which failed to find a relationship between the shortfall and the degree of implementation of TVEI (Diag. 5a and Table 5). On the other hand the Exam Total score may well have been negatively influenced by membership of the TVEI scheme. This was particularly true of the more able pupils (Dig. 4d) and of those schools which implemented TVEI most extensively in terms of Student Centred Learning. However, to interpret the data as showing that Student Centred Learning caused poor Exam Totals would seem to be wrong in this context since the evidence is against an effect on average Exam grade and surely, if such a teaching methodology influenced exam results it would show itself directly in the exam grades. The DiffSCL measure is aggregated and therefore more realistically interpreted as a measure of how far the 'TVEI philosophy' was prevalent than as a measure of the amount of Student Centred Learning per pupil.

The second observation concerns the within school contextual effect. There has been much evidence to show that schools with more able intakes tend to get better results than expected (e.g. Willms 1988, Willms 1986, Rutter, Maughan, Ouston and Smith 1979) but exactly how the effect arises is not clear. Willms (1986) suggests that "to some extent they are associated with factors that lie outside teachers' control, such as peer influences or the effects of local community 'deprivation'". On the other hand Hauser (1970) has demonstrated that contextual effects can arise artifactually. Diag. 5a and Table 5 point to a within school contextual effect where the relatively less able TVEI cohort in a school perform less well than the rest of the school having controlled for ability and aspirational levels. The effect is not so clear for the Exam Total score as it is for the Exam mean. Once again the difference between these scores could be seen as resulting from school policy differences. In which case it could be that there is an underlying contextual effect but that within school exam entry policy could override its influence when the number of entries is considered.

This apparent effect has important implications for school organisation and is something which should ideally be followed up with experimentation. The Hauser argument is an important one and although there were good controls in the data (Raven's SPM and LSE) it is never possible to control for all variables and in any case those that are measured are measured with error.

EVALUATION

This paper has considered the exam results of TVEI and non-TVEI pupils in the database established at Newcastle during the evaluation of the pilot TVEI schemes in the North East of England during the period 1985-87. The pupils took national exams in the summer of 1987, when they were about 16 years old. Their average grades (Exam Means) and their accumulated grades (Exam Totals) were considered separately.

The Exam Mean and Exam Total scores were generally worse for TVEI than non-TVEI pupils even when pupils with the same ability test scores (Raven's Standard Progressive Matrices) and aspirational level (Likelihood of Staying in Education) were compared. These findings were reported in Fitz-Gibbon et al. (1988).

However, as was noted in that article, it is possible that factors other than TVEI itself caused the discrepancy between the TVEI and non-TVEI pupils and it is those factors and the search for other explanations which has been the subject of this paper.

Firstly, the possibility that the TVEI/non-TVEI difference was an artefact of regression analysis was considered. Such a discrepancy can appear when the means scores on the predictor, for two groups, differ. Now, the two groups were similar on their ability scores, but the mean aspirational level of the TVEI pupils was lower than that of their non-TVEI peers. Furthermore, the aspirational level had proved to be a good predictor of examination success. It was argued, using a graphical approach, that the data were consistent with there having been no TVEI effect on the Exam Mean score but a negative TVEI effect on the Exam Total.

Secondly, the data were considered using multilevel models to see if associations could be found between the TVEI/non-TVEI within school differences, on the one hand, and three other measures of TVEI/non-TVEI within school differences on the other. The idea behind this was that schools varied considerably in the degree to which they implemented TVEI policy and if there were a TVEI effect it would appear most clearly in those schools where there was a differentiated cohort of TVEI pupils to whom the TVEI philosophy was applied. Two variables were available in the database which measured the implementation of TVEI policy; one related to the amount of work experience and the other to the degree of Student Centred Learning. Both measures were reported by pupils. For each school, an aggregated measure of the difference between the TVEI and non-TVEI on the two measures of implementation was calculated (DiffWE and DiffSCL respectively). A third school level measure was considered and that was the difference between the mean ability scores of the two groups in each school (DiffRAV).

The TVEI/non-TVEI within school Exam Mean discrepancy was strongly associated with the DiffRAV variable but it bore little relationship to the measures of TVEI policy implementation. By contrast, the Exam Total score was less associated with the DiffRAV variable but significantly related to DiffSCL.

Once again the results are consistent with there having been no TVEI effect on Exam Mean but a negative TVEI effect on Exam Total.

Schools in England and Wales, at this time, were able to operate their own polices with respect to subject options and exam entries, and whilst most schools entered most pupils for several subjects there was some variation between schools, as regards their policy on the number of subjects which a pupil might enter. There were schools which tended to emphasise the importance of many qualifications, whilst others tended to encourage different aspects of the curriculum, so restricting the number of possible passes.

The data would seem to support the view that in some schools TVEI pupils were timetabled for activities other than examination subjects and so their Exam Totals were lower than their non-TVEI peers.

CONCLUSION

The initial analysis of the dataset by Fitz-Gibbon et al. (1988) was accompanied by several warnings about the interpretation of the apparently negative influence of TVEI. In particular it was noted that the initiative was not a true experiment in the Campbell and Stanley (1966) sense, and that the lack of equivalence of the TVEI and non-TVEI pupils meant that any comparison must be viewed tentatively. The present paper has gone some way towards showing that the lack of equivalence was indeed an important factor in making the TVEI pupils seem to perform less well than their peers. The TVEI cohorts generally contained pupils who expected to leave education earlier than the non-TVEI pupils of the same ability and whilst this was picked up in the survey data, it would seem likely that the correction procedures undercorrected and that the analysis therefore showed a negative TVEI effect. The present analysis, whilst being unable to overcome the possible undercorrection has been able to relate the apparent TVEI effect to the degree of implementation of TVEI policy within TVEI schools. This goes some way towards elucidating the causal nature of the relationship.

The best conclusion which can be drawn is that the TVEI initiative had no detectable influence on the average Exam grades of pupils but that there was a negative effect on Exam Total for more able pupils. One possibility being that in concentrating on the new ideas of TVEI, such as Student Centred Learning, schools entered these pupils for fewer examinations.

REFERENCES

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

Bryk, A.S., Raudenbush, S.W., Seltzer, M. and Congdon, R.T. (1986). *An introduction to HLM Computer Program and Users' Guide.*

Campbell, D.T. and Stanley, J.C. (1966). *Experimental and quasi-experimental designs in research.* Chicago: Rand McNally.

Cuttance, P. (1985). Methodological issues in the statistical analysis of data on the effectiveness of schools. *British Educational Research Journal 11* (2: 85), 163-179.

Edwards, T. and Whitty, G. (1989). *The State and Private Education An Evaluation of the Assisted Places Scheme.* The Falmer Press.

Fitz-Gibbon, C.T., Hazelwood, R.D., Tymms, P.B. and McCabe, J.J.C. (1988). Performance Indicators and the TVEI pilot. *Evaluation and Research in Education 2* (3), 49-60.

Goldstein, H. (1987). *Multilevel Models in Educational and Social Research.* London: Griffin.

Gleeson, D. (1987). Ed. *TVEI and secondary education: a critical appraisal.* OU press.

Hazelwood, R.D., Fitz-Gibbon, C.T., and McCabe, J.J.C. (1988). Student Perception of

Teaching and Learning Styles in TVEI. *Evaluation and Research in Education 2* (3), 61-68.

Hauser, R.M. (1970). Context and consex: a cautionary tale. *Amer. J. of Sociology 75,* (January), 645-664.

Manpower Services Commission (undated). *Technical and Vocational Education Initiative Operating Manual.* Sheffield: MSC.

Marks, E. (1982). A note on Geometrical Interpretation of the Correlation Coefficient. *Journal of Educational Statistics 7* (3), 233-237.

Pirie, M.F. (1989). The Current Development of TVEI in Scotland. *Scottish Educational Review 21* (1).

Raffe, D., (1988). Making the Gift Horse Jump the Hurdles. *British Journal of Education and Work 2* (3).

Rasbash, J., Prosser, R., and Goldstein, H. (1988). *ML2 Software for Two-level Analysis.*

Raudenbush, S. and Bryk, A.S. (1986). A hierarchical model for studying school effects. *Sociology of Education 59,* 1-17.

Raven, J.C. (1958). *Standard Progressive Matrices.* London: H.K. Lewis and Company Ltd.

Rutter, M., Maughan, B., Mortimore, P., Ouston, I., and Smith, A. (1979). Fifteen Thousand hours. Somerset: Open Books.

Tenne, R. (1989). *TVEI Students and Studies: 14-16 Students.* The Training Agency, Sheffield.

Willms, J.D. (1988). A longitudinal Hierarchical Linear Model for Estimating School Effects and their Stability. *Journal of Educational Measurement,* in press.

Willms, J.D. (1986). Social Class segregation and its relationship to student's examination results in Scotland. *American Sociological Review 6* (3), 289-306.