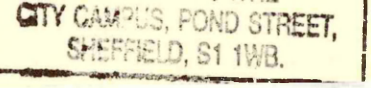


REFERENCE

Value-added as a Performance Indicator of Teaching in Higher Education in the UK

Li Wang

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Doctor of Philosophy

June 2000

ABSTRACT

This research explores the issue of the development of a value-added performance indicator of teaching in higher education in the UK. The empirical work of this research is based on the data on university student entry qualifications and degree results from two new universities and the interviews with 18 members of academic staff from a new university.

This research contains two main parts both aimed at the question of the feasibility of constructing and using value-added as a performance indicator. The first part of this research developed a method of calculating value-added in higher education in the UK, and this method was used to measure value-added at the two institutions at course, school / departmental, and institutional (cohort) level. This was intended to discover the feasibility of developing a method to quantify the relationship between entry qualifications and degree results. The second part of the study used the value-added results obtained from the first part of the study to interview (semi-structured interview) academic staff to investigate their views on whether these value-added results can be used to indicate quality of teaching. In this case the question addressed was the perceived utility of the measure and its acceptability.

The first part of the study found that the method of calculating value-added developed can be used to identify variations in value-added at course (course with large number of students), school / departmental, and institutional (cohort) level, and this method has advantages over the main existing value-added measurements, Index methods and the Comparative method. The second part of the study suggested that academics found it acceptable if the value-added results were used to identify problems, nevertheless, they argued that the value-added results can not directly indicate quality of teaching. The difficulties with directly using value-added results to indicate quality of teaching are summarised into the following aspects: the concept of value-added, comparability, factors which have impact on student academic achievements, factors which have impact on how accurately degree results can reflect students' true achievements.

It is concluded that the method of calculating value-added developed in this research can be used to identify problems in higher education in the UK, but it can not be used directly to indicate quality of teaching. The findings of this research imply that a value-added performance indicator would derive its significance from the link between value-added results and specific processes of teaching and learning.

PREFACE

Before coming to the UK, I had worked at the Peking Normal University in China for five years. During the five year period, I worked on how to assure and to enhance quality of teaching at the University. I was also involved in a number of research projects on teaching evaluation in higher education in China. My MEd study in the UK gave me the opportunity to observe the quality of teaching and the management of quality of teaching of an UK university. Of course, I also learned more theory about teaching evaluation through my MEd. All these experiences, time and time again, made me feel the importance and difficulties of measuring the quality of teaching. Therefore, when I was offered a research studentship and was asked to do the research project on value added as a performance indicator in higher education in the UK as my Ph.D. study, I immediately took interest in this project.

ACKNOWLEDGMENTS

Thanks are due to:

Professor Peter Ashworth - for providing continual guidance and encouragement throughout the whole project. And to whom I am deeply indebted

Mr. Mike Grimsley - for all his guidance, particularly in statistics. Without his guidance, this project would not have been possible.

Dr. Neil Challis - for all his guidance and also for giving me many kinds of support, polishing my English was an obvious one. And to whom I am most grateful.

Mr. John Brennan and Dr. Philip McGeevor - for their interest in the project, and their valuable comments on the project and the thesis. I deeply appreciate it.

The registrars at two institutions who allowed me to use their data and the administrative staff at the two institutions who downloaded the data from their database for me. To the members of the staff who so generously gave their valuable time to be interviewed and so openly shared their experiences. Without their help, this project will not be completed.

The members of staff in LTI, particularly **David Mowthorpe**, **Maurice Teasdale, Meg Handscombe**, and **Hazel Oliver**. To other research students in LTI: **Theresa, Mark, Ian, Shauna**, and **Rachel** - for all help they have given me.

My family and my friend, **Miss Elizabeth Adams** - for their great support and encouragement.

CONTENTS

Chapter Four: Measuring Value-added at Course Level

Chapter Five: Measuring Value-added at School and Institutional Level

Chapter Six: Value-added Results and Quality of Teaching

Chapter Seven: Conclusion

CHAPTER ONE INTRODUCTION

The national economic crisis of 1974/75 led to the imposition of financial constraints on higher education. The British Government determined to make the higher education sector more accountable to the taxpayers. Thus, performance indicators were externally imposed into higher education policy in the UK.

The framework of performance indicators is based on production theory in economics. Degree results as a most obvious outcome of teaching activity are proposed as a performance indicator of teaching. However, critics quickly recognized the limitations of comparison of degree results that failed to take entry qualifications into account. It was argued that some institutions consistently attract better-qualified entrants, therefore some institutions had higher achievement 'because of what happens in the admission office rather than what happens in the classroom.' Thus the concept of value added was introduced as a possible performance indicator in higher education.

The idea of value added is attractive. However, to quantify the relationship between inputs and outputs is a formidable problem. The measurements proposed to measure value added (e.g. the index methods and the comparative method discussed below) and the suggestions of using value added results as a performance indicator of teaching have met with considerable criticism. As a result, value added was dropped off the list of performance indicators published by the Joint Performance Indicator Working Group (JPIWG) in 1995 on the grounds that there is a need to do more research on them. Cave *et al* (1997, p135) concluded that

> *Research in this area is still in its infancy and by no means at a stage where we can say that value added measures can or cannot be made operational at some level*

The present research attempts to make contributions towards this debate.
The aim of the thesis is to explore the feasibility of constructing and using a value-added performance indicator of teaching. The study has explored this issue from two different aspects. The first part of the study has explored the issue from the aspect of technique and has focused on measuring value added. The second part has stepped back to critically assess problems with using the value added results obtained from the first part of the study to indicate quality of teaching.

The remainder of the thesis is presented as follows.

Chapter 2 Literature review

The literature review is divided into three sections. The first section discusses why and how performance indicators were introduced into higher education policy in the UK. Next it introduces that the framework of performance indicators in higher education is based on production theory in economics. However, when this production theory is applied to higher education sector, it meets the difficulties with measuring input and output of higher education. It is revealed that because of these difficulties, measuring the output of universities is therefore an extremely complex problem. The arguments about definition of performance indicators and the intended use of performance indicators in higher education are also discussed in this section.

The second section focuses on discussing teaching performance indicators. Universities are essentially funded as teaching and research institutions, yet research is given much more status and priority. Therefore there is an increased pressure for evaluating teaching in UK and elsewhere. The development of main teaching performance indicators is discussed.

The literature on value added as a performance indicator of teaching is extensively reviewed in the third section. How the concept of value added was introduced into higher education is examined first. Then, main methods of calculating value added in higher education in the UK are discussed and a

value added program in higher education in USA and methods of calculating value added in school education are also reviewed. Main arguments and difficulties concerned with developing a value added performance indicator in higher education in the UK are summarized. Literature reveals that research in this area is still in its infancy and there is a call for more research. A feature of this research project emerges that this research is an exploratory study, tackling issues of measurement, acceptability, and limitations in the claims about teaching quality that can be made.

Chapter 3 Methodology

This thesis contains two main parts. The first part of the study has designed a method of calculating value added which is based on entry qualifications and degree results. This method is then used to measure value added at course, school and institutional level with actual data in order to examine the feasibility of the method in actual use. The second part of this research has used the value added results obtained from the first part of the study to interview (semi-structured interview) the academic staff, who have taught on these courses or in these schools measured, in order to assess acceptability and perceived legitimacy of the value added method developed as a performance indicator of teaching.

The main feature of this research is that unlike most of the research on performance indicators in higher education most of which were carried out at highly aggregated level (e.g. between institutions), this research is conducted at the lower levels of aggregation (e.g. within an institution), and is a very specific study. Therefore the empirical work of this research provides some insights about developing a value added performance indicator of teaching.

Chapter 4 Measuring value-added at course level

Chapter 4 reports on a study which aims to investigate the strengths and weaknesses of the model developed in the methodology chapter when it is applied to actual data. The study uses small-scale data to measure value added at course level. It is found that although there is some diversity in entry

qualifications, there is a pattern and that students can be placed into groups in such a way as to represent the majority of the population of the institution. Therefore the value added analysis can be based on these major groups rather than the whole population. Another main finding is that the model is only applicable to courses with a large number of students.

Chapter 5 Measuring value-added at school and institutional level

On the basis of the findings of chapter 4, Chapter 5 reports on the use of the model to measure variations in the value added in an institution over the period 1988 - 1993. The measurements are made both at school / departmental level and institutional level. The variations in value added between schools / departments and variations in value added in the institution between different cohorts are identified. These variations in value added raise a series of questions about quality of teaching in the institution during 1988 - 1993. At school level, for example, the A level students from School 2 entered the university with an average A level points 9.24 and about half of them achieved a good degree, while the A level students from school 6 started with an average of 13.61 A level points, but only 39.2 per cent of them obtained a good degree. Why is this so? Does this result indicate that the quality of teaching in school 2 is better than that in school 6?

Comparisons between the model developed in this study and main existing method, the comparative method, are also made. It is found that although the two methods are different, the value added results generated from the two methods are similar. However, conversely, working with the same value added method (i.e. the model developed in this study), but using different methods to measure degree quality (per cent of good degrees as opposed to degree scores) can lead to different value added results.

Chapter 6 Value-added results and quality of teaching

Chapter 6 focuses on investigating whether the value added measurement developed in first part of the study can be used to indicate the quality of

teaching. The qualitative research technique of semi-structured interviewing is used. This study emphasizes the value of academic staff views in assessing a performance indicator of teaching, because it is academic staff who deliver the teaching and who are ultimately responsible for degree classification decisions, and therefore they are most knowledgeable about whether a performance indicator can reflect the reality of performance of teaching. It also indicates whether such a performance indicator would be regarded favourably by teaching staff.

The academic staff concerns about using the value added measurement as a performance indicator of teaching may be summarized into the following aspects:

Firstly, the academic staff argued that degree results could not reflect all the qualities gained by the students through teaching, and there is some value which can not be classified into an upper second or lower second (degree class).

Secondly, they were concerned that the diversity and flexibility which exist under the current higher system in the UK, make it seem impossible to 'compare like with like' even at course level. They argued that even students who have the same entry qualifications and study on the same degree course should not be treated the same when we measure value added, because they may take different routes to study.

Thirdly, the academic staff argued that there was not a one to one relationship between the quality of teaching and what students achieve, because there were other factors that come in to it. The factors suggested by the academic staff may be summarized into the following categories: motivation, the 'aura' of the class, some attainments which are not measured by entry qualifications, domestic situations of the students, increasing amount of time of working in term time, social class, and management of teaching.

Fourthly, the academic staff suggested that degree classifications did not necessarily indicate how good the students were or how good the teaching was. There are factors that interact to influence degree classifications. The factors they suggested are summarized into following categories: 'academic community', 'the way the examination boards behaved', 'attitudes of external examiners', 'how we assessed students', 'standard of marking', and 'coaching students for examinations'. Therefore when a figure / score shows the improvement of degree results, it could be a product of the improvement of teaching or a combination of the improvement of teaching with these factors.

Chapter 7 Conclusion

There is evidence that indicates that the method of calculating value added developed in this research can be used to identify problems of teaching and the value added results should be treated as a symptom which needs to be further investigated. However, it would not be acceptable using the value-added results to directly indicate quality of teaching.

The main implications of the findings of this research for the development and use of a value added indicator in higher education are as follows. Firstly, this research suggests that development of a value-added performance indicator needs well defined objectives first. A value-added performance indicator should measure the achievement against the objectives in terms of value-added. Secondly, it is suggested that the significance of the value added results would derive from the link between them and the process of teaching and learning.

CHAPTER TWO: LITERATURE REVIEW

This chapter reviews literature on performance indicators in higher education and in particular, literature on value-added as a performance indicator in higher education. The aim of the survey of literature is to define the research problem.

The first section, Background, discusses general issues about performance indicators in higher education. The second section then focuses on discussing teaching performance indicators. Finally, the literature on value-added as a performance indicator of teaching is extensively reviewed in the third section.

The literature reveals that research in the area of value-added as a performance indicator of teaching in higher education is still in its infancy and there is a call for more research. The research reported in this thesis is an exploratory study, aimed at clarifying, and suggesting solutions (where these exist), and the matters which threaten the feasibility of this performance indicator.

2.1 Background

2.1.1 Introduction of performance indicators into higher education in UK

Performance indicators were externally imposed into higher education in the UK by government over a decade ago. This section will briefly review why and how performance indicators were introduced. It will discuss the social and economic background of the introduction of performance indicators, the link between the university funding system and introduction of performance indicators, and link between the binary system of higher education and introduction of performance indicators.

2.1.1.1 The social and economic background

The national economic crisis of 1974/75 led to the imposition of financial constraints across the public sector - higher education, health service, local government and police. The financial difficulties were coupled with a distrust by

politicians of self-governing professional groups (such as university teachers) and a desire to increase the control exercised over them. Thus the British government determined to make the public sector more accountable to the taxpayer and emphasise efficiency, effectiveness, value for money, and accountability.

Government policy towards higher education has changed rapidly. Higher education has been expected to make a greater contribution to national economic development. Institutions can no longer expect a continuation of funding regardless of past performance.

The pressure for fundamental improvements in the contribution of the higher education sector to national economic development led directly to the setting up of the Jarratt Committee by the Committee of Vice-Chancellors and Principals (CVCP). The work of the Committee involved an investigation of the efficiency and effectiveness of universities. The Jarratt report made some far-reaching recommendations about the governance and management of universities. One of the recommendations was that

> *A range of performance indicators should be developed, covering both inputs and outputs and designed for use both within individual institutions and for making comparisons between institutions. (Jarratt, 1985, p36)*

The Department of Education and Science (DES) warmly welcomed the Jarratt Report's suggestions that a range of performance indicators should be constructed:

> *The Government believes there would be advantage in the regular publication of a range of unit cost and other performance indicators by institution and by department. It therefore welcomes the Jarratt Report's suggestions for developing reliable and consistent performance indicators designed for use both within individual universities and for making comparisons between them.* (DES, 1985, p31)

The DES also discussed how the performance of institutions should be measured. The Green Paper suggested that the Government would be interested in three main outcomes of higher education - highly qualified manpower; research; and other social benefits - and that it would wish these to be produced with proper regard for value for money. The performance indicators discussed in the Green Paper included: student numbers and participation rates; unit costs; recurrent costs; the number and costs of successful students.

To respond to the recommendation of the Jarratt report, the University Grants Committee (UGC) and CVCP set up a joint Working Group. The Working Group produced three broad types of indicators: input indicator, process indicators and output indicators. It also suggested that it should publish a range of quantitative indicators to 'assist universities in the running of their affairs'.

2.1.1. 2 The university funding and introduction of performance indicators

In the UK the university sector has been publicly funded since 1889. However, as noted earlier, because of the economic crisis of 1974/5, financial constraints were imposed on universities. The exchequer funding of universities fell dramatically. The university sector lessened its dependence on government funding; more emphasis was given to overseas student fees, industry, and research foundations. The system of funding was changed from student-led, in which student numbers determined the total amount required by the universities, to finance-constrained, in which this income was regarded as insufficient.

The UGC adopted a highly selective approach when implementing the reductions in recurrent grant for the period 1979/80 to 1983/84. However, it did not explain how the decisions on the proportion of cuts in funding between institutions were made in detail. It reported some factors that were taken into account when the decisions were made and stated that (UGC, 1982, p9)

for the best distribution of resources to the universities there is no substitute for judgement based on experience and repeated review of relevant information.

Therefore the entire evaluation exercise was viewed as inadequate. Many academics felt the criteria used by the UGC needed to be more transparent. It should be ensured that 'future judgements, while certainly selective, would be carried out on the basis of criteria as far as possible known in advance to the university community' (Harris 1986). In response to the criticisms, the UGC promised (UGC, 1984, p43)

We intend to be more open about decision and advice than in the past. In particular we aim to give a rather more detailed explanation of the grants to individual institutions.

The pressures for more detailed explanation made the UGC turn its attention to more systematic performance measurement.

The introduction of performance indicators was also linked to allocation of research funding. In 1984, the UGC stated its intention to move towards a more selective funding of research in order to ensure that resources for research were used to the best advantage. A selectivity exercise was then launched in 1986. The research rankings were taken into account when the total grant was allocated to universities. The total grant for each institution was based on student numbers in each cost centre and each cost centre's research record. However, the ways by which the UGC used to produce the research ranking were not clear and highly subjective. Therefore there was a demand to search for more explicit and objective decision aids (Rogers and Scratcherd, 1986 and Evans and Clift 1987).

2.1.1. 3 The binary system and introduction of performance indicators into higher education in the UK

During the period of introduction of performance indicators, there was a binary line between the universities, and polytechnics and colleges. There were some main differences between the university sector and polytechnic and college sector.

Although the universities receive government grants, the universities in the UK are legally autonomous institutions (DES, 1987a). The university grants are allocated to individual institutions through the University Grants Committee (UGC, which was replaced by the University Funding Council in 1988/89, UFC). Nevertheless, the universities retained their autonomy. How the public money is to be used within the sector is based on the negotiation between the universities and the UGC. The universities can validate their degrees without referring to any outside body. The UGC distributed the financial allocations by using largely informal evaluations, which were carried out by its expert sub-committees, in terms of 'peer review'.

The polytechnics and colleges had substantially less autonomy (Taylor, 1987). They were funded publicly since foundation. Local education authorities were given the duty to allocate their resources (Pratt, 1982). The degrees and qualifications awarded in the polytechnics and colleges needed to be validated by the Council for National Academic Awards (CNAA). The Polytechnics and Colleges Funding Council (PCFC) was created as a parallel body to the UFC in the 1987, and The Polytechnics became gradually more autonomous of the CNAA.

In 1992, the binary line between the universities and polytechnics and colleges was abolished. All polytechnics and colleges that could meet certain criteria became universities. But, in the wake of this liberalisation, all higher education institutions were required to work to stated objectives and to demonstrate that they had met them. A more comprehensive system of quality control on higher education institutions was imposed by the government.

In general, the introduction of performance indicators in UK higher education has been perceived as a threat to entrenched values of autonomy of the institution, the department and the individual. Many academics' attitude towards the usefulness of performance indicators is one of scepticism. Despite this, the widespread and growing use of performance indicators in the public sector in the UK has made it difficult for university authorities to resist for long the introduction of performance indicators.

On the other hand, the pressure of financial constraint also forced individual institutions to accept or develop performance indicators in order to gain a bigger allocation. For the same reason, the higher education system has had to adopt performance indicators: 'If universities wish to receive increased sums of public money- and they must receive more- they must exhibit evidence that what has been received has been well applied' (Page 1987).

2.1.2 The development of performance indicators in higher education in the UK

It is over a decade since performance indicators were externally imposed into higher education in the UK. Here we briefly review their development. A series of working groups have been set up to develop performance indicators in higher education. The review will use major working groups as a line to present development of performance indicators and focus on presenting the establishment of these working groups, their statements, and some comments to their statements. The review shows that development of performance indicators in higher education has been under political pressure, and that there has been a tendency that development of performance indicators places prime emphasis on economy and efficiency. It also reveals the difficulties concerned with developing performance indicators in higher education in the UK.

2.1.2.1 Joint CVCP/UGC Working Group on Performance Indicators

The joint CVCP/UGC Working Group was established in July 1985 in response to the recommendation of the Jarratt Committee. In its Statement, *Performance Indicators in Universities* (CVCP/UGC 1986), it recommended a range of

performance indicators which could be used for the internal management of individual institutions and the evaluation of university performance. The list of performance indictors includes those for which the information was already available and those for which the information needed to be collected.

The Statement claimed that teaching and research were the major focus, but it did not specify performance indicators related to teaching outcome apart from cost and wastage rate which indicate cost saving. As Elton (1987) argued, performance indicators should relate to stated objectives and the objectives for universities are primarily teaching. Nevertheless, the actual focus of the Statement was the expenditure of money, efficiency. This type of managerialist scheme may actually reduce the chances of achieving other kinds of objectives, particularly those concerned with effectiveness, professional development and collegiality (Pollitt, 1987).

The Statement provided the lists of caveats attached to each performance indicator. The lists were useful, but the Statement did not provide additional measures to prevent uncritical use of performance indicators. Therefore the lists will help those of good intent, but they may well have the opposite effect on others (Elton, 1987). In fact, until now, such additional and specific measures to prevent uncritical use of performance indicators have not been developed.

In a second Statement (CVCP/UGC 1987), 39 performance indicators were proposed to be published in the autumn of 1987. The statistics were indeed published under the title *University Management Statistics and Performance Indicators* , an interesting change of emphasis from *Performance Indicators in Universities.*

Appropriate indicators for teaching and research were still not found in the second Statement. Elton (1987) criticised that 'the second Statement gets round the problem of devising performance indicators for teaching and research by almost omitting them and concentrated on what is easily quantifiable.'

Page (1987), a member of the working group's technical committee, accepted that the second Statement was primarily concerned with the data that would be published soon. Therefore emphasis may have been given to certain performance indicators simply because the data were readily available. This was open to the criticism that most performance indicator systems in the public sector are 'data driven' (Carter, 1989). However, Page argued that universities need to demonstrate that they were properly spending the money they had received and should therefore receive increased funding. Therefore 'of course, there was an imperative to publish some useful figures by the end of 1987' (Page 1987).

Performance indicators should be developed on the basis of the specification of objectives. They can only derive their significance when they are deployed within a context where the overall objectives of the system are defined (Pollitt, 1990). However, the second Statement failed to discuss objectives at all.

Nevertheless, the second Statement contained some sound analysis. The stress on trends rather than on a 'snapshot' was important and such analysis was in line with general conceptualisations about the role of performance indicators in public service (Cave et al, 1997).

The warning in the second statement that 'uncritical use of these indicators may seriously damage the health of your university' was welcomed by the academic community and has been frequently quoted.

Since 1987, the technical work of the development of performance indicators has been focused on refinements of what had been presented in the 1987 and 1988 edition, and in developing the new indicators. The list had been an annual publication until 1995. Some changes were made in the early years. For example, indicators for entry qualifications were introduced. The working group did manage to introduce additional indicators that had a greater relevance to teaching and learning. These included:

- Non completion rates
- Degree results
- First destinations of new graduates
- Research quality rating.

The 1990 edition (CVCP/UGC 1990) continued to warn that users of the volume need to consider carefully the inferences that can be properly and usefully drawn from the figures. The various editions included an illustrative commentary on the interpretation of some of the indicators along with the usual list of caveats.

The main problem with the performance indicators system developed by those working groups was that it failed to take consumers into account (Pollitt, 1990). The university system has a variety of 'stakeholders': payers, management, academic researchers and teachers, students, and employers, but few performance indicators developed address their interests. This issue was not directly addressed in the development of performance indicators until 1999.

2.1.2.2 The Morris Committee

Performance indicators for UK higher education were mostly developed initially in the university sector. The development of performance indicators in the former polytechnic sector was also under political pressure. The Secretary of State sent a letter of guidance to the chairman of the Polytechnics and Colleges Funding Council (PCFC) and stated: '*I look to the Council to develop further indicators of both quality and quantity of institutions' teaching and would be grateful if it could consider how these might be used as an input to its funding policies and decisions'* (Morris, 1990, para.1.2). The Morris Committee was subsequently set up by the PCFC .

The Morris Committee drew attention to both macro and institutional performance indicators. On the one hand, it emphasised the function of macro performance indicators to protect the interest of the whole polytechnic sector in public expenditure allocation. These indicators were to be designed to enable

the PCFC 'to illustrate that the sector is well managed, accountable, performance conscious, a good investment and that its claims for resources are credible' .Thirteen macro indicators relating to scale and effectiveness, level of resourcing, efficiency and source of funds were suggested. The PCFC endorsed and refined ten of them. The remaining three, which need further development work, were: **value-added**, quality profiles, and employer / client satisfaction.

On the other hand, the Morris Committee recognised the limitations of performance indicators, and emphasised that institutional performance indicators should be chosen by and relate to the *missions, aims and objectives* of each individual college or polytechnic and reflect their differing priorities.

2.1.2.3 The Joint Performance Indicators Working Group (JPIWG)

With the end of the binary line in UK higher education, the Joint Performance Indicators Working Group (JPIWG) was created in order to conform to the letters of guidance from the Secretaries of State.

The Group stated that the main purpose of performance indicators was ' to inform *institutional managers* about the performance of their institution in its various aspects, and particularly in relation to other institutions' (CVCP, 1995 para1.15). This seems different from the main purpose of performance indicators stated in the White Paper of 1991 where it is stated that 'the *Funding Councils* will be responsible for the development and monitoring of performance indicators in higher education'(DES, 1991).

The JPIWG recognised the importance of developing performance indicators based on the objectives, and also considered the issue of feasibility of developing performance indicators. The JPIWG suggested that the proposed statistics 'should be sufficiently flexible in their construction and application to reflect the wide diversity of institutional missions and objectives' and that the data 'should wherever possible, be available from existing sources' (CVCP, 1995).

In 1994, the JPIWG issued a consultative document in which it listed 88 proposed indicators in five categories: 10 teaching; 14 research; 23 financial health; 4 estate; and 37 macro. This received a generally favourable response from institutions. 'This explicitness about the various categories of indicators was something never achieved in *University Management Statistics and Performance Indicators* and reflects the considerable level of technical development work undertaken by the JPIWG' (Cave, 1997, p64).

It is worth noting that value-added was not included in the list of teaching performance indicators proposed. It is stated that

> *Comparison of inputs and outputs flowing from indicators of students progression / achievement could allow a relative crude value-added measure to be calculated, using entry and exit qualifications. However, we appreciate that this is a complex and potentially awkward area and we recommend that further work should be undertaken to develop a more sophisticated value-added indicator...* (JPIWG, 1995, p17).

Student progression rates were introduced. "Standardised comparators" was proposed to be used to calculate progression rates in order to take into account the factors: entry qualifications, subject mix and gender mix. Nevertheless, it is also noted that

> *In some cases differences may reflect teaching quality. In other cases differences may reflect other factors, including students' individual characteristics and circumstances* (JPIWG, 1995, p13)

The introduction of "Standardised comparators" was based on Johnes' and Taylor's (1990) research. In their research, Johnes and Taylor tested the extent to which inter-university variations in a number of proposed measures of performance can be explained by inter-university variations in other factors. In the light of these findings, they suggested 'Standardised comparators' which

can take some factors into account.

2.1.2.4 Higher Education Management Statistics Group (HEMS)

To take the work forward and publish the data, HEMS was created and took over responsibility in this field from the JPIWG. It took the JPIWG's report as its starting point and reviewed and refined some of the macro statistics. However, it did not intend to produce a successor to the *University Management Statistics and Performance Indicators* volume. Institutional statistics and macro statistics were planned to be published separately.

The HEMS Group (1995) developed the JPIWG indicators, added a few specific measures and proposed the four sets of management statistics for publication in 1996. The four sets of management statistics were macro statistics for teaching; macro finance statistics; institutional statistics for teaching; and institutional financial profiles. The statistics listed are narrower than had been initially proposed by the CVCP/UGC Working Group in 1986.

2.1.2.5 Performance Indicators Steering Group (PISG)

In November 1997, the Secretary of State for Education and Employment asked the HEFCE to develop suitable indicators and benchmarks of performance in the higher education sector. The Performance Indicators Steering Group was established.

The terms of reference of the working group were

> *... to develop appropriate performance indicators and benchmarks which recognise the diversity of the sector and*
>
> *to identify and develop indicators of performance in higher education at both institution and sector levels, which will meet the requirement of the different stakeholders with an interest in such performance (PISG 1999, p3)*

It can be seen, from their terms of reference, that this working group places the emphasis on 'the diversity of the sector' and 'different stakeholders' which is different from previous working groups.

The working group was able to take account of the work of HEMS. Publishing institutional-level indicators for teaching and research is a priority of the working group, and the progress has been made in measuring : participation of under-represented groups, student progression, learning outcomes (including non-completion), efficiency of learning and teaching, student employment, and research output.

It should noted that the indicator, 'participation of under-represented groups' may address the objective of 'widening access' to higher education, but it can only indicate achievement of the objective in terms of quantity, and does not measure the achievement in widening access to higher education in terms of quality. A value -added performance indicator would be a good indicator to measure such an achievement.

As far as technique is concerned, the 'adjusted sector' benchmark approach, which takes some factors into account, has been adopted in preference to more sophisticated modelling techniques. The Group suggested that sophisticated models should not be used directly in constructing performance indicators. Three reasons were given for this suggestion:

> *The indicators should, if at all possible, be understandable by those using them. They should not come from a 'black box'. Secondly, the method, once adopted, should continue to be used so that year on year comparisons can be made. Any attempt at modelling the higher education process will be complex. It is very unlikely to be once and for all operation. Finally, even if a very sophisticated model is developed, the differences between the actual and adjusted outcomes are likely to be due to what the model has missed out as much as to any real differences in the performance of institutions. (*Performance Indicators

Steering Group, 1999, p35)

The above reasons indicate that the development of performance indicators in the UK places more emphasis on the feasibility of use of performance indicators in practice than before. This is consistent with experience in the USA. Ewell and Jones (1994) noted that many promising indicator systems fail simply because they are too expensive, too complex, too time-consuming, or too politically costly to implement. Often the simplest is the best, even if it initially seems less technically attractive.

In December 1999, the higher education sector's first performance indicators were published (HEFCE, 1999). The publication was generally welcomed by the university community such as the Committee of Vice-Chancellors and Principals, and the Association of University Teachers. The performance indicators are seen as a way of demonstrating some of the major achievements of the sector. The Higher Education Funding Council suggested that universities have the autonomy to decide what to do about their relative performance, and the role of the funding council is to provide support to identify those which are performing the best and how they are getting it right. The performance indicators published, reflecting the work of the PISG, include participation of under-represented groups in higher education, drop-out rate, share of research output, and learning outcomes and efficiencies. Indicators of employability are absent (HEFCE, 1999). It is criticised that the calculation of the benchmark figure does not condemn universities with higher entry qualifications and certain subject mixes for recruiting fewer students from lower socio-economic groups (Goddard, Thomson and Wojtas, 1999). A certain bias is implied which value-added might reduce.

2.1.3 The framework of performance indicators in higher education

The framework of performance indicators in higher education is based on production theory in economics, which is also called the 'input - process - output approach' (see Bottrill and Borden, 1994 or Cave et al. , 1997) or the

'inputs - outputs - outcomes' approach (see Richardson, 1994). According to a conceptual production theory, the process of higher education is a production process which transforms inputs (e.g. students, academic and non academic staff, building, equipment, heating, and telephone etc.) into outputs (e.g. graduates, research publications, patents...etc.) and higher education itself is a part of a wider economic and social process.

The production function indicates the maximum possible output which can be obtained by a production unit (e.g. a firm) from a given set of inputs (e.g. raw material, labour input, capital input, technical knowledge, and consumables etc.). Therefore when production theory is applied to the university sector, it would allow an assessment of the efficiency of universities.

However the university sector is not like other industries. There are some problems which arise when this production theory is used in the university sector. We will discuss this in next section.

2.1.4 Measuring output and input of universities

2.1.4.1 Difficulties with measuring output and input of universities

In most commercial activities, output can be measured in monetary units, but this is difficult or impossible in the university sector. Johnes and Taylor (1990) suggested that universities aim to produce four main categories of output: output derived from teaching activities, output derived from research activities, output derived from consultancy and related activities, and cultural and social outputs. However, these outputs are very different and can not be added together in any meaningful ways. Therefore it is necessary to specify the individual outputs a university produces and inputs which these individual outputs are dependent upon. Nevertheless, in the university sector, inputs are often used to produce more than one output, therefore it is very difficult to attribute a specific input to a specific output.

Another difficulty is of separating research and teaching activities. Many activities in universities are the joint product of teaching and research such as

research degrees. Research output can affect teaching output by affecting the inputs which determine teaching output. For example, good research records usually attract students with high academic ability (see Johnes and Taylor, 1990 and Cave, 1997).

A further difficulty is that many outputs and inputs are not quantifiable. For example, the quantity of research publications produced by a university may be a measure of the amount of research output, but it ignores the quality of research output. As far as teaching output is concerned, we can measure the number of graduates, the degree classifications, and the first destination of graduates, but how should all this information be added together to produce a measure of the teaching output of an institution?

Johnes and Taylor (1990) concluded that 'measuring the output of universities is therefore an extremely complex problem.' Nevertheless, higher education can not give up the attempt to measure its output because of the demand for evaluation of universities.

2.1.4.2 The techniques of measurement used for the evaluation of efficiency

Historically, a number of techniques have been used to evaluate efficiency in the public sector. These techniques are regression analysis, cost - benefit analysis, and cost-effectiveness (Cave, et al 1997). All these techniques involve establishing some relationship between inputs and output.

Regression analysis is one of the techniques which has been applied to evaluate performance (see Johnes and Taylor, 1990 and Taylor, 1995). Regression analysis has been used to test hypotheses that certain factors have affected the production process in higher education. For example, a hypothesis may be that the departments which have higher staff-student ratios produce more graduates with good degrees. Each observation of average staff - student ratios and proportion of good degrees can be presented by a point, and using standard techniques of statistical regression, one may fit a line through the

points. The closer the fit of points to the line, the greater the variation in proportion of good degrees 'explained' by the factor, staff - student ratios.

Regression analysis can be used to evaluate the performance of individual departments or institutions by comparing the actual output with expected outputs for a given level of inputs. This may be explained by using the above example. If the above hypothesis is proved to be true, given a level of staff - student ratio of a department, we will be able to predict the expected proportion of students with a good degree. Thus, the differences between actual and expected proportion of students with a good degree can be used as a measure of performance of the department.

The technique of cost - benefit analysis (see Layard and Glaister, 1994, or Weale, 1992) attempts to evaluate the efficiency of higher education by measuring the rate of return on investment in higher education either for the economy as a whole (the social return), or for the individual students. It is seen as one of the most ambitious techniques which has been used to appraise the efficiency of higher education. However, in practice, there are some difficulties with using this approach. For example, it would be difficult to establish the economic benefits associated with research output.

Cost - effectiveness analysis is another technique in which inputs are measured in cost terms and output in physical units (e.g. number of graduates), or both output and input are measured in physical units and are combined to produce productivity measures (e.g. staff - student ratio). In many cases, the cost-effectiveness measure itself is a form of performance indicator. The disadvantage of this approach is that when it uses physical units to measure output, quality of output can not be reflected.

The difficulty with both cost - benefit and cost - effectiveness analysis is that they are based on the assumption that particular costs can be associated with particular returns in higher education. However, in reality, it is very difficult to distinguish the cost of teaching and the cost of research. For example, money

spent on libraries or computers are both teaching and research costs. As noted earlier, it is also difficult to separate output of research and output of teaching.

2.1.5 Definition of performance indicators

Despite the introduction of performance indicators into higher education in the UK for over a decade, there is no single authoritative definition or interpretation of their nature (Cave, 1997).

Laurillard (1980) defined performance indicators in the broadest sense and suggested that the very properties of performance indicators, which make them useful, are that 'they reduce a complexity of subjective judgements to a single objective measure' and 'they are context - free.' Nevertheless, she went on 'these very properties also call into the question the validity of the indicators themselves'. The property of performance indicators, that they reduce complexity, was also pointed out by Frackmann (1987): ' The more complex the reality, the more a need to 'abstract' prevails, if manageability is felt as at all a necessity', therefore ' performance indicators stand for simplified information that is needed for management and organisation.'

In the survey carried out in the mid 1980s under the Organisation for Economic Co-operation and Development (OECD)'s Institutional Management in Higher Education (IMHE) programme, performance indicators were defined as:

> *'numerical values which provide a measurement for assessing the quantitative or qualitative performance of a system and which can be derived in different ways'* (Cuenin, 1987, p6).

Cuenin explained that this definition was deliberately very broad since the international survey covered seventy institutions in fifteen countries in which the concepts of performance indicators used were not identical.

Cuenin then drew a distinction between simple indicators, performance indicators and general indicators. Simple indicators, he suggested, are usually

expressed in the form of absolute figures, and are intended to provide a relatively unbiased description of a situation or process. For example, the number of graduates would be a simple indicator. Compared with performance indicators, simple indicators are more neutral, but a simple indicator may become a performance indicator if a value judgement is involved.

Performance indicators differ from simple indicators in that they imply a point of reference, for example a standard, an objective, an assessment, or a comparator, and are relative rather then absolute in character. Cuenin pointed out that there is ambiguity in some cases, therefore he proposed a general rule to avoid such ambiguity. He suggested that performance indicators should have the following property:

> *when the indicator shows a difference in one direction this means that the situation is better, whereas, if it shows a difference in the opposite direction, then this means that the situation is less favourable. The way in which the data are to be interpreted ought to be obvious* (Cuenin, 1987, p10)

General indicators are those which in the main are derived from outside the situation and are not indicators in the strict sense- they are frequently opinions, survey findings or general statistics. Although they may not conform exactly to the definition of an indicator they are used in decision making (Cuenin, 1987).

The CVCP/UGC (1986) Working Group in their first statement defined performance indicators as ' statements, usually quantified, on resources employed and achievements secured in areas relevant to the particular objectives of the enterprise'. They emphasised that indicators are signals or guides rather than absolute measures and indicators do not necessarily provide direct measurements of inputs, processes and outputs, but they can offer valuable information in relation to them.

As we have see in 1987, the CVCP/GUC Working Group published the

statistics under the title *University Management Statistics and Performance indicators.* Page (1987) claimed that the term performance indicators 'are an abbreviated way of referring to all those numerical data which are useful in managing a university, assessing its operations, costs and performance' therefore this title, *'University Management Statistics'* is better than *Performance indicators.*

Yorke (1991) and Sizer (1992) noted that there was something of a distinction to be drawn between performance indicators and management statistics. Sizer (1992) suggested the relationship between indicators and statistics is that

> *While indicators which are more or less valid expression of objectives formulated by one of the parties involved, statistics merely have to meet the standard of reliability. Statistics must be organised in such a way that a number of indicators can be constructed on their basis, since objectives vary and will therefore produce a variety of indicators. The reliability of the statistical records will 'reflect' on the indicators and without reliable basic information, we are unable to produce reliable indicators.*

Yorke (1991) also pointed out that a management statistic at one level can become a performance indicator at another. The Joint Performance Indicators Working Group (JPIWG,1995, para. 1.8) then, in their statement, suggested that

> *...proposed indicators might be better described as management statistics rather than performance indicators. Institutions could, if they so wished, convert these statistics into performance indicators...*

However, Cave *et al* (1997) comment that 'perhaps the JPIWG has taken things too far with their statement.' They brought various authors' definitions and interpretations of performance indicators together, and defined performance indicators as

> *a measure - usually in quantitative form - of an aspect of the activity of a higher education institution. The measure may be either ordinal or cardinal, absolute or comparative. It thus includes the mechanical applications of formulae... and can inform, and be derived from, such informal and subjective procedures as peer evaluations or reputational rankings (*Cave *et al,* 1997, p24)

This definition looks very ambitious, but 'it seems to run out of steam as it is elaborated' (Yorke, 1996, p1). It mentions that the measure may be ordinal or cardinal, but fails to mention that the measure may also be interval or categorical. Yorke points out that a performance indicator becomes such only where there is a criterion explicitly or implicitly applicable to the activity concerned, and that criteria themselves reflect value-positions regarding activities.

Yorke suggested that it is perhaps wise not to be too ambitious in attempting a definition, and that for the purpose of a particular text, programme quality, in his research.

> *a performance indicator is taken to be a marker of the extent to which a particular purpose is being achieved: the purpose may, of course, be articulated at any of a number of levels ranging from the sectoral to the personal* (Yorke,1996, p1)

He then explained that it is necessary to be able to state what the various purposes are, to articulate the values underlying these purposes, and to know what evidence will be taken as indicative of good and poor performance.

Given the main definitions of performance indicators in the literature, I think that Yorke's definition of performance indicators has defined the properties of performance indicators well, because it addresses the issue of objectives (purposes) and value judgement, which are, I believe, main properties of

performance indicators.

I also think that Cuenin's distinction between simple indicators and performance indicators is very important in defining performance indicators. I believe that performance indicators are made of simple indicators. A simple indicator or a number of simple indicators becomes a performance indicator or performance indicators when one chooses to use it or them as a measure of achievement of an objective. For example, the proportion of good degrees is a simple indicator, but when one chooses it as a measure of quality of teaching, it becomes a performance indicator. Performance indicators always consist of simple indicators. So in this sense, performance indicators are those simple indicators which are chosen to be used as a measure or measures of achievement of an objective.

Therefore performance indicators are both objective and subjective. The objective property is from simple indicators. As Cuenin noted, simple indicators are neutral and more objective. The decision about which simple indicators are used as a measure of achievement is subjective.

2.1.6 The intended use of performance indicators in higher education

Performance indicators may be used at different levels of the higher education system. They can be used to indicate performance of the system as whole, performance of sectors, or performance of individual institutions, departments, programmes, subjects and individuals.

Performance indicators may be used for different purposes. Size, Spee and Bormans (1992) summarise five core uses of performance indicators:

i) Monitoring: the ability to register developments in the system.

ii) Evaluation: the ability to comment on the degree of goal attainment

iii) Dialogue: the improvement of mutual administrative relationships.

iv) Rationalisation: the accomplishment of a coherent policy making process and planning processes.

v) Resource allocation: performance indicators can be used as parameters in the resource allocation model.

Nedwek and Neal (1994) suggest six types of use of performance indicators some of which overlap with the above categories of use. For example, monitoring condition, measuring progress toward specific goals, and allocation decision making. Nevertheless, they also suggest that performance indicators can be used to forecast problems, and diagnose problems and can be used as political symbolism by those politicians who are concerned that they are 'doing something' about post-secondary education.

However, in practice, performance indicator schemes have placed a heavy emphasis on the purpose of resource allocation (purpose v). They were mainly developed to indicate areas of potential cost saving or raise questions about the organisation of resources. The dominant concern appeared to be efficiency. In contrast, those purposes about effectiveness (e.g. purposes ii and iii) appeared in some schemes, but less commonly. This can be seen from 'the development of performance indicators in higher education in the UK', discussed in section 2.1.2.

Pollitt (1987) pointed out that the question 'What is the scheme for?' is often close to the more political one of 'Who is it for?' Most of the schemes in British public service organisations were 'for' politicians and top management. The main concerns of politicians and top management were 'waste' and 'inefficiency'. *'In practice, this strategy has involved tactical moves to weaken the powers of unions and professional bodies...'*(Pollitt, 1987, p88) In contrast, those schemes developed by public service were more concerned with quality and effectiveness than with efficiency. Pollitt also suggested that professional groups themselves should be involved in producing the performance assessment schemes, otherwise, the schemes would have an 'efficiency' character.

However, Cave et al (1997) noted that in the UK a prime concern with the potential of performance indicators to aid funding bodies to allocate resources or to call institutions to account for the efficient use of resources, seems to have given way to concern that individual institutions can effectively assess their own performance and act upon that assessment.

Nevertheless, on the whole, as Banta and Borden (1994) noted, many external agencies responsible for funding higher education seem more interested in performance indicators as a means of demonstrating that colleges and universities are using their resources wisely, whereas faculty and campus administrators are most supportive of outcomes assessment that will help them improve teaching and learning, the overall student experience, and administrative processes. The quality improvement literature suggests that performance indicator systems must move from certifying competence to improving institutional quality (Nedwek and Neal, 1994).

As we noted earlier, in the UK, recently, the higher education sector's first performance indicators were published (HEFCE, 1999). Three reasons were given for publishing these performance indicators: to give information to the public about the performance of universities and colleges; to give managers data to help them manage their institutions; and to ensure public accountability (Bekhradnia, 1999). The reasons may indicate that performance indicators are intended to be used for the purposes of both accountability and improvement.

How should performance indicators be used? Several researchers (e.g. Moravcik, 1986, Jesson and Mayston, 1990, Banta and Borden,1994 and Cave et al , 1997) suggest their standards for the use or development of performance indicators. The most significant one is suggested by Banta and Borden. Drawing on their own experience, Banta and Borden (1994) suggest five standards for developing performance indicators for use within institutions:

a) *Purpose of performance indicators.* A clear purpose is essential to success of a system of performance indicators. How will the indicators be used?

b) *Aligning performance indicators throughout the organisation or system.* Within an institution, constituent colleges or schools should have goals that contribute to the institutional mission, and departments or other units that make up each college or school should likewise have goals that complement those of their schools and of the institution (vertical alignment).

c) *Aligning performance indicators across inputs, processes, and outcomes* we should give more attention to the intervening processes that use resources to produce outcomes. We need to examine carefully the processes that lead to outcomes if we hope to improve them.

d) *Co-ordinating a variety of methods.* Performance indicators should be derived from a variety of co-ordinated methods.

e) *Using performance indicators in decision making.* Performance indicators should be used to inform decision making. They should be used to provide evidence of accountability or direction for improvement, preferably both.

2. 2 Context

2.2.1 The role of teaching in higher education and the need for developing performance indicators of teaching in higher education

Universities were essentially funded as teaching and research institutions, yet research was given much more status and priority. The main reasons for this phenomenon may be summarised into two aspects. The first aspect is concerned with funding and staff promotion. In UK, the allocation of research funding from the funding council is based on research assessment. In contrast, teaching assessment and audit do not carry funding implications. Therefore research activities become the activities which earn grants. In higher education in the UK, performance in research has become increasingly more important than teaching performance in staff promotion and job offer decisions. The second aspect is concerned with the features of teaching and research. It has been the conventional view that research is easier to evaluate than teaching.

The best research advances knowledge (Aitken, 1991). All these created the risk that teaching would receive less attention than research. This would possibly lead to teaching quality decline (Elton, 1987). Therefore there is an increasing pressure to evaluate teaching performance in the UK. In fact, such a pressure exists not only in the UK but also in other countries such as the USA, Australia and New Zealand (Miller, 1986; Boyer, 1989; Moses 1989; Clift *et al* 1989).

Additionally, as noted earlier, in UK and before 1992, there was binary line in higher education. On the one side of the binary line, old universities see teaching as induction of students into disciplinary knowledge and the concepts. They emphasised research and scholarship. On the other side of the binary line, the polytechnics, new universities have a strong teaching culture, and 'research is not a primary concern for the majority of academic staff'(Sizer, 1989). Nevertheless, the abolition of the binary line has made new universities compete for research funding. Their traditional educational culture seems to be undermined. Therefore, there is a need to give more attention to teaching and a need for evaluating teaching in UK.

The 1987 White Paper (DES 1987) suggested that the quality of teaching should be appraised. This White Paper suggested the list of performance indicators which include: non-completion rates; the subsequent employment patterns of students; and students' achievement compared with their entry standards. Nevertheless, as Sizer (1989) pointed out that *to date, publicly available PIs of comparative teaching quality are little more than a desirable objective.* Despite this, some potential performance indicators of teaching have been developed. The next section will discuss main performance indicators of teaching.

2.2.2 Main performance indicators of teaching in higher education in the UK

Value-added is one potential performance indicator of teaching in HE. Before systematically reviewing the literature on value-added as a performance

indicator, it may be worth briefly sketching research on the development of other performance indicators of teaching in higher education in the UK. By doing so, we can see the general technical and practical difficulties with the development of performance indicators of teaching in higher education. Cost measures, student progression rates and exit qualifications, and employment and first destinations are those indicators which are widely discussed and appear regularly in the lists of potential performance indicators of teaching in higher education (e.g. Johnes and Taylor, 1990, Cave et al, 1997, Jarratt ,1985, CVCP, 1987, and CVCP, 1995).

Unit costs / Cost measures

Unit costs are a measure of cost per unit of output. In the university sector, it should be the cost per graduate (output). Nevertheless, in practice, cost per student (input) is more widely used. The two are different when not all students become graduates. Naturally, in principle, a comparison of the average cost per student or graduate between institutions would indicate performance of institutions in terms of efficiency.

However, measuring unit costs is not easy. One simple method of measuring unit costs is the staff-student ratio. One obvious problem with this method is that staff (lecturers) is just one of the inputs (costs) required to transfer students into graduates (output). There are other inputs involved such as equipment, and administrators. One could try to measure other costs, but another difficulty arises in that not all costs are teaching costs, and some expenditures are not directly related to the output of graduates, such as expenditure on academic services. There are also some difficulties with the measurement of output. University output includes undergraduates, taught postgraduates, and research postgraduates. To aggregate these outputs into a single measure is problematical, 'traditionally the weights adopted in the UK were chosen fairly arbitrarily' (Cave et al, 1997). Nevin (1985) suggested an approach to measuring output for calculating cost per unit of output of universities. He argued that staff produce research output and so should be included in output. He combined staff, postgraduates and undergraduates as an index of output by

giving arbitrary weights to staff and students.

Activity based cost (ABC) is a new approach to measuring costs. The basic idea of the ABC method (Rimson 1991) is first, establishing what activities are required to produce particular outputs ; second, establishing what inputs are required for each activity; third, calculating cost of these inputs. Although, the ABC method has been used by some UK universities (Mitchell, 1996), it is still in its infancy (Dehayes and Lovrinic 1994).

Cave et al (1997) point out some difficulties with the interpretation of high cost per student when average cost is used as a measure of comparative efficiency. By one interpretation high unit cost (a high staff-student ratio) may be taken as an indicator of a high quality education process, because high staff-student ratio may be associated with more time allocated to students, and there may be a direct relationship between teaching time and student quality. By the second interpretation high unit cost leads to opposite conclusions. 'If degrees of the same grade are of the same quality, irrespective of the awarding institution, and if the value-added to an individual of obtaining a degree of the same class is the same for all institutions, then average cost may, in certain conditions, be used as an index of efficiency. However, the conditions required for such a conclusion to be valid for inter-institution comparisons are fairly restrictive.'

However, although there are difficulties with using unit cost as a performance indicator, some unit cost data are still included in the list of Higher Education Management Statistics for publication in 1996.

Student progression rates and exit qualifications

The report by the JPIWG in 1995 suggests using student progression rates and exit qualifications as one of the performance indicators of teaching instead of non-completion rates or wastage rates (CVCP, 1995).

Several main factors may have influenced this change. One factor is associated with the interpretation of using wastage rates as a PI. A high wastage rate, on

the one hand, may indicate poor quality of teaching, but on the other hand, maintenance of a high academic standard may mean that a certain level of wastage is unavoidable (Cave 1997). Another factor is associated with the policy of widening access to higher education. The policy of widening access to higher education encourages institutions to lower their admission standards (Baker 1987). Nevertheless, '...as a result of taking more 'higher risk' students, institutions had a higher wastage rate for which they are penalised' (Cave *et al* 1997). The third factor, as the JPIWG stated, the change from wastage rates to students' progression had been strongly influenced by a profile framework. The profile model (Williams 1994) is a student centred approach which suggests a comprehensive measure of student progression from the moment entry is being contemplated to the point at which the student moves on from the completed programme of study to further study or employment. Therefore the approach measures not only output, but also input and process.

Three indicators for student progression and exit qualifications recommended by JPIWG are:

> ***Inability to progress**: measures the proportion of students assessed as being unable to progress.*
>
> ***Leave in good standing**: measures the proportion of students who are assessed as being able to progress but who chose to leave the institution or switch the subject of their qualification aim*
>
> ***Proportion who qualify**: means the proportion of students at each institution who qualified in the particular qualification aim.* (CVCP, 1995, p13)

However, the diversity and flexibility of the higher education system make the progression rate very complicated to calculate in practice. The JPIWG report identifies 12 possible pathways from which indicators of completion rates might be calculated. For example, it is suggested that students who are assessed (either through exam failure or less formally) by the institution as *not able* to progress to the next level may follow the following pathways:

pathway A: choose to leave HE
pathway C: retake the course
pathway D: take another course at the original level
pathway E: go to another institution at original level (CVCP, 1995, p45)

Nevertheless, these 12 pathways still have not solved the problem. The JPIWG consultation exercise in 1994 found that

> *'one set of critiques drew attention to the huge variety in patterns of enrolment for study and options for progression that now exist in higher education. This makes it increasingly difficult to decide whether or not students can be said to have left a programme or succeeded in a qualification aim. The growth in numbers and types of off-campus learners also makes it more difficult to define what is meant by a student.'*(JPIWG 1994, quoting from (Cave et al 1997).

Above all these made it very difficult to calculate progression rates in practice.

More importantly, the same problems with using non-completion rate as a performance indicator of teaching still exist, when progression rates are used as indicators of performance relating to the quality of teaching. Some research (e.g. Johnes 1990 and Entwistle and Wilson 1977) found that individual students leave higher education for many different reasons. Research (e.g. De Rome and Lewin 1984 and Johnes 1990) also indicates that the non-completion rate varies with subjects. Cave et al (1997) pointed out that 'the problems of drawing a clear line between external or individual reasons for students leaving without a qualification that are beyond the control of the institutions and reasons to do with the educational opportunities are also partly conceptual'. Johnes' and Taylor's research (1990) shows that at an institutional level, differences in non-completion rates between universities can be explained by three main factors: *the average A level score of each university's new entrants, the proportion of each university's students taking business*

studies or language courses (subject mix), and the proportion of each university's students accommodated in a hall of residence. They found that two student-related factors and four university-related factors can explain the variations in degree results and these factors are: *the mean A level score of entrants; the percentage of students who live at home; library expenditure as a percentage of total expenditure; whether or not a university is ex-college of advanced technology; whether or not a university is one of the new greenfield universities established in 1964/65; and whether or not a university is located in Scotland.*

Johnes and Taylor suggested that a standardised non-completion rate can be constructed to take into account these factors. Once a standardised non-completion rate has been calculated, this can be used as the bench-mark against which each university's actual non-completion rated should be compared.

This method has been adopted by JPIWG. JPIWG stated that 'the measured indicators (progression rates) will be affected by a variety of factors, including teaching quality. Particularly important will be differences between institutions in the entry qualifications of their students, the mix of subjects studied and differences in gender mix. It is proposed to make provision for this by the calculation of "standardised compactors" or "benchmarks" for each institution.' (CVCP 1995).

Nevertheless, how to interpret the results of these standardised indicators may be still the problem. Ball and Wilkinson (1994) pointed out that 'we may be concerned that the regression "explains" only 83 % of the variation...' JPIWG suggest that 'in some cases differences may reflect teaching quality. In other cases differences may reflect other factors, including students' individual characteristics and circumstances.'

The HEMS Group (1995) also questioned whether other input variables (e.g. student's ethnicity and socio-economic background) should be taken into

account.

The publication of the higher education sector's first performance indicator used drop-out rates which are based on tracking students from the year they enter an institution to the following year, and provides information about where they are in the second year. The benchmark for each institution is provided (The Times, Dec, 1999).

Employment and first destinations

Whether universities make graduates employable is a concern not only for government and institutions but also employers and potential university students. Naturally, graduate destinations were among the first list of indicators produced by the CVCP/UGC (1986). Since then the success of graduates in the job market has been an important indicator of performance of the university sector.

The debates have been focused on how to measure graduate destinations. One of the widely discussed issues is at what stage employment data should be collected. CVCP/UGC in their first statement proposed to collect data on the occupation of graduates after 12 months and five years post-graduation, but in their second statement, CVCP/UGC proposed to provide only figures for graduate employment after six months. The JPIWG suggested that a graduate employment follow-up survey two years after graduation would provide a useful source of information (CVCP 1995). Brennan and McGeevor (1988) carried out a survey on CNAA graduates at work and found that during the first three years of employment 58 percent of CNAA graduates had changed their jobs. Therefore they concluded that the use of first destination statistics to imply anything more than first destination could be misleading.

The second issue is about the categories of employment indicators. The JPIWG adopted categories of first destination used by the Association of Graduate Careers Advisory Service (AGCAS) and subdivided one category. These categories include: permanent employment, unemployment, short-term

employment, further academic study, other types of further study and training, not available for employment, and unknown destination. Brennan et al (1994) argued that other types of employment indicator should be collected and suggested three main kinds of information: objective indicators of income, proportions of unemployment, subjective indicators such as graduates' perceptions of their career paths and aspirations.

The third issue is that there are some factors which may have an effect on the first destination of each university's graduates. Johnes and Taylor (1990) found that about 90 per cent of inter-university variations in first destination of graduates could be statistically explained. Subject mix is a major factor. Other factors include the age of each university, the number of students per employer visit on the annual milk-round, whether a university is located on the geographical periphery of the UK labour market etc. The CVCP/UGC (1988) proposed to refine the first destination indicators to take into account the factor, subject mix. The JPIWG suggested that it is necessary to make some provision in the results for the factors, subject, gender, age, and entry qualification (CVCP 1995).

In summary, a number of potential performance indicators of teaching have been developed. Efforts have been made to analyse and define what is being measured, and the problems of using those measurements as performance indicators of teaching.

The diversity and flexibility of the current higher education system, the complexity of the process, and the untidiness of reality have made it very difficult to develop a performance indicator of teaching that can accurately measure what it is intended to measure. In fact, sometimes, 'what should be measured' is unclear. The interpretation of these performance indicators is also problematical. The extent to which individual indicators or even clusters of indicators provide a meaningful indication of holistic teaching performance is doubtful.

2.3 Value-added as a performance indicator of teaching in higher education in the UK

2.3.1 The introduction of value-added as a performance indicator of teaching in higher education

The traditional approach to evaluating teaching outcomes has been to examine the level of student achievement, for example, degree classifications. Higher classes of degree are seen to indicate better achievement of an institution or department. In recent years this approach has received increasing criticism. It was argued that the traditional approach can

> *...reward students through normative highly-competitive grading practices that disregard individuals' starting point differentials and concentrate, instead, solely on output differentiations between and among people* (Taylor, 1985, p191)

In America, much research (e.g. Astin, 1965, 1982b and Karabel & Astin, 1975 quoted in Pascarella and Ternzini, 1991) suggests that 'high quality' colleges start with a distinct advantage in terms of the academic ability, educational aspirations, level and clarity of career ambition, and family financial resources of the students they recruit and enrol and these particular students are characterised by pre-college traits that make them especially likely to obtain the bachelor's degree and in many cases move on to graduate or professional school. The league tables, as they are now called, showing the top higher education institutions(HEIs) by achievement caused strong reaction and evolved the introduction of a value-added approach in education. It was found that the top fifth of HEIs had better qualified entrants, therefore students in these 50 HEIs had higher achievement 'because of what happens in the admissions office rather than... what happens in the classroom' (Egan 1986, p10).

The value-added approach in education is derived from the Value-Added Tax

economic proposition which suggests that a commercial enterprise should be able to determine, for tax purpose, the value-added to a product or service as it moves along production channels. Therefore, in a similar fashion, an educational institution should be able to determine the value-added to a student as he or she proceeds through his or her years of study at the institution, and record such individual growth accordingly (Taylor, 1985).

As most economists originally argued that the VAT was too complicated to implement, many educators similarly argued that it was too difficult to determine added educational value. Therefore value-added educational assessment would not be implemented in higher education. However, the VAT system has in fact been implemented in many European nations for years, and equally there are some examples of value-added assessment in higher education in the USA. For example, Northeast Missouri State University has developed and used the value-added Educational Assessment for twelve years. The university has received the prestigious award for Excellence and Innovation in Higher Education for its value-added program (Taylor, 1985).

In the UK, students' A-level entry scores have been used to show how different institutions fare in the market place to attract potential entrants, and higher scores have been associated with better performance of institutions or departments. However entry scores only reflect the academic level of institutions or departments on entry and provide no evidence on how much value has been added to students' knowledge and skills.

Until recently, value-added measures received great attention in the UK. In 1987 the White Paper stated that

> *Evaluation of institutional performance also requires students' achievements to be set alongside their entry standards* (DES 1987)

The introduction of value-added as a performance indicator in HE in UK has also been associated with differences in culture between old universities and

new universities. Old universities usually recruit young students with A level entry qualifications. In contrast, new universities, former institutions in the public sector have defined themselves partly in terms of the admission of non-traditional entrants. Before the binary line was abolished, institutions in the public sector believed that the value-added approach would provide a macro PI to show performance at inter-sectoral level . The Morris report (1990) suggested that 'relatively high performance in terms of 'value-added ' is a distinctive characteristic of the PCFC sector, and of individual institutions within it.'

The attempt to find some ways of measuring the value-added can also be linked with a shift in the concept of quality (Cave et al 1997). The concept of quality has shifted from the traditional absolute concept of quality, the aim to maximise academic excellence, towards relative and transformative notions. Yorke (1991, p14) suggested that

> *'embedded in the notion of quality in higher education are two polarities:*
> *- academic excellence versus 'value-added'; and*
> *- threshold acceptability versus judgements of relative worth.'*

2.3.2 Main methods of calculating value-added

The idea of value-added is attractive. However, to quantify the relationship between inputs and outputs is a formidable problem. The following will focus on discussing main methods of calculating value-added in higher education in the UK. Nevertheless, a value-added program in higher education in USA and methods of calculating value-added in school education are also reviewed

2.3.2.1 Value-added measurements based on entry qualifications and degree results

Cave et al (1997) noted that main effort in the UK towards devising methods of calculating value-added has concentrated on finding a way to measure the difference between entry and exit qualifications. Methods for the calculation of value-added achieved by students in higher education can be divided into two main approaches--- 'Index' methods and the comparative method

(PCFC/CNAA, 1990).

All 'Index' methods attribute scores to measures of academic inputs (usually based on A level points) and academic outputs (based on degree results) and then calculate value-added by relating these two measures in some way. There are six types of 'Index methods', labelled A to F.

For example, Index method A, developed by the performance indicators project team at Nottingham Trent University , calculates value-added as follows

VASAI (value-added student attainment indicator)=SAI (student attainment indicator) /II (input indicator)x100
where SAI = ((Total scores awarded) / (number of students x10)) x100;
and II =(A level points /15) X 100

To calculate value-added, Index method B uses A level points as the input indicator which is then multiplied by an output weighting. Index methods C, D, E and F are all based on a formula which subtracts input from exit scores. The scoring systems for Index methods are described in the methodology chapter.

The main problem with 'Index' methods is that they combine scores measuring two different things, entry and exit qualifications. The value-added scores produced by the Index methods would favour courses with low A level and non-A level recruitment. There is no theoretical or empirical basis for the selection of scores. Consequently the results can be biased by the weighting given to certain types of entry qualification at the expense of others (PCFC/CNAA 1990, and Hadley & Winn 1992).

The CNAA/PCFC working party attempted to solve this failing of Index methods by developing the comparative method. The comparative method measures value-added by comparing degree results predicted for students with particular entry qualifications with the actual degree results achieved. The predicted degree class is the expectation of a student with a particular entry

qualification gaining a particular degree class and is derived from national data. The value-added score of a course, department or an institution is a function of the difference between the degree results achieved and the result predicted from entry qualifications (PCFC / CNAA 1990).

It is claimed that there are two advantages of the comparative method. *'Firstly, it is not an arbitrary score, but is based on an empirically derived expected value. Therefore the claim that a particular course did better or worse than expected when compared to national data is likely to have a robust acceptability. Secondly, because the playing field has been levelled all institutions have an incentive to improve their value-added scores whatever their current recruitment profile'* (PCFC/CNAA 1990, p11).

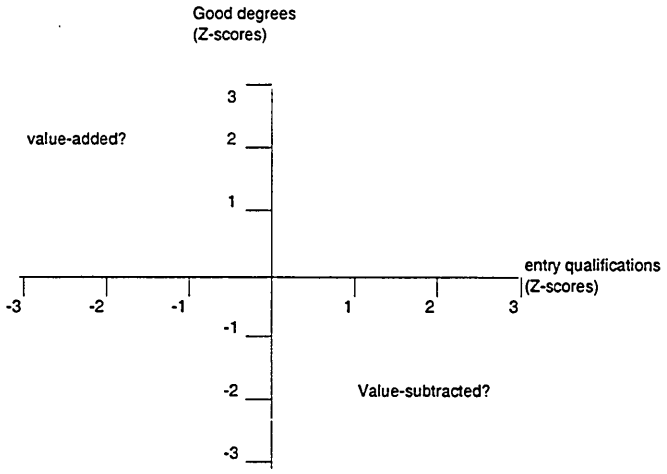
Additionally, it is also claimed that the comparative method has flexibility. It allows the comparison of a cohort of students with a range of other groups of students such as a previous cohort from the same course, cohorts from similar courses within the same institution or from similar courses in other institutions.

Although the comparative method does not make arbitrary decisions about the relationship between input and output and adopts national data as a yardstick, it still requires the arbitrary weighting of degree classifications. It is also based on the assumption that the same distance is travelled between different degree classes. With the comparative method, one student moving five classes is the same as five students moving one class using the comparative method (Gallagher 1991). (Further details of the disadvantages of the Index methods and the Comparative method will be discussed in the methodology chapter).

Additionally, Chapman (1996) suggested a framework for measuring value-added at institutional or subject level in higher education which uses Z -scores. The methodology is represented by figure 2.1. The axes are calibrated with reference to the proportion of good degrees awarded and the original entry qualifications of graduating cohort. The units on each axis are Z-scores expressing the position of individual observations in a data set relative to the

mean in standard deviation units. By locating individual universities (or departments) relative to the two axes, it is possible to establish the characteristic position of an institution or department within this framework. A position in the top left quadrant indicates a negative relationship between entry qualifications and degree results and has implication for the concept of value-added.

Figure 2.1 Value-added: a conceptual framework



2.3.2.2 Value-added measurement based on earnings differentials between graduates and non-graduates

Mallier and Rodgers (1995) proposed a measure of value-added in higher education based on the earning differentials between graduates and non-graduates. The purpose of this value-added measurement is firstly, to estimate the social rate of return for different degree classes, and, secondly, to propose a monetary-based performance indicator that could be used in the process of allocating resources in higher education.

The value-added for the individual graduates is calculated as follows:

Value-added per year	=	Average income of graduate with a given degree class	_	Average income of employee educated to A-level standard

The value-added for the individual who participates in higher education is

calculated in terms of the difference between the graduate's expected income and expected income of someone educated to A level standard. The individual graduate's expected income depends on the degree class.

Mallier and Rodgers (1995, pp 121) criticised the comparative value-added measure, arguing that 'it makes no attempt to measure directly the value-added of the individual student, and therefore gives us no benchmark to measure value for money against.'

Nevertheless, their method was questioned in that the income differences may be based on differences in aptitude rather than the value-added of higher education. Cave et al (1997, p130) pointed out that it is quite clear that

> *an analysis conducted on such an aggregated scale can not provide insights into the performance of individual institutions in dealing with particular cohorts of varied entry qualifications.*

2.3.2.3 The Method Used at Northeast Missouri State University (NMSU)

Northeast Missouri State University has utilised a value-added program (VAP) for over a decade. The VAP has been studied in detail by diverse sources such as about 300 individual colleges and universities in the USA, the United States Army Reserve Officer Training Corps, several prominent education , business, and economic publications. The common appraisal is that as an assessment approach to ensure educational effectiveness and accountability, it not only discovers problems within an organisation, but also becomes part of the managerial solutions (McClain et al 1986).

The purpose of the VAP is to measure gains in the total person - knowledge, analytic ability, skills, values, cultural awareness, and other personal development. NMSU attempts to measure a wide range of perspectives on value-added from every practicable angle including 'Attitude surveys, interviews, objective standardised tests, course-taking patterns, subjective tests, and extensive performance sampling' (Taylor, 1985). The VAP began with a pre-test (the ACT entrance examination), continued in a subsequent

post-test which is administered at the end of the second year, and extended to final year undergraduates. Students are also assessed before entering (pre-test) and after leaving university.

The following actual case study shows the diagnostic potential of the VAP. Prior to and during 1979, tests for all business students revealed a weakness in the area of mathematics. Similar results were also shown by tests for business graduates. Thus curriculum committees within the various business disciplines discussed several approaches which could improve the mathematical skills of business students and recommended that a stronger mathematics foundation should be required for all four-year business majors and these would be laid down in the curriculum requirements. The recommendation became effective during the 1979-80 academic year. Test results in each of the subsequent years showed students' improvement in mathematics. The improvements achieved were translated into fiscal benefits because the internal resource allocation was linked to student outcomes.

The VAP not only made it possible to diagnose a problem but also allowed NMSU to demonstrate to funding authorities and the public that education added value to individuals and that resources allocated to NMSU have been invested well.

However, there are several problems raised in VAP. First, Cave et al (1991) noted that although the program attempts to measure many aspects of student characteristics from different practicable angles, what relative weights should be applied to these aspects and which components are most important? Assuming it accurately reflects individual benefit from education, can it reflect social benefit from the education process?

Second, it is difficult to use the VAP to make a comparison between different institutions. In order to enable inter-institutional comparison to be made, tests need to be standardised. This will lead to the value-added score becoming the definition of quality. Institutions will inevitably reorganise their teaching practices

to obtain high value-added scores in the assessment, 'teaching to the test' rather than adding *real value* to students. Value-added measures may no longer measure what they were intended to measure (Taylor, 1985; Bauer 1986; Cave et al 1991).

Third, in order to prove the quality of institutions or departments, students would have to take extensive tests and obtain no direct positive benefit from them. Is it ethical? Could or should students take such comprehensive tests (Bauer 1986; Cave et al 1991)?

Finally, it would be time- and resources- consuming to operate a concurrent value-added assessment. The operation of the VAP involves a lot of work: developing and using tests, analysing and registering test data, and this work has to be carried out for a large random sample of students at least twice , 'before and after' measurements. As a result, the operation of the VAP will affect all activities in institutions or departments. The cost may exceed the benefit from such an assessment (Bauer 1986; Cave et al 1991).

2.3.2.4 Methods used for measuring value-added in school education in the UK

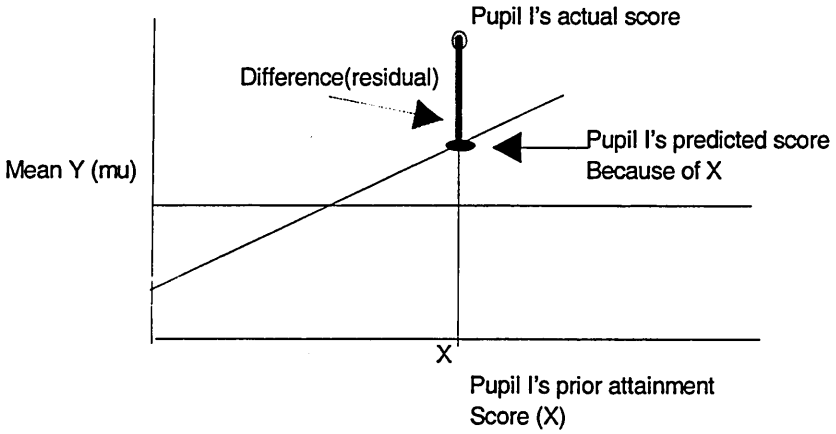
Although the type of information measured in value-added in school education is different from that in higher education, it worth noting methods of calculating value-added in school education.

Two main methods or statistical techniques have been used to calculate value-added in school education. One is simple regression analysis (or residual gain analysis), another is multi-level modelling analysis. Both methods have their advantages and disadvantages. There have been a lot of debates concerning which of the two methods should be used to calculate value-added in school education (e.g. Trower, and Vincent , 1995 and 1996, or Raudenbush and Willms, 1991).

Simple regression analysis defines value-added as the difference between a pupil's attainment statistically predicted from knowledge of the data, and actual attainment (see Figure 2.2). The statistical term is 'residual' which may be explained as that which is left over after prior attainment has been taken into account. The statistically predicted score can be roughly described as the average score obtained by similar pupils in other schools, so it provides a fair and understandable comparison for the score obtained by one's own pupil. The main advantage of this method is that it is easy to understand and therefore more accessible.

However, some statisticians (e.g. Raudenbush and Willms, 1991) have argued that pupils within a school are more alike than randomly allocated groups would be. This fact can be taken into account by using sophisticated statistical technique - multi-level modelling, developed in 1980's (Goldstein, 1987 gives an explanation of use of multi-level modelling in education). Therefore multi-level modelling will provide more reliable analysis when we try to estimate school effects. It is also capable of distinguishing students' effect and school effect. Multi-level modelling, however, is a very complex and advanced statistical technique. To understand it, people need to have advanced statistics knowledge.

Figure 2.2 The value-added residual as the difference between predicted and actualscore



This last point is a problem, because, as a performance indicator, measurement of value-added is expected to be simple and understandable. The research of Fitz-Gibbon (1997, p 23-24) found that *the views of head teachers were very largely positive towards the use of value-added data within schools..., particularly if it was kept* ***simple and understandable*** *and used cautiously (Fitz-*Gibbon 1997, Executive Summary).

Therefore, the Value-added Advisory Group (for schools), which included representatives from schools, the School Curriculum and Assessment Authority, DfEE, OFSTED and the Curriculum Evaluation and Management Centre (which was at University of Durham), recommended '*readily understandable*' as one of the criteria for developing a national value-added system in school education in the UK .

Furthermore, extensive statistical trialing carried out by the team of Value-added National Project (Trower and Vincent, 1995) demonstrated that very simple and accessible models of value-added yielded information that was, for all practical purposes, consistent with the information that would be derived from the application of more complex and sophisticated models.

Fitz-Gibbon, project director of Value-added National Project, also argued that

> *to model 'the school' is to make an assumption that there is some consistent 'effect' from the school on each pupil, either a constant effect, exactly the same for all pupils ... or, as the modelling becomes more complex, some other effect working in the same way on all pupils...such an assumption is unlikely to be valid and until effects are traced to causes, the interpretation of the residuals should be left to the schools, not modelled, by multi-level modelling...* (Fitz-Gibbon, 1997, p106)

2.3.3 Main arguments and difficulties concerned with using value-added as a performance indicator of teaching in higher education

The above have discussed specific arguments concerned with individual value-added methods. The following section will focus on general arguments and difficulties concerned with development of value-added as a performance indicator of teaching in higher education.

2.3.3.1 The relationship between students' entry qualifications and degree results

Most attempts to implement the value-added approach focus on academic attainment at entry and exit. This then is based on an assumption that there is a relationship between entry qualification and degree results.

Barnett (1988) examined the relationship between entry qualifications and degree performance, not by looking in detail at the statistics but by standing back and asking questions about the terms in which the debate has been framed. He argued that there is no reason to expect that there should be any relationship between entry qualifications and degree performance.

Barnett pointed out that the assumption that there is a positive relationship between entry qualifications and degree results is based on a number of additional and extremely odd hypotheses. These include:

> *(a) the process of higher education simply moves everybody forward more or less equally;*
> *(b) what is assessed at the point of entry is roughly the same kind of thing that is assessed at the point of exit; the logic is that higher education simply develops further the skills that enable the students to gain entry. Higher education, in this sense, is just further education, just more of what has gone before, rather than the development of higher order skills;*
> *(c) that students, despite being adults (of varying ages, from different socio-economic backgrounds, and with different social responsibilities) are likely to develop at the same pace;*
> *(d) that the educational process itself - the way in which the educators*

organise the learning experience - is unlikely to affect individual students differently. (Barnett, 1988, p18)

Barnett noted that these additional hypotheses run counter to many findings on student learning and student development.

He also noted that the issue about the relationship between A level points and degree results is connected with the educationalists' concern to demonstrate the extent to which A levels are predictors of success in higher education. He argued that the kinds of abilities that schools are trying to develop are not identical to those which the higher education system expects to be able to foster in honours graduates. Therefore A level may not be a good indicator of success in higher education.

However, I would argue that the all of the abilities that schools are trying to develop certainly are not identical to those which the higher education system expects in honours graduates. Nevertheless good basic knowledge, study skills and study habits that are developed in school education would provide a basis for students to learn high order knowledge and skills at universities, and these might well be necessary (but not sufficient) to degree success. A student starts university with high A level points means that the student has a good basis to do his or her degree course.

There are also some quantitative studies which have examined the relationship between entry qualifications and degree results. Sear (1983) suggested that the correlation between A -level and degree results does exist although it varies with subjects: the correlation for scientific subjects is generally stronger than that for arts and social studies, and the correlation for languages falls in between.

Johnes and Taylor's work (1987 and 1990) suggested that the mean A level scores of a university's students is highly significantly related to degree results. They noted that

using regression analysis, it was found that over 80 per cent of the variation between universities in degree results can be explained (statistically) by a set of plausible explanatory variables, the main one being the mean A level score of each university's student entrants (Johnes and Taylor 1990, p113).

2.3.3.2 The Comparability of entry qualifications and degree classifications

There are also arguments about the comparability of A level and comparability of degree classifications. With comparability of A level, Barnett (1988) argued 'to what extent would the 'same' candidate have received similar grades if the examinations had been taken with another Board?' Thus one institution recruiting a substantial proportion of its entrants holding qualifications awarded by one examination board may have a better intake than one with apparently similar qualifications, but awarded by another board.

The comparability of A level grades over time is also questioned. Barnett also noted that there are also suggestions that 'A'-level standards are slipping and top 'A' level grades can be achieved more easily. If this is the case, value-added results may be distorted when comparison of value-added is made over a period of time.

Barnett is right, there are questions about the consistency and comparability of A level grades both between examination boards and over time. Nevertheless, I believe that the consistency and comparability of entry qualifications should be relative. In reality, it may be difficult to find absolute consistency and comparability. One could also question the comparability of A level grades within an examination board because different examiners may have a different understanding and judgement of the standards of marking required. This is despite the very strenuous efforts that are made to 'standardise' the scoring of exams. We may also have to accept a philosophical principle that everything is

changing. We may not be able to find any absolute consistency and comparability of A level grades between different years. The question is to what extent the variations in consistency and comparability of A level are acceptable or not acceptable. Furthermore, we also have to appreciate that these A level grades have been used as an important element of the criteria for university admissions and offering of jobs, although the problems with consistency and comparability of A level grades are recognised.

Additionally, there is also a debate about the extent to which different qualifications are equivalent. Diversity is a main feature of the current higher education system in the UK. Universities are encouraged to recruit students with non traditional qualifications. Particularly, the new universities have a tradition of admitting students with qualifications other than A level qualifications. However, to what extent, are these different qualifications equivalent? The natures of these qualifications are different.

Concerning the reliability of degree classifications, it is argued that most disciplines use constructed response approaches (essays, problem solving, coursework and project) in their assessments which are dependent on the professional judgement of the examiner (Bennett, 1993). Barnett (1988) questioned 'How far is the pattern of degree classifications a function of the particular institution or of the particular subject or the particular department?'

Furthermore, any value-added measurements which involve comparison between institutions or comparison over a period of time would assume that a specified degree class indicates an equivalence of achievement across institutions or between different periods of time. However, it is argued that although the external examining system is intended to maintain the standard of degree classifications across institutions, there are serious doubts about its credibility (Silver et al 1995).

The above argument may be true, nevertheless, degree classifications are final results of students for their three or four year study in higher education and are

used as a reference for employment. Apart from these, what other educational attainment measurements are available ?

2.3.3.3 The concept of value-added

What should be measured in terms of value-added by higher education

What value is added to students who enter into higher education? Fincher (1985) suggested that value of college graduates is evidently reflected in the salaries they can command upon receipt of a college degree, however, the missing factor is obviously the difference between learner capabilities prior to their educational experiences and graduate capabilities after earning a college degree. Thus, it makes sense to judge value-added by education in terms of achievement after-education less student achievement before-education. Fincher then explained that in economic terms, the value of education is the contribution that educational programs and institutions make to the skills, competencies and accomplishments of the individuals.

The 1987 White Paper suggested that what needs to be measured is not only improvements in specialist knowledge but also in communication skills, and in positive attitudes towards enterprise and employment patterns.

Cave *et al* (1997) suggested a broad concept of value-added by higher education. They noted that value of education encompasses private benefits and social benefits. In terms of private benefits, education increases the earning potential of individuals and makes a contribution to personal development. The benefits to society derive from having one more highly educated individual, and from any positive external effects. An example of a positive external effect is that a well-educated person may increase the productivity of a less well-educated person, either by adopting more efficient methods of work, or by the less well-educated person learning from the highly educated employee (Le Grand and Robinson, 1979 quoted in Cave *et al* 1997).

The above discussions about value-added in higher education are at different levels. The discussion by Cave *et al* is at a very general level. They classify the

value of higher education into private benefits and social benefits. The value of college graduates is also discussed by Fincher. 'The salaries, skills, competencies, and accomplishments of the individuals', may be seen as 'private benefits' as suggested by Cave. The 'skills, competencies, and accomplishments of the individuals' may be considered 'personal development' as defined by Cave, but they are still very abstract. The suggestion of the 1987 White Paper about value-added, 'specialist knowledge, communication skills and attitudes towards enterprise', is very specific, nevertheless, it is not comprehensive.

What has been measured in terms of value-added by higher education

Although in theory, the concept of value-added can be broad, in practice, most attempts to implement value-added approaches in the UK have adopted a much narrower concept of value- added which compares the academic attainments of students at entry to higher education (i.e. entry qualifications) with their attainments at exit (i.e. degree results) such as the comparative value-added approach and Index methods noted earlier.

Another attempt to measure value- added in terms of earnings increase is made by Mallier and Rodgers (1995) as we have seen. They attempted to estimate a monetary value of the better qualifications by comparing the earning of a graduate employee with that of an A level educated employee. Their attempt to measure value-added is based on the premise that the employees will be paid according to their marginal product. Therefore, individuals educated to graduate level will be paid more than those educated to A level standard by virtue of their greater productivity resulting from value-added in higher education. However, as Mallier and Rodgers recognised, factors other than relative productivity often explain pay differentials. For example, employers may not match marginal productivity and wages because it is too complex and expensive to measure individual contributions to production. Wages are actually based on simpler and less expensive criteria, and individuals accept the consequences of such administrative mechanism (Frank 1984). Nevertheless,

Mallier and Rodgers (1995, p112) argued that

> *...earnings data shows a strong relationship between education and earnings levels and so we believe that the basic principle is still applicable.*

However, even if we accept the assumption that employees are paid according to their marginal product, there are other difficulties when using income data to construct a measurement of value-added in higher education. Mallier and Rodgers noted that one of the difficulties is that except for the marginal employee, the value of the marginal product will be above the wages' rate, therefore it may underestimate the value of the employee to the employer, and hence the true value-added by higher education. Another difficulty is concerned with distinguishing between the consequences of value-added due to education and the effects of subsequent on-the-job training.

In USA, as noted above, a significant attempt to measure value-added in higher education was made in the 1970s at the Northeast Missouri State University (NMSU). It is claimed that NMSU value-added assessment model is not the pure value-added assessment program, 'it is however, purer than most educators realise has been implemented to date'(Taylor, 1985).

Indeed, the NMSU value-added program has attempted to measure 'gains in the total person'. Their measurements include cultural awareness, interpersonal skills, self-confidence, problem solving, and functioning in the larger society. To accomplish such a task, a variety of methods have been used to collect data such as 'Attitude surveys, interviews, objective standardised tests, course-taking patterns, subjective tests, and extensive performance sampling'(Taylor, 1985). However, as we noted earlier, the NMSU value-added program has its problems in practice such as what relative weights should be applied to these aspects and which components are most important?

The above shows that the conception of value-added itself may be simple, and can be broad. However, in practice, we lack the ability and the capacity to realise the full conception of value -added.

2. 3.3.4 Value-added and gain scores

As it was described earlier, the concept of value-added was introduced into higher education as one of the performance indicators which was intended to measure the efficiency and effectiveness of higher education. Nevertheless, it is suggested that '*The value-added measure is a crude version of an extensively researched concept in the educational measurement literature, namely, the 'gain score'*'(Morrison *et al* 1995, p131*)*

Gain scores (or change / difference scores) have been used to assess the impact of a teaching programme. Gain scores are often measured by testing a student at the beginning of a programme and repeating the test (using the same test or a parallel test) at the end. Morrison *et al* (1995) argued, 'A level and degree scores aren't even measured in the same test metric'. Therefore it is suggested that if value-added is to be measured, some new, fairly standardised examination will be required at entrance to university (Pollitt, 1990).

Nevertheless, the feasibility of implementing such a new exam is doubtful. The financial, administrative, and political implications of operating such a new exam will be huge. There will be also an ethical problem. This can be seen from experience of the NMSU value-added program, as discussed earlier.

Assuming that we could operate a pre- and post- test to measure student achievements at the start and end of a study period, educational measurement literature suggests that there are other difficulties. Fincher (1985) discussed some of these difficulties. Firstly, gain scores seldom have consistency, stability, or reliability. This is because the variance that is common to the pre- and post-test makes a duplicative contribution to gain scores, leaving the specific variance and random error in each test as the major determinants of

the observed difference.

Secondly, a significant correlation of gain scores and pre-test scores may imply that the initial abilities of students are related to the amount or extent of learning that takes place. Therefore the interpretation of gain score is greatly dependent upon instructional objectives and/or intended outcomes

Finally, pre-and post -testing paradigms imply a two stage test, but if the intent is to assess changes in performance as a result of the institution, it is often advisable- *but seldom feasible*- to have at least one intermediate test in order to plot change or growth curves and depict in graphic form the academic progress of students. Therefore, even if we can operate such tests, the measurement of the value-added is still not accurate and interpretation of the value-added is problematical.

I believe that the question is whether value-added as a performance indicator really has to be seen so strongly related to gain score. The purpose of developing a value-added measurement is to provide information for management. The gain score is developed in the context of educational measurement or research for different purposes. A crude measure of value-added may be still useful as a management tool.

Further more, Morrison (1995) also argues that since gain scores are negatively correlated with initial scores, under certain conditions, it can have zero reliability. With a value-added approach, a lower A level points at entrance will generate a high value-added score, while a higher A level points will lead to low gains.

2. 3. 4 A call for research

Despite the difficulties and complexities with developing a value-added performance indicator in higher education described above, Cave et al (1997, p135) concluded that

research in this area is still in its infancy and by no means at a stage where we can say whether value-added measures can or cannot be made operational at some level.

In practice, the JPIWG recognised some of the difficulties of developing a value-added performance indicator and decided to drop 'value-added' from the list of performance indicators recommended in their report published in 1995 before more research in this area has been done.

From the above literature review, the conclusion emerges that there is a need to undertake further research in this area in terms of how to measure value-added and whether value-added can be used as a performance indicator in higher education in the UK. It has also emerged that this research would be exploratory because of the difficulties and complexities of the issues revealed by the literature review and because this research is a PhD project with limited resource and time.

Therefore the present study will not to attempt to solve all these questions as described, in other words, we will not attempt to provide an off-the-shelf, ready-made value-added performance indicator. However, what can be done is to explore the feasibility of construing and using a value-added performance indicator to lay out clearly some of the strengths and weaknesses of a value-added approach.

CHAPTER THREE: METHODOLOGY

The last chapter (Literature Review) has revealed the problems with existing methods of calculating value-added and some difficulties of using value-added as a performance indicator of teaching in higher education. It indicates that the research in this area is still in its infancy. This research aims to explore the issue of development of a value-added performance indicator of teaching in higher education. It does not intend to provide a ready-made value-added performance indicator for immediate use in the assessment of teaching in higher education.

The aim of this chapter is to outline how the project has been carried out. The first section, research design, further defines the research question and discusses how this project studies it. It notes that this research explores the issue of value-added as a performance indicator in higher education from two aspects. The first aspect focuses on measuring value-added, which develops an alternative method of calculating value-added in higher education. The second aspect focus on investigating problems over the feasibility of using the new value-added measurement as a performance indicator, which employs a qualitative method to investigate academic staff views. The second section of this chapter reports the development of the alternative method of calculating value-added, and the data used for testing the proposed method of calculating value-added. The third section discusses how the qualitative study was conducted. This includes the objectives of the qualitative study, the research tool chosen to collect data, the participants, the procedure of the data collection, and the method used to analyse the data.

3.1 Research design

3.1.1 Methodologies of research in the social sciences

Lincoln and Guba (1985) suggested that there are three paradigm eras in relation to enquiries about the world: prepositivist, positivist and post-positivist. They (1985) outlined what they see as five assumptions on which the positivist

paradigm is based.

An ontological assumption of a single and tangible reality 'out there' that can be broken apart into pieces capable of being studied independently; the whole is simply the sum of the parts.

An epistemological assumption about the possibility of separation of the observer from the observed-the knower from the known

An assumption of linear causality; there are no effects without causes and no causes without effects.

An axiological assumption of value freedom, that is that the methodology guarantees that the results of an inquiry are essentially free from the influence of any value system

An assumption of the temporal and contextual independence of observations so that what is true at one time and place may , under appropriate circumstance (such as sampling), also be true at another time and place.

Based on these positivist assumptions, the quantitative methodology of the social sciences has developed. Quantitative research methods are intended to identify correlation between variables and specify 'causal relationships'. To establish a causal relationship, the variables are systematically and artificially manipulated to see if the experiments will produce expected results.

Nevertheless, positivists do not rule qualitative techniques out of court. They regard qualitative techniques as particularly valuable in situations where there is no possibility of experimentation, or the variables have not yet been specified and an open-ended process to discover the relevant factors is needed (Galtung, 1967).

Filstead (1970) criticized pre-determined measurement theory and indicated that social scientists 'have tended to blend, reshape, and distort the empirical social world to fit the model they use to investigate it.' Another challenge to the claims of positivism came from Ions (1977). He has put forward an objection of quantification when it becomes an end in itself and he pointed out:

> *The argument begins when we quantify the process and interpret the human act... However high-minded the intention, the result is depersonalization, the effect of which can be felt at the level of the individual human being, not simply at the level of culture.*

Lincoln and Guba also (1985) indicated the problem with objectivity underlying positivism. They were concerned that the research procedures based on the positivistic paradigm lead to the exclusion of the necessary humanness of the participants by refusing to give them equal rights of determination of the research. Research is carried out with an outside (objective) perspective and this virtually undermines the importance and significance of an inside perspective (subjective).

One of the critics of the positivist social science view of man is Hampden-Turner (1970). He concludes that the social science view of man is biased in that it leads the social scientist to concentrate on the repetitive, predictable and invariant aspects of the person but ignore the subjective world.

Post-positivism emerges in opposition to positivism. Although post-positivists within social science itself have a variety of schools of thought, in terms of epistemological viewpoint, they all reject the belief that human behaviour is governed by general laws and characterized by underlying regularities. Their common belief is that the social world can only be understood from the standpoint of the individuals who are part of the ongoing action being investigated. Social science is seen as a subjective rather than an objective undertaking, as a means of dealing with people's direct experience in specific

contexts. People act in terms of socially constructed rules of meaning, not as a lawful outcome of determining variables (Cohen and Manion, 1989).

3.1.2 Research design

The broad interest of this project is to explore the issue of value-added as a performance indicator. Given the difficulties and arguments concerning the measurement and use of value-added as a teaching performance indicator, which have been discussed in the literature review, this project will explore the issue from two aspects.

The first aspect focuses on measuring value-added. On the basis of the evaluation of difficulties with measuring value-added in higher education in the UK and the evaluation of the existing value-added measurements, an alternative method of calculating value-added is developed. This method is then applied to actual data to evaluate the feasibility of the method in actual use. The first part of study may be classified as a quantitative study.

The second aspect of the study steps back and critically assesses whether the method developed in the first part can be used as a performance indicator of teaching. It does so by employing the qualitative research method to investigate academic staff views on the issue. The second part of this study is qualitative.

The above design of this research indicates that positivism is the methodology to be used for developing a method of calculating value-added. It is assumed that one student who has a particular 'entry qualification' or 'exit qualification' is comparable to another student with that qualification and the value-added is an outcome of certain causal factors relating to departments and institutions. Therefore the first part of this research is positivist quantitative research. The second part of the study uses qualitative techniques to collect data. The data from the qualitative study are tentatively analyzed in a positivistic way. For example, the data are used to identify some variables or factors which may have impact on students' academic achievements. Nevertheless, it is found that

there are some things that are not tangible, and 'variables' can not always be separated out from specific context. The academic staff interviewed are expressing the meanings of the notion of value-added in the context of their understanding of the course and of higher education: meanings not variables. At this point, the second part of the study may be taken as post-positivist qualitative research.

3.2 Methodology of the quantitative study

3.2.1 The objective of the quantitative study

The quantitative study focuses on measuring value-added. The objective of the quantitative study is to develop an alternative modest method to quantify the relationship between entry qualifications and degree results using available data, and then to use this method to measure value-added of institutions.

3.2.2 Development of an alternative method of calculating value-added in higher education in the UK

3.2.2.1 Some considerations for the development of the alternative method of calculating value-added in higher education in the UK

It is important to emphasise that the development of the method of calculating value-added in this research is based the following considerations.

The need for a value-added performance indicator

The literature review has extensively discussed the need for performance indicators and why the concept of value-added was introduced, nevertheless it is worth emphasising some points here.

I believe that a value-added performance indicator plays an important role in measuring the achievement of the objective of widening access to higher education. One of the objectives of higher education in the UK is widening access. Among the proposed performance indicators, some indicators appear to measure the achievement of this objective, but only measure the performance of the higher education sector or individual institutions in achieving

such an objective in terms of quantity rather than quality. For example, in the recently published, first higher education sector's performance indicators (HEFCE, 1999) 'participation rate' provides information on the proportion of entrants to higher education who come from certain groups that are under-represented relative to the population as a whole. However, this performance indicator does not indicate how well these students from under-represented groups are doing after their study at universities. If, for example, an institution recruited a large number of students from under-represented groups, but at the end, these students all failed, we can not say that this institution performs well in achieving the objective of widening access to higher education. A value-added performance indicator will be able to indicate performance of the sector or institutions in achieving the objective of widening access to higher education, in terms of quality.

The data

It should be appreciated that only limited data are available for this research (the detail of the data obtained for this research will be further reported in a later section of this chapter), thus the development of the value-added approach is limited by the availability of data. For the development of performance indicators as a whole, performance indicators should, if at all possible, be based on readily available data. The literature review has revealed the problems arising from collecting data. Of course, performance indicators as management tools, once adopted, are not expected only be used once, and are expected to be used in a large scale. Currently, in the UK, in both school education and higher education, there are already some concerns that quality assessments have created undue burdens on schools and universities. Indeed, performance indicators themselves should be cost effective. In terms of measuring value-added, it should be appreciated that only data on entry qualifications and degree results are available.

The simplicity of the measurement

The literature review (e.g. that concerned with the experience of value-added measurements in school education in the UK) indicates that performance

indicators as management tools should be as simple as possible. Because a simple performance indicator will be easily understood by those using it, therefore it will be more accessible than a complicated one, and a simple performance indicator will be easy to put into operation in terms of time and resource. Thus it can be continually used so that year on year comparisons can be made. As Ewell and Jones (1994) noted:

> *Many promising indicator systems fail simply because they are too expensive, too complex, or too politically costly to implement. Often* ***the simplest is the best****, even if it initially seems less technically attractive.*

The Performance Indicators Steering Group in their report (1999) suggest that it is best not to use a sophisticated model directly in constructing performance indicators although it is necessary to use a sophisticated model to carry out some research in constructing performance indicators.

The accuracy of the measurement

The method of calculating value-added developed is based on entry qualifications and degree results although it is recognized that in a strict sense, entry qualifications and degree results can not be regarded as pre-test and post-test of gain scores which measure student starting points and finishing points of a study of program. The arguments for this are firstly, the literature review (especially recounting the experience of NMSU) suggests that it is not feasible to conduct special tests which may be counted as pre-test and post-test of gain scores in line with the educational measurement literature; secondly, the research indicates that there is correlation between entry qualifications (A level) and degree results. In reality, entry qualifications are a major criterion in university admission. Therefore, entry qualifications and degree results can be treated as a crude measure of starting and finishing points of a study of a program, nevertheless, a crude version of value-added performance indicator may be better than nothing.

Furthermore, the survey of literature found that there are problems with other

proposed performance indicators of teaching and none of them is accurate. The problem with comparability of degree results between institutions as discussed in the literature review is actually applicable to all performance indicators which have something to do with a degree such as progression rate and exit qualification.

It was based on the above considerations and evaluations of major existing methods of calculating value-added, that this research has developed an alternative value-added method.

3.2.2.2 An alternative method of quantifying the relationship between entry qualifications and degree results

Given the disadvantages of existing value-added methods, this research suggests that the relationship between entry qualifications and degree results may be quantified in the following way:

Step one: Show differences in academic attainments between departments or institutions at entry. This will enable comparison of entry qualifications between departments or institutions to be made.

Step two: Indicate differences in academic attainments between departments or institutions at exit. This will enable comparison of degree results between departments or institutions to be made.

Step three: the difference in academic attainments of a department or an institution compared with other departments or institutions between entry and exit is value-added for this department or institution.

The differences in academic attainments at entry and exit may be measured by ranking.

The fundamental difference between existing methods and the method in this study is that existing methods calculate value-added by directly measuring entry

qualifications and degree results. Both the Index and the Comparative method calculate value-added for each institution, or department first, and then examine the difference in value-added between different institutions (or departments). In contrast the method in this study calculates value-added by measuring **differences** in entry qualifications and **differences** in degree results between institutions or departments. The alternative method enables comparisons of entry qualifications between different institutions (or departments) and comparisons of degree results between different institutions (or departments) to be made first, and then the relative value-added scores to be calculated. The value-added results generated from the alternative method indicate change in the differences in academic attainments between institutions or departments.

The value-added scores calculated by the alternative method are relative rather than absolute because they are based on comparison between institutions or departments. The main purpose of measuring value-added is to make a comparison of performance between institutions or departments. Any comparative results depend on objects between which comparisons are made. Therefore, value-added scores produced for evaluating performance may inevitably be relative as Cuenin (1987) noted.

The relative value-added score has implications for the policy of widening access, a policy that encourages institutions to take students without traditional qualifications. The alternative method clearly indicates differences in academic attainments between institutions or departments at entry, and differences in academic attainments between institutions or departments at exit.

3.2.2.3. How the alternative method deals with different entry qualifications

How to deal with different entry qualifications is a difficult issue in measuring value-added. The policy of widening access to higher education leads to students entering universities with diverse entry qualifications. These entry

qualifications are not equivalent. Therefore it is very difficult to construct a measurement to encompass students with different entry qualifications.

Six types of 'index' methods are based on an arbitrary choice of how entry and exit qualifications are scored (PCFC/CNAA, 1990). Index methods C and D, for example, give different scores to a range of entry qualifications. Index Method C does not differentiate between the number of A level points. The entry scores are shown in table 3.1. The entry qualification scores of Index Method D are derived from Further Education Statistics Record (FESR) coding. For the highest A level point score, which in the FESR is 13-15, the mid-point of 14 is taken as the sum. Thus the entry qualification score ranges from 0 to 14 (see table 3.1). An HNC/D is scored as 3 A levels, 3-8 points, and 5.5 is the mid-point. An ONC/D art and design foundation is scored as 2 A levels at 2-4 points, 3 being the mid-point.

Table 3.1 Entry qualification scores

Qualifications	Index method C	Index method D
3 As 13-15pts	20	14.0
3 As 9-12pts	20	11.5
3 As 3-8 pts	20	5.5
2 As 8-10 pts	15	9.0
2 As 5-7 pts	15	6.0
2 As 2-4 pts	15	3.0
HNC/Cert Ed	25	5.5
ONC/A&D Foundation	not be scored	3.0
AS level or RSA	5	not be scored
Univ or OU credits	30	not be scored
No formal quals	0	0.0

There is no theoretical and empirical basis for the selection of scores. Why, for instance, was the HNC qualification given a score of 5.5 but the Art and Design foundation course was given a score of only 3.0?

Because of the weighting difficulty, some of the 'index' methods of calculating

value-added only measure value-added for A level students and then give scores on basis of A level points. Index method B, for example, converted A level points to entry qualification score on the basis of 15 A level points = 1; 14 = 2,...1 A level point = 15. Index Method E calculated entry qualification scores by taking the A level points times 3, thus 15 A level points =45, 14=42...1=3, 0=0. These methods take no account of other entry qualifications.

The PCFC / CNAA project report (1990) pointed out that 'with the index methods a movement of one degree class varies in its effect according to entry qualification. All graduates are not equal in their impact on the total score.' It is claimed that the Comparative method uses national data to provide a national benchmark covering a variety of entry qualifications and derive the relevant 'expected' results against which the actual results for entrants with qualifications in the same category are compared, thus, with the Comparative method, the movement up or down of one degree class has the same effect on the score regardless of entry qualifications. The Comparative method also eliminates the need for scoring entry qualifications.

It is claimed that the Comparative method provides a level playing field. The Comparative method can tell us how well a course is doing compared with similar courses with similar intakes. It can calculate value-added for different kinds of entrant.

However, this advantage actually causes other difficulties. The Comparative method frequently involves dealing with small numbers in calculating CVA at institutional and course level. Straw and Kaye (1995) noted that an unrealistic 'yo-yo' effect can occur between years when few data are involved. For example, the data published in the CNAA/PCFC report (1990) indicates that the CVA for law students with foreign qualifications moving from +48 in 1988 to -52 in 1989. This shows a dramatic decrease in value-added. Nevertheless the CVA is based on only two students graduating in 1988 and one in 1989. Therefore the number of students is too small to draw any conclusions. The small number of students also leads to the problem of testing significance.

Given the above difficulties and arguments, the present study examines the characteristics of entry qualification data obtained and finds that although entry qualifications are diverse, the majority of students' entry qualifications are concentrated on several major groups. For example, students entered a BA Business Studies course in one institution with nine different types of entry qualifications over the period of 1988-1991. However, there is a consistency over the whole period in that the majority of students' entry qualifications(about 80%) fall into two major groups, A level and BTEC. The number of students with other entry qualifications is very small. For example, the number of students with a European Baccalaureate on the course was only one in 1990. Therefore value-added measurement may concentrate on these majority students (e.g. A level and BTEC groups) rather than all students (e.g. students with all nine different entry qualifications).

To avoid attributing arbitrary scores to different entry qualifications, the comparison in input may be made within each entry qualification group. For example, for the BA Business Studies course, separate comparisons may be made between A level students and BTEC students. Thus it overcomes the disadvantages of the index methods (arbitrary scores) and the Comparative methods (the number of students with some entry qualifications is too small to draw conclusions).

On the basis of results on value-added for the major groups, value-added for the courses can be measured by calculating the sum of value-added for the individual groups. For example, the calculation procedure for the BA Business study course is as follows:

$$VA = VA_{\textit{for A level}} \times \text{Weight}_{\textit{for A level}} + VA_{\textit{for BTEC}} \times \text{Weight}_{\textit{for BTEC}}$$

$$\text{Weight}_{\textit{for A level}} = \text{Total no. of A level students / Total no. of students of two groups}$$

$$\text{Weight}_{\text{for BTEC}} = \text{Total no. of BTEC students / Total no. of students of two groups}$$

3.2.2.4 How the alternative method measures outputs

Table 3.2 shows how six types of 'Index' methods score exit qualifications. It can be seen that exit qualification scores are arbitrary just as entry qualification scores are.

Although the Comparative method eliminates the need for scoring entry qualifications. It still gives arbitrary scores to degree results (1st =6, ...fail = 1). This scoring system is based on the assumption that the difference in value between a first and upper second class degree is the same as that between an unclassified degree and a fail (Hadley and Winn, 1992). Under the CVA, one student who was expected to fail but actually obtained a first has the same effect on CVA as five students who were expected to obtain 2.2 degrees but actually obtained 2.1. This is because CVA takes into account both movement from one degree classification to another and distance travelled between classifications (Gallagher, 1991).

Table 3.2 Exit qualification scores

	Index method A	Index method B	Index method C	Index method D	Index method E	Index method F
1st	100	5	70	80	70	50
2.1	80	5	65	75	65	40
2.2	60	4	60	70	55	30
3rd	40	4	55	65	45	20
Pass	20	3	50	60	/	15

Note: / = not be scored

Therefore as Cave *et al* (1997, p129) concluded

> *Clearly this is contestable, and it demonstrates the impossibility of devising any system of value-added which is wholly free of arbitrary weights.*

Given these difficulties with measuring degree results, in this research, the quality of a degree is measured by calculating both percentages of good degree

(namely first and upper second degree) and degree scores. One of the advantages of using the percentage of good degrees is that it avoids scoring different degree classifications. Nevertheless, one of the disadvantages is that it loses information about the achievements of students with other degree classifications. Degree scores would provide more comprehensive information about the quality of the degree; nevertheless, as it is shown above, it is difficult to be free of arbitrariness. To calculate degree scores, the present study attributes scores to degree class according to the criteria (median mark) that are used to award the degree class by the institution: 1st = 72, 2.1 = 64, 2.2 = 54, 3rd = 44, pass = 35, fail = 25. Those students who did not complete courses are not included in the value-added measurement.

In short, the main advantages of the method of calculating value-added developed in this study over the major existing value-added measurements are as follows:

Firstly, it compares like with like and avoids the disadvantage of Index methods which attribute arbitrary scores to measure two different things. It links entry qualifications to degree results by comparing **differences** (or position) in academic attainments between institutions or departments at entry, with **differences** (or position) in academic attainments between these institutions or departments at exit.

Secondly, it measures value-added by making comparisons between students with the same type of entry qualifications, thus it retains the advantage of the Comparative method which claims to provide a level playing field and calculates value-added similarly for different entrants. Nevertheless the method developed in this research focuses on major groups of diverse entry qualifications, Thus it overcomes the difficulty in calculating CVA at institutional and course level, which frequently involves dealing with a small number of students.

The disadvantage is that the alternative method uses rank to measure differences in entry qualifications and degree results between institutions or departments, thus it reflects only order and does not possess the property of equal intervals. This measurement hides some differences. For example, the differences between institutions ranked 1 and 2 may be much bigger (or smaller) than that between institutions ranked 3 and 4.

Additionally, although in this research we argue that a simple performance indicator has its advantages, it is a crude measurement. It is recognized that a measurement based on sophisticated statistical techniques would provide a more accurate description of variability in degree results and distinguish between student attributes and institution and department attributes. If data are available, there is a need for **research** in this area as in corresponding research in school education, although the sophisticated approach may not be suitable to be directly used as a performance indicator. That is, sophisticated measures are appropriate as research tools, but probably not as management indicators.

The alternative method may be used in two ways. Firstly, it may be used to enable direct comparison between institutions or departments to be made.
Secondly, it may be used to enable comparison between the average of national data and individual institutions, or comparison between the average of an institution and individual departments to be made.

3.2.3 The data used to test the proposed method and constraints of the study

To test the above framework of measuring value-added, empirical data concerned with students' entry qualifications and degree results were needed.

Originally, the present study attempted to focus on using the framework to make comparison of value-added between institutions. Therefore the initial intention was to obtain data from 20 universities. Two potential sources of data were identified. One was the University Statistics Record (USR) and another

the Further Education Statistical Record (FESR). Most of the research in this area used this data resource such as Johnes and Taylor's work (1990) and the PCFC / CNAA (1990) project. Another source was individual institutions. When I tried to collect data, unfortunately, it was the time when HESA took-over USR and FESR. This increased the difficulty, in terms of finding the right persons to contact, and the waiting time for response.

It might be worth noting a specific experience concerned with collecting data. In order to obtain the data, I have tried to contact people in the USR and FESR by letter in my own name as a Ph.D. student, but I did not receive any response. After that, we contacted them again in the name of my supervisor, a professor, and we received the response. However, even so this source of data had to be given up. The main reasons are firstly, the information on A level points for many institutions is not available because of change of organization at the time, and secondly, because of confidentiality constraints. HESA will provide the data but only subject to an individual institution's permission even though it was emphasised that individual institutions would not be identified.

Efforts were also made to obtain data directly from individual institutions. The following methods were used to contact the relevant persons: e-mail, phone, letter, visit, and personal contact. To obtain the data, it was explained to the relevant people what the aims of the study were, the data required, and the methods used to analyse data. They were also assured that the study would preserve the anonymity of the individuals and institutions concerned. About 25 institutions were contacted. Given the sensitivity of the research topic, most institutions are reluctant to release data which may be used to assess their performance. Additionally, many institutions' data were not computerized until the 1990s. Providing data required by me would presumably increase the relevant staff workload.

The data from two institutions were finally obtained. It is right to express very grateful thanks to these two universities. The detail of these data will be noted in the following chapters where analysis of these data is presented. With the

given data, the framework proposed has mainly been tested within one institution. Therefore the aim of the study is now focused on developing a performance indicator for use within one institution. In fact, developing performance indicators for use within individual institutions is also needed, as Jarratt report suggested that

> *A range of performance indicators should be developed, covering both inputs and outputs and designed for use both within individual institutions and for making comparisons between institutions (Jarratt 1985:36).*

The framework is used with the two sets of data. With the first set of data, the data is used to measure value-added at course level. With the second set of data, the framework is used at two levels. First it measures variations in the value-added of the institution between different years by aggregating all schools. Second, it evaluates the value-added of different schools within the institution by aggregating different years.

To some extent, the comparison of performance of institutions between different years may be seen as comparison at institutional level; the comparison of performance between different schools may be seen as comparison at subject level across institutions.

It also should pointed out that the data obtained are from two new universities, and this gives a chance for the present study to reveal some characteristics of value-added of universities, while much research in this area is based on data from old universities.

The data obtained are in the form of individual student records and broken down by course and year. They demand a lot of editorial work such as changing variable type (from string to numeric), recording variables, matching individual students' entry qualifications with their degree results, and selecting useable data. Not surprisingly, some research in this area has had a team to

clean and edit raw data to make data ready to use (e.g. Cuttance, 1987). This may be due to the reason that data were not collected specifically for the purposes of research of this kind.

3.3 Methodology of the qualitative study

3.3.1 The objective of the qualitative study

The aim of the qualitative study is to assess whether the value-added measurements developed in the quantitative study can be used as a performance indicator. This study will focuses on investigating academic staff views on problems with using the value-added results to indicate the quality of teaching by employing a qualitative method.

The objective above is guided by the following considerations.

Firstly, academic staff play a crucial role in teaching and learning, and therefore they are most knowledgeable about the reality of how a student starts with a certain entry qualification at entry and comes out with a certain class of degree at exit, in other words, they are most knowledgeable about the 'departmental black box' which delivers the teaching and which is ultimately responsible for degree classification decisions. Secondly, value-added is intended to be used as a measure of academic staff performance, therefore it is important to know what academic staff think about it. Do they think that value-added results can reflect the quality of their teaching, or what their concerns are? Do they accept it? One may argue that whether academic staff like it or not, performance indicators are here to stay. Nevertheless, I believe, understanding the academic staff view would be significant in terms of how to construct a value-added performance indicator and how to use value-added as a performance indicator. Furthermore, if the academics do not accept value-added as a performance indicator of their teaching, it would undermine the effectiveness of this performance indicator in practice.

3.3.2 The research method chosen

Interviewing is a very good way to access people's perceptions, meanings, definitions of situations and constructions of reality (Punch, 1998). There are

many different types of interviews. Fontana and Frey (1994) suggest that interviews can be categorized into three types in terms of the degree of structure: structured interviews, semi-structured and unstructured interviews.

As far as the present study is concerned, the semi-structured interview would be a good tool for collecting data, because the purpose of the research is to investigate academic staff views on whether the value-added measurement can be used as a performance indicator of teaching. In a structured interview, interview questions have been formulated and standardised before the interview, the respondent is expected to answer in terms of the interviewer's framework and the interview itself does not attempt to go to any great depth (Fontana and Frey, 1994). For example, if we wish to find out how many people oppose a nuclear repository, structured-interview is the best tool, and we can quantify and code the responses and use mathematical models to explain our findings (Frey, 1993). However, given the purpose of the interview in the present research, the interviewer does not intend to or cannot provide the categories or frameworks for answering the interview questions. The interview questions are necessarily open-ended, and the interview attempts to go in-depth. Therefore, a structured-interview is not suitable for the present research.

An unstructured-interview is also not suitable for the present research, because in an interview questions are not pre-planned and standardized, but instead there are general questions to get the interview going. Specific questions will then arise from the respondent's reaction to the broad issue raised by the inquirer. The interviewer does not seek normative responses (Fontana and Frey, 1994). For example, if we wish to know and understand about the lives of Palestinian women in the resistance (Gluck, 1991), we need to interview them at length and in depth in an unstructured way. However, the present research attempts to focus on investigating the academic staff's views on the value-added measurement. The discussions in the interview have to be limited to certain parts of the academic staff's experience. The academic staff's views are more likely to be expressed in a relatively openly designed interview situation.

A semi-structured interview will allow the interview not only to focus on the issue of whether value-added results reflect the quality of teaching, but also to probe more deeply to gain a more thorough understanding of the academic staff's opinions and reasons behind them on the value-added measurement. Therefore a semi-structured interview is chosen as the tool to collect data in this research.

3.3.3 The semi-structured interview

3.3.3.1 Objectives

The interviewing employed in the present study was directed by the following objectives: (1) to obtain academic staff's comments on the value-added measurement proposed, (2) to find out academic staff's concerns if the value-added measurement is to be used as a performance indicator of teaching.

The interview was intended to meet two criteria during the design of the interview guide and the implementation of the interview itself. Firstly, the interview had to proceed with open-mindedness within the areas of the research questions. The interviewer guidance should be minimal. Secondly, the interview had to focus on specific comments which are based on the particular members of staff's specific teaching experience, in order to be involved in a 'departmental black box' where teaching is delivered and degree classifications are decided, and to prevent the interview from remaining on the level of general statement.

3.3.3.2 The participants

The three schools from an institution, Engineering, Business Studies, and Health and Community Studies were chosen as a sample. The purpose of choosing the three schools is that quantitative data were obtained from those same three schools, therefore the qualitative study can be more directly linked to the quantitative study, and the results of the quantitative can be used as an example in the qualitative study. Another reason for choosing the sample is that

the institution chosen for the qualitative study is close to the town where I live. This would make the interview feasible in terms of time and cost.

All members of staff who taught / teach on the courses sampled were contacted. 18 of about 30 members of academic staff agreed to be interviewed. Most of the interviewees are senior lecturers or principal lecturers and have many years teaching experience in higher education. Particularly, some of them have taught the courses sampled for over 5 years. This means that they have experienced the teaching period (88/89-92 /93 academic years) measured in the quantitative study. Therefore the interviewees know what has been going on in the teaching.

3.3.3.3 The procedure of the interview

As the interview was semi-structured, the main questions were open and they were primarily constructed according to the objectives of the research along with previous research and public debate on performance indicators.

Several researchers (e.g. Sizer, 1979, Ewell and Jones, 1994, and Cave et al, 1997) and the CVCP / UGC working group suggested a number of standards, questions or criteria to test potential performance indicators. The main purpose of their suggesting these criteria are to standardize existing practice since some of the information used by universities as performance indicators may not fall within the definitions of performance indicators. Nevertheless, some of these criteria are useful to help the present study to formulate the focus of the interview. For example, Cave et al. proposed that 'ambiguity' and 'manipulability' should be the criteria to evaluate performance indicators. In terms of ambiguity, it is suggested that a performance indicator should be able to identify a high or low value of the indicators as unambiguously favourable or unfavourable. If a performance indicator can be manipulated by the individual or body that it is intended to assess, its value is reduced. These criteria have helped to formulate some of the interview questions.

Before the interview, the relevant information was sent to each of the interviewees. The information explains the background of the research, the purposes of the interview, the value-added model developed, the results of using the value-added model to analyse the teaching performance of the particular course which they teach. It may worth noting that several of members of staff mentioned during the interview that they were very interested to see this quantitative analysis of the particular course which they have taught. Most of the interviewees had read the information before the interview. To ensure that the interview would achieve the objectives, the interview guide was tried out with one member of academic staff whom I knew and he gave me some comments about the interview which included checking of some important items of vocabulary used in the interview, and I noted his interpretation and reaction to my questions. I then evaluated the whole procedure of the interview.

To make interviewees feel at ease and free to talk, the interview started with introducing myself as a research student, emphasizing that what they are going to say in the interview would only be used for the research, that their personal identity would be kept confidential, and explaining the purpose of the interview emphasizing that any kind of comments concerned with the value-added measurement was welcomed. Their permission to use a tape recorder was obtained. The interviewees then were asked if they had had a chance to read relevant information sent to them before the interview and if they had any questions about it, usually followed by a brief explanation of the model again. During the interview, the interviewees were asked to give any comments on the value-added measurement and on the results of using the model to analyze performance in the courses they teach. This question with probing often led some of the interviewees to start to talk about all their concerns about using value-added as a performance indicator of teaching. The interview usually lasted 30 -60 minutes. The interviews were all tape-recorded and their accounts were transcribed verbatim.

3.3.4 How the data were analyzed

To obtain a general understanding of what is there, all the interview transcripts were thoroughly read several times. The main objective of the analysis of the interview was to establish respondents' categories which illustrated academic staff's concerns with using the value-added measurement as a performance indicator of teaching in HE.

In practice, the following specific procedures were undertaken to process the data and to delimit the conceptual constructs.

Stage 1, the points made by individual interviewees were placed on the analysis sheet. The analysis sheets also recorded this information: number of sheet, name of interviewee, school which interviewee is from, number of paragraph where the specific points were made, categories (to be used at later stage of analysis). An example is displayed below.

No. 1 Name: D. Adam School: A

Paragraphs	Points	Categories
1-3	*We are adding things to these students' education experiences, skills that you can not quantify in terms of upper second, lower - second etc.*	
4-5	....	

Stage 2, categories were named. This was done by

- firstly, marking all the transcripts with coloured stripes down the left hand margin;
- secondly, cutting up the paragraphs which deliver the points,
- thirdly, putting the most and least relevant paragraphs together as a group and giving provisional categories to each group;
- fourthly, after reading the transcripts of individual groups and referring to the previous literature and the public debates, naming the categories more firmly and extracting some categories from the data;

- fifthly, reviewing categories, dealing with overlap categories, making revision of some categories,
- sixthly, categories were examined for identification of possible relationships between them, finally putting name of categories onto the analysis sheets.

Stage 3, points, groups, the names of categories were all checked again. Some modifications were made. The categories which were ultimately delimited were considered for ways in which logical sequences could be identified.

3.3.5 Questions of validity

Validity has long been a key issue in debates over the legitimacy of qualitative research. 'Validity refers to a number of different things, all to do with whether the measuring instrument is actually measuring the variable which it was intended to measure' (Ashworth, 1997). To enhance the credibility of qualitative research in the fashion of naturalistic inquiry, Lincoln and Guba (1985) suggested five tasks to be undertaken. Lather (1986, and 1993) outlined a number of validity criteria. Maxwell (1992) working within a realistic frame of reference derived five different forms of validity from practices of qualitative researches. Cohen and Manion (1989) suggested that perhaps the most practical way of achieving greater validity is to minimize the amount of bias as much as possible. They then identified some sources of bias when interviewing is used as a research instrument.

In the case of this research, the following aspects have been considered in order to enhance the internal validity of the present qualitative study.

Openness of interview procedure

The purpose of the present qualitative study is to investigate what academic staff's concerns are if the value-added measurement is to be used as a performance indicator of teaching in higher education. The interview process is open to the respondents' frame of reference. Given the aim of openness to interviewee framing, the strategy used is to ask an initial 'access' question and then follow this up by non-directively facilitating the interviewee's elaboration

and expansion of the viewpoint they have started to express.

Although the present qualitative study is to investigate academic staff's views on whether the value-added measurement, which I developed, can be used as a PI of teaching, I am aware of the difficulties with measuring value-added in higher education and the controversial introduction of performance indicators into higher education. The purpose of the study is not to provide a ready-made or off-the-shelf performance indicator. The purpose is to explore the measurement of value-added in higher education and to discover difficulties of using it. In fact, I was expecting negative comments on the value-added measurement which I developed. This presupposition has helped avoid a tendency to seek answers that support my preconceived notions or see the respondents in my own image.

Additionally, it is unlikely that I, an interviewer, as a Ph.D. student, would influence the concerns, attitudes or opinions of the interviewees, academic staff on whether a measurement is to be used as a performance indicator of their teaching. However, on the other hand, to some extent, since I was an 'outsider' of the interviewees, I had the advantage that I do not take for granted things which an 'insider' may do.

Contextual nature of the interview

The interview presented interviewees with actual analyses by using the value-added measurement rather than posing questions in general or abstract terms. Such *contextual nature* (Kitwood, 1980) of the interviewing attempted to tap into an interviewee's value reasoning in action, as he or she attempted to grapple with a specific problem instance or context, to promote ecological validity.

Descriptive validity

Since English is my second language, to assure myself that the meanings of the questions were crystal clear, and that my language-use fitted into the culture of academic staff in HE, the interview guide was tried out with a member of academic staff before the interview. On the other hand, to avoid

misunderstanding what the interviewees are saying, whenever I was in a little doubt about them, I would check with the interviewee or my colleagues (British). This has been done throughout the process of the interview and data analysis. Checking transcriptions with tapes was done whenever it was necessary; above all this would enhance the descriptive validity (Maxwell, 1992) of the interview.

In terms of the external validity of the present study - the extent to which the understandings generated from this study might be generalisable and applicable to other contexts - strictly speaking, the outcome of the qualitative research is not generalisable. This qualitative study is based on interviews with some members of academic staff from three subjects in just one institution and their comments are about a particular value-added measurement, therefore the external validity of this study is limited. However, it is likely that many of the findings may be applicable to other value-added measurements which are based on entry qualifications and degree results. Some of understandings generated from this study may be useful in the development and use of other performance indicators in HE.

However, it is important to re-emphasize that the post-positivist takes the view that 'realities' are multiple, and in a social world, what we can find are multiple truths, and multiple understandings. Ashworth (1997) argued that ' I cannot produce a neat schema for doing valid and reliable qualitative work within the positivist understanding of qualitative research.' He believes that such a schema is impossible to lay down exhaustively.

Although some actions have been taken to enhance 'validity', the interview data provided in this study are inevitably a specific joint product of me, as a researcher with my cultural standpoint, and the members of staff, as interviewees within their culture. As Ashworth (1997) noted, interview data were

> *the record of a joint process by which two individuals have negotiated a 'fusion of horizons', or each came to appreciation of a way in which their initial fore-understanding was transformed by bringing it up against the*

others' conversation.

Although the above view on interview data is within the post-positivist framework, and the present research is based on positivism, I believe that that is a reality because I, as interviewer and the academic staff as interviewees are not machines.

CHAPTER FOUR
MEASURING VALUE-ADDED AT COURSE LEVEL

4.1 Introduction

In the last chapter, the framework of the proposed value-added measurement has been discussed. This chapter will use the value-added method to analyze actual data in order to explore the strengths and weakness of the model.

This chapter is divided into four sections. In the first section, the characteristics of the population of the courses were examined first. The intention of this is to identify major groups of students' entry qualifications and provide some background information for measuring and understanding value-added. Following the first section, the sections two, three, and four report the results of the measurements of value-added by using the method developed in the methodology chapter. As described in Chapter 3, the model can be used to measure value-added based on direct comparisons between individual cohorts. So sections two and three report the results based on this measurement. Section four reports the results of value-added based on comparisons between the average of previous cohorts and the current cohorts. Finally, the last section will draw conclusions about whether the method can be used to measure value-added at course level when it is applied to actual data, what the value-added results are, and the strengths and weaknesses of the model.

The set of data first obtained includes information from three schools: Business, Health and Community Studies, and Engineering. Since the number of students on some of the courses is small, some of the courses are combined into one course. The detail will be presented in the following sections.

The data were obtained from the University Student Records of an institution in the form of individual student records. The data included information on student registration number, entry qualifications, A level points, degree results, age, gender, mode of attendance from year one to year three or year four. Since the content of the course for sandwich and full-time students was exactly same,

and the major difference between these two groups of students was that the sandwich students spent one year in placement, these two groups of students were put together when calculating value-added.

It should be noted that the last year (year four, or year three) data did not include the information on entry qualifications. However the Statistical Package for Social Science (SPSS) software was used to match entry qualifications with degree results by individual students by using ID numbers.

4.2 Characteristics of the population of the courses

4.2.1 Business Studies Course

The characteristics of the population of the Business Studies students are shown in Table 4.1. The course has a large number of students – over a hundred per year. The students entered the course of study with nine different types of entry qualifications over the five year period. However, there is a consistency over the whole period (88- 91) in that the majority of students' entry qualifications fall into two major groups, A level and BTEC.

Few students graduated with a first class degree. The majority of them obtained upper second or lower second class degree. There was a small percentage of students who had results 'referred' and 'deferred'.

From 1988 to 1990, the number of female students was slightly higher than male students, but in 1992, there were more male students than female students. Mature students were a minority of the course members.

4.2.2 Applied Social Studies course

The data obtained for the School of Health and Community Studies included the two courses, BA (Hons) Applied Social Studies and BA (Hons) Applied Social Studies with Diploma in Social Work. The cohorts for which records were obtained were only four cohorts: 1987, 1988, 1990, and 1991 entry cohorts.

Table 4.1 Business Studies Course

	1988 cohort		1989 cohort		1990 cohort		1991 cohort		1992 cohort	
No. of students	108		191		162		127		114	
Students with:	No.	%	No.	%	No.	%	No.	%	No.	%
A level	67	62	146	76	118	73	77	61	101	88.6
BTEC	11	10	18	9	24	15	24	22	7	6.2
Access	4	4	4	2	0	0	3	2	1	0.9
O level	1	1	0	0	1	0.6	3	2	0	0
no formal qualifi.	0	0	0	0	2	1.2	0	0	0	0
other qualifi.	0	0	0	0	1	0.1	0	0		1.8
European Bacal.	0	0	0	0	1	0.1	0	0	0	0
degree ordinary	0	0	0	0	0	0	1	0.7	0	0
oversea qual.									2	0.9
missing	25	23	22	12	15	9	19	15	4	1.8
Students with:	No.	%	No.	%	No.	%	No.	%	No.	%
first	1	1	0	0	1	0.6	0	0	0	0
upper second	46	43	108	57	44	27	30	24	36	31.6
lower second	15	14	38	20	65	40	44	35	72	63.2
Third	0	0	0	0	0	0	0	0	4	3.5
Unclass.	0	0	0	0	0	0	0	0	1	0.9
referred	0	0	7	4	8	5	10	8	0	0
deferred	0	0	4	2	4	2.5	7	5.5	0	0
missing	46	43	34	18	40	25	36	28	1	0.9
Sex	No.	%	No.	%	No.	%	No.	%	No.	%
Male	33	31	74	39	75	46	62	49	66	57.9
Female	50	46	96	50	73	45	48	38	48	42.1
missing	25	23	21	11	14	9	17	13	0	0
Age	No.	%	No.	%	No.	%	No.	%	No.	%
Mature	20	19	23	12	20	12	22	17	8	7
Non-mature	63	58	147	77	128	79	88	69	106	93
missing	25	23	21	11	14	9	17	13	0	0

The data for 1989 were missing for some reason when the data were provided to me. Some important information (e.g. A level points) for the data of the 1987 cohort was missing, therefore they cannot be used either. The students included full-time and sandwich students. Since the number of students for the each course was small (under 50 students), the two courses were combined as one course.

The characteristics of the Applied Social Studies students' population are shown in Table 4.2. Like the Business Studies course, there is a consistency over the whole period (89, 90, and 91) in that the majority of students' entry qualifications fall into two major groups, but the two major groups here are A level and Access or Conversion Courses.

Few graduates obtained a first class degree, most of them graduated with upper second and lower second class degrees. No students obtained the third class degree. Over 50 percent of graduates are women. There has been a change in the age profile. In 1988, the number of mature students was around double the number of non-mature students. In contrast, in 1990, there were 49 non-mature students out of a total 78 students. However, the number of mature and non-mature students were about equal in 1991.

4.2.3 Engineering courses

The data obtained for the school of Engineering were from four courses: B.Eng. (Honours) Computer Aided Engineering and Design, B.Eng. (Honours) Engineering with Business Studies, B.Eng. (Honours) Materials Engineering, and B.Eng. (Honours) Mechanical Engineering. Data for four cohorts for these four courses were obtained. They are the 1989, 1990, 1991, and 1992 cohorts.

The overall aims of the four courses are to produce graduates who can be employed as engineers with the engineering applications emphasis of the course. The course content of the first year is common to the four courses and provides a base knowledge of engineering. Year three is professional work-based experience. Since numbers of the students on each of the four courses

Table 4.2 Applied Social Studies Course

	1988 cohort		1990 cohort		1991 cohort	
Number of students	76		84		107	
Entry qualifications	No	%	No.	%	No.	%
A level	49	64.5	62	73.8	72	67.3
Access /conversion	17	22.4	10	11.9	21	19.6
BTEC	5	6.6	5	6.0	5	4.7
O level	1	1.3	4	4.8	2	1.9
Open University credits	1	1.3	0	0	2	1.9
professional qualifications	1	1.3	0	0	0	0
other qualification	0	0	1	1.2	2	1.9
no formal qualification	0	0	1	1.2	1	0.9
other university credits	0	0	1	1.2	0	0
missing	2	2.6	0	0	0	0
Degree results	No.	%	No.	%	No.	%
first	2	2.6	3	3.6	3	2.8
upper second	48	63.2	44	52.4	58	54.2
lower second	26	34.2	34	40.5	30	28
referred	0	0	1	1.2	9	8.4
deferred	0	0	1	1.2	4	3.7
missing	0	0	1	1.2	3	2.8
Sex	No.	%	No.	%	No.	%
Male	19	25	19	23	22	20.0
Female	57	75	65	77	86	80.4
Age	No.	%	No.	%	No.	%
Mature (age>=21)	48	63.2	34	40.5	52	48.6
Non-mature (age<21)	28	36.8	50	59.5	55	51.4

are very small, the four courses were combined as one. The course leaders of the four courses have no objection to combining these four courses.

The number of students on the four courses varies over the four years. There were about 100 students on the courses in 1989, and then the number of the students dramatically dropped to about 50 in the following year. From 1991 to 1992, the number of students had been about 80 (See Table 4.3).

Over the whole period (89, 90, 91 and 92), about 80 percent of students' entry qualifications fell into two major groups, A level and BTEC. The propotion of students with A Level qualifications slightly decreased over the years, while the percentages of BTEC students and Access or Conversion Course students appeared to be slightly increased. This may reflect the policies of widening access to higher education and encouraging diversity in higher education. In the future, this tendency may be changed. For example, the Engineering Council has announced that students who are to become an accredited chartered engineers must have 24 A level points.

Over the four years, the percentage of students who obtained a first class degree was under 10 percent. This proportion is consistent with the national average proportion of first class degrees in the Engineering subject for all new universities (Chapman, 1994). Students who obtained the upper second and lower second class degrees have been the two dominant groups (about 80 %).

Non-mature and male students had dominated the four courses over the four years. This may due to the stereotype of the subject. The numbers of mature students on the courses reached 22 (30%) in 1992, while female students were still less than 20 percent of all the students.

4.3 Value-added based on direct comparisons between individual cohorts

This section focuses on measuring value-added based on direct comparisons between individual cohorts. The analysis in the last section of this chapter

Table 4.3 Engineering Courses

	1989 cohort		1990 cohort		1991 cohort		1992 cohort	
Number of students	107		57		80		74	
Number of students with	No.	%	No.	%	No.	%	No.	%
A level	77	72	39	68	51	64	33	45
BTEC	28	26	15	26	24	30	28	38
Access / conversion course	1	1	2	4	1	1	7	10
other overseas qualification	1	1	0	0	2	3	1	1
O level	0	0	1	2	0	0	1	1
other qualification	0	0	0	0	2	3	2	3
missing data	0	0	0	0	0	0	2	3
Number of students with	No.	%	No.	%	No.	%	No.	%
first	10	9.3	3	5	5	6	6	8
upper second	42	39	30	53	41	51	33	45
lower second	45	42	13	23	25	31	32	43
third	2	2	0	0	1	1	0	0
unclassified	1	1	0	0	0	0	1	1
fail	0	0	1	2	1	1	0	0
referred	4	4	1	2	5	6	0	0
deferred	1	1	1	2	1	1	0	0
missing	2	2	8	14	1	1	0	0
Sex	No.	%	No.	%	No.	%	No.	%
Male	93	87	48	84	74	93	64	87
Female	13	12	9	16	2	6	10	14
missing	1	1	0	0	0	0	0	0
Age	No.	%	No.	%	No.	%	No.	%
Mature(age>=21)	12	11	8	14	12	15	22	30
Non-mature(age<21)	95	89	49	86	68	85	52	70
missing	0	0	0	0	0	0	0	0

reveals that although students on the three courses started with diverse entry qualifications, the majority of the students' entry qualifications were concentrated on several major groups (see table 4.4). Therefore the performance of the courses may be represented by the performance of their major groups of students. We will measure value-added for these major groups of each course first, then on the basis of value-added for these individual groups, the over all value-added for the courses are calculated.

Table 4.4 The major groups

Course	Major group of entry qualifications
Business Studies	A level, BTEC
Applied Social Studies	A level , Access
Engineering	A level, BTEC

4.3.1 Value-added for the major groups

4.3.1.1 A level group

As described in the methodology chapter, the calculation of value-added for A level students involves three steps:

- The first step calculates mean A level points for each cohort and then ranks the mean of the A level points (input rank).
- The second step calculates percentage of good degrees for each cohort and then ranks the percentage (output rank).
- The final step, takes input rank and subtracts output rank. The results are the value-added scores (see Table 4.5).

Within the Business Studies Course, on the whole, from 1988 to 1992, there was no clear upward or downward trend in value-added (see figure 4.1). Nevertheless, figure 4.1 shows a dramatic decrease in value-added between the 1989 and 1990 cohorts, and between the 1991 and 1992 cohorts. It should be noted that the figure uses unequal intervals on the axes, since it is just for illustration purposes. From table 4.5, it is can be seen that the 1990 cohort recruited students with a mean A level points of 17.32, but only 43 per cent of

them obtained a good degree, while 75 per cent of the 1989 cohort achieved good degrees and their mean A level point score was 15.93. Similarly, only 32

Fig. 4.1 Trend in value-added for A level group of Business Studies course

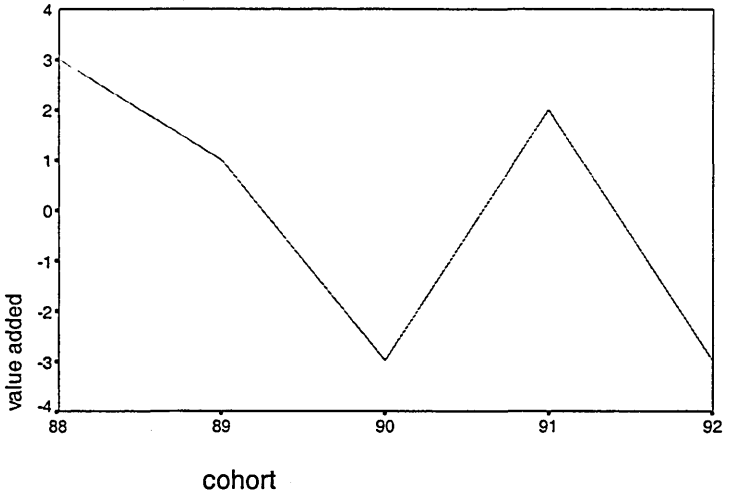


Table 4.5 Value-added for A level group of Business Studies Course

	1988 cohort	1989 cohort	1990 cohort	1991 cohort	1992 cohort
Number of students:	52	117	82	45	101
Mean A level points:	15.69	15.93	17.32	15.02	17.11
Input rank :	4	3	1	5	2
Percentage of good degrees:	79%	75%	43%	47%	32.7%
Output rank :	1	2	4	3	5
Value-added:	3	1	-3	2	-3

per cent of the 1992 cohort graduated with good degrees, which is the lowest in the five cohorts and which is less than half of the proportion of students who achieved good degrees in the 1988 cohort, but the mean A level points for this cohort are 17.11, which ranks 2 in the five cohorts. Therefore questions are raised here: Why did the Business Studies course recruit better qualified entrants in the 1990 and 1992 cohort, but there were fewer students on the course achieving good degrees in these two cohorts? The answers to these questions will be discussed in the second part of the research (chapter 6).

Within the Applied Social Studies course, from Table 4.6, it can be seen that the academic attainment (i.e. mean A level points) of the three cohorts of students was different when they entered the university. For example, the 1990 cohort had the highest mean of A level points, 14.08, while the mean of A level points for the 1988 cohort was 11.76. After three or four years, about seventy-three percent of the students of the 1988 cohort obtained good degrees, but only fifty-three percent of students of the 1990 cohort obtained good degrees. Therefore the 1990 cohort appeared to have much less value-added than the 1988 cohort did. Is this because the teaching quality of the 1990 cohort was poor? The value-added for the 1991 cohort is 0 which means that the difference in academic attainment between the 1991 cohort and other cohorts had not changed.

Table 4.6 Value-added for A level group of the Applied Social Studies Course

	1988	*1990*	*1991*
number of students	49	62	68
mean of A level points	11.76	14.08	12.09
rank of input	3	1	2
percentage of good degrees	73.47%	53.3%	68.29%
rank of output	*1*	*3*	*2*
value-added	2	-2	0

Within the Engineering courses, Table 4.7 shows that over the four year period (1989-1991) there was no upward or downward trend in value-added for A level students on the Engineering courses. The mean A level points for the 1991 cohort was 7.98 which was the lowest of the four cohorts, but over 63 per cent of the students achieved good degrees, which is the highest in the four cohorts. Is teaching quality of the Engineering courses for the 1991 cohort the best in four cohorts? Why did the mean A level points of the Engineering courses improve from 1991 (mean A level points = 7.98) to 1992 (mean A level points = 10.38), but the proportion of good degrees decrease?

Between the three courses, on the whole, clearly, mean A level points of the three courses during the period of 1988 -1992 were different, mean A level points of students of the Business Studies Course was the highest, 16.21, the

Applied Social Studies Course was second with 12.64, and the Engineering

Table 4. 7 Value-added for A level group of Engineering Courses

cohort	1989 cohort	1990 cohort	1991 cohort	1992 cohort
No. of students with A level:	74	38	47	34
Mean of A level points:	9.41	10.11	7.98	10.38
Input rank:	3	2	4	1
Percentage of good degrees:	52.6%	57.9%	63.8%	58.8%
Output rank:	4	1	2	3
Value-added:	-1	1	2	-2

courses were the lowest, about 9.47. Nevertheless, on the whole, the proportion of good degrees on the Business Studies course was the lowest at 55.34%, the Engineering courses had 58.28 % which was higher than Business Studies percentage, and the Applied Social Studies Course was 65.02 %. Therefore the question is raised as to why the Business Studies Course recruited better qualified entrants than other two courses did, but it produced fewer graduates with good degrees? Does this indicate that quality of teaching on the Business Studies Course was poorer than the other two courses?

A number of statistical techniques were used in order to test whether the differences in mean A level points between those cohorts of the three courses, and differences in percentages of good degrees between them were statistically significant. The results for each course are reported below.

For the A level group of the Business Studies Course, the analysis of variance (ANOVA) was used to test whether the differences in mean A level points between the five cohorts (1988-92) are statistically significant, because mean A level points are represented by an interval variable, and distributions of A level points are approximately normal (See Fig 1.1 in Appendix 1). The ANOVA test results indicate that the differences in mean A level points between the five cohorts are statistically significant at the 5 per cent level (F value = 4.36). This means that there are significant differences in mean A level points between the five cohorts, but the ANOVA test results did not indicate between which cohorts the significant differences exist. Thus a Tukey – HSD (honestly significant

difference) test was used to further test the results. The Tukey – HSD test allows us to compare each pair of the five cohorts to see if the difference in each case is significant.

Table 4.8 displays the test results. It suggests that the mean A level points score for the 1991 cohort was statistically significantly lower than that for the 1992 cohort and the 1990 cohort . The differences in mean A level points between the other pairs of cohorts are not statistically significant.

Table 4.8 Tukey - HSD test result for the Business Studies Course

Mean A level	Cohort	91 cohort	88 cohort	89 cohort	92 cohort	90 cohort
15.02	91 cohort					
15.69	88 cohort					
15.93	89 cohort					
17.11	92 cohort	*				
17.31	90 cohort	*				

Note * = significant at the 5% level.

Since the percentage of good degrees is represented by a nominal variable, the chi-square was used to test the differences in this between the five cohorts. The result suggested that the differences in percentage of good degrees between the five cohorts are significant at 1 percent level.

For the A level group of the Applied Social Studies course, since the distributions of A level points for the three cohorts are approximately normal (see fig. 1.2 in Appendix 1), the ANOVA test was used to test if differences in mean A level points between the three cohorts (88, 90, and 91 cohorts) are statistically significant. The result indicates that the differences are statistically significant at the 1 per cent level (F value = 4.49; D.F. = 2). A Tukey – HSD (honestly significant difference) test further suggests that the mean A level points for the 1990 cohort was statistically significantly higher than that for the 1988 cohort and the 1991 cohort. However, a chi-squared test carried out on the observed differences in percentages of people obtaining good degrees between the three cohorts does not reach statistically significant levels (5%).

For the A level group of the Engineering courses, the non-parametric Kruskal - Wallis test has been used to test the observed differences in mean A level points between the four cohorts (89 - 92 cohorts) since the distributions of A level points of the cohorts are not Normal (see fig.1.3 in Appendix 1). It shows that the differences in mean A level points between four cohorts are statistically significant at the 5% level (value = 8.3, DF= 3). Nevertheless, the analysis of variance (ANOVA) is also used to test the significance of the differences in mean A level points, because the analysis of variance is robust. The ANOVA suggests the different result that the differences in mean A level points are not statistically significant. In this case, it would be safe that both test results are taken into account. The chi-square test has shown that the observed differences in percentages of good degrees for the four cohorts are not statistically significant.

4.3.1.2 Value-added for BTEC group

BTEC students were one of the major groups of students on the Business Studies Course and the Engineering courses. Since the information on grades of BTEC students was not available, the academic attainments of all students with BTEC were taken as the same. Therefore the value-added for BTEC is measured by comparing differences in degree results (see Table 4.9).

It can be seen from Table 4.9 that value-added for BTEC students on the Business Studies course increased from 1988 to 1989, then, like the A level group of the course, it decreased dramatically in 1990. The percentage of good degrees decreased from 82 percent to 35 percent. It slightly increased in 1991 and decreased to 29 per cent in 1992.

With BTEC students on Engineering courses, value-added had been fluctuating during the period. The 1990 cohort appeared to have most value-added. From 1990 to 1992, there was a downward trend in value-added (see table 4.9).

However, chi-squared tests suggest that the observed differences in the percentage of good degrees between the different cohorts of both courses are not statistically significant. Nevertheless, it can be seen that the differences in

Table 4.9 Value-added for students with BTEC of the Engineering Courses and the Business Studies Course

	1988 cohort	1989 cohort	1990 cohort	1991 cohort	1992 cohort
Percentage of good degrees for Business Studies Course	67%	82%	35%	46%	29%
Percentage of good degrees for Engineering courses	*N/A*	44%	64%	55%	50%

percentages of good degrees between the 1989 and 1992 cohorts of the Business Studies course is quite big. Therefore insignificant results may be due to the fact that the numbers of students with BTEC for each cohort are very small (see table 4.10). The range of number of students with BTEC on Engineering courses was 14 - 28, although the BTEC group was a second major group of the course.

Table 4.10 Number of students with BTEC entry qualifications on the Business studies course

cohort	1988 cohort	1989 cohort	1990 cohort	1991 cohort	1992 cohort
Total no. of students on the course:	108	191	162	127	114
Number of students with BTEC:	9	11	17	13	7

4.3.1.3 Value-added for Access group

Access students appeared to be a major group of students on the Applied Social Studies course. The calculation of the value-added for Access students

is the same as that for BTEC students. The percentage of good degrees for the 1988, 1990 and 1991 cohorts are 47.1, 66.67, and 61.1. Access students of the 1990 cohort had the most value-added, while the 1988 cohort of the course had the least value-added. The difference in percentage of good degrees between the 1990 and 1991 cohorts is much smaller than that between the 1990 and 1988 cohorts. The question may be asked as to why there was a much lower proportion of students from the 1988 cohort achieving good degrees.

A chi-squared test indicates that the differences in percentages of good degrees between the three cohorts are not statistically significant. Again, the number of students with Access entry qualification is very small (between 9 and 18). This may be the reason for the insignificant results.

4.3.1.4 The comparison of value-added between A level students and BTEC students on the Engineering courses

Although A level and BTEC are seen as two different entry qualifications and it is found difficult to give scores relating the two entry qualifications, as noted earlier, currently, the Engineering Council and professional bodies have announced that they intend to develop 'accredited chartered engineer programs' and students who study these programs must have 24 A level points, while Engineering courses at some institutions usually recruit students with BTEC and have difficulty in recruiting students with high A level points. Therefore the comparisons of value-added between A level students and BTEC students become more important. The comparison of value-added between these two groups will focus on the Engineering courses.

Because of the difficulty with scoring these two different entry qualifications, the comparison between these two groups of students may be made under the following assumptions.

First assumption: students with A level entry qualifications and students with BTEC have similar capability to do their degree course. Table 4.11 shows the

differences in percentage of good degrees between A level and BTEC students on the Engineering courses. A level students appeared to do better than BTEC students from three cohorts (1989, 1991, and 1992), but the 1990 cohort of BTEC students on the whole achieved better degree results than A level students. The observed differences in the proportion of good degrees between A level and BTEC students have been tested by using a chi-squared test. It is discovered that the differences are not statistically significant. This may be interpreted as meaning that BTEC students have achieved the same or similar academic attainment as the A level students. This may indicate that BTEC students have the same opportunity to achieve a good degree result as the A level students. Nevertheless the number of students with BTEC for some cohorts is very small. This could be an explanation of the statistical insignificance of the observation.

Table 4.11 The proportion of good degrees obtained by A level and BTEC students on Engineering courses

	1989 cohort	1990 cohort	1991 cohort	1992 cohort
Percentage of good degrees obtained by *A level* students	52.6%	57.9 %	63.8%	58.8%
Percentage of good degrees obtained by *BTEC* students	44%	64%	55%	50%

BTEC entry qualifications are usually regarded as a less favourable entry qualification than A level in university admission. This may be interpreted as meaning that the academic attainments or ability of students with BTEC is lower than that of A level students. Under this assumption, it may concluded that the BTEC students of the Engineering courses gained more value than the A level students during 1989-1992.

4.3.2 Value-added for the courses

On the basis of the above results on value-added for the major groups, the value-added for the courses can be measured by calculating the sum

(weighted) of value-added for the individual groups, as described in the methodology chapter. For example, the calculation procedure for BA Business Studies course is as follows:

VA = VA *for A level* x Weight *for A level* + VA *for BTEC* x Weight *for BTEC*

Weight *for A level* = Total no. of *A level students* / Total no. of students of two groups

Weight *for BTEC* = Total no. of *BTEC students* / Total no. of students of two groups

The results are demonstrated in Table 4.12 and are also presented more effectively by Fig. 4.2, which uses unequal intervals on the axes for the illustration purpose. It is suggested that there was no upward or downward trend on value-added for the three courses over the period, 1988-1992.

Table 4.12 Value-added for the courses

cohort	1988 cohort	1989 cohort	1990 cohort	1991 cohort	1992 cohort
VA for the Business Studies Course:	2.87	1.00	-2.12	2.13	-2.00
VA for the Applied Social Studies courses:	1.39	N/A	-1.39	0	N/A
VA for the Engineering Courses:	*N/A	-.22	-1.16	1.89	-0.52

*note * = not applicable*

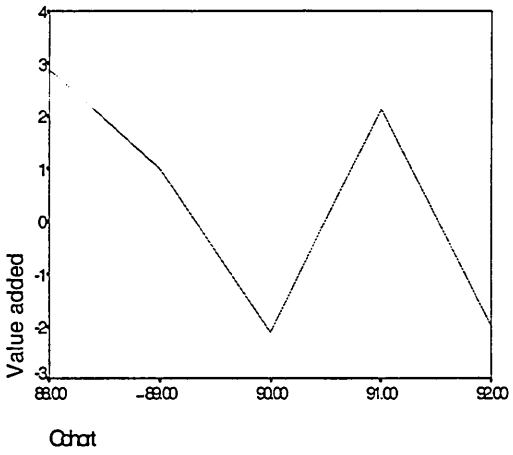
The 1990 cohorts of the three courses appeared to have had least value-added. The question should be asked why this cohort of all three courses did more poorly than other cohorts. Attention should be drawn to this cohort for all courses. Is there any change in examination policy or recruiting policy at the institution as a whole? The 1988 cohorts of the Business Studies and Applied Social Studies had most value-added.

4.4 Value-added based on comparisons between average of previous cohorts and current cohort

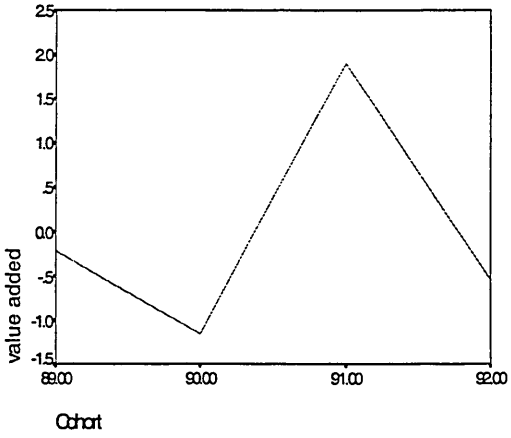
Value-added for the current cohort can be measured by a comparison in entry

Fig 4.2 Value-added for the three courses

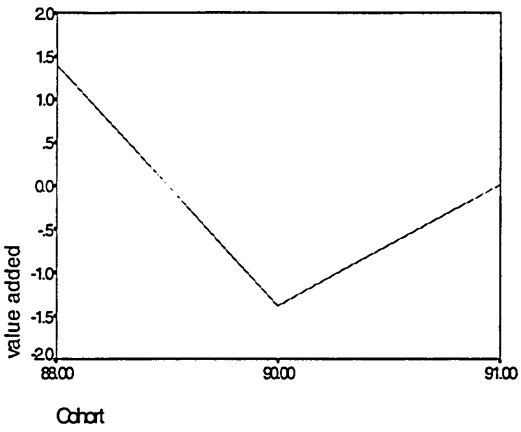
Business Studies course



Engineering courses



Applied Social Studies



qualification and exit qualifications between the current cohort and average of previous cohorts. This method was used to measure value-added for the 1992 cohort of A level students on the Business Studies Course as compared with the previous four cohorts, 1988-1991. It was found that the mean A level points for the 1992 cohort (17.09) was higher than the average of the previous four cohorts (16.41), but the proportion of students who gained good degrees for the 1992 cohort (33%) was smaller than the average of the previous four cohorts (59%). Therefore the 92 cohort had less value-added as compared with the average of the previous four cohorts. This result is consistent with the results presented in table 4.5. The analysis of variance test suggests that the difference in mean A level points between the 1992 cohort and the average of the previous four cohorts is statistically significant at the 5 per cent level (F value = 4.7). A chi-squared test indicates that the differences in percentage of good degrees between the 1992 cohort and the average of the previous four cohorts is statistically significant at the 1 per cent level.

4.5 Conclusions

In this chapter, the method of calculating value-added developed in the methodology chapter is used to measure value-added for the three first degree courses, Business Studies, Applied Social Studies, and Engineering. It concludes that the method can be used to measure value-added at course level when the course has a large number of students. The test in this chapter shows that the method can identify variations in value-added by making direct comparisons between different cohorts or by making comparisons between the average of previous cohorts and current cohorts. The two measurements serve different purposes and provide different information.

To measure value-added, the characteristics of the three courses have been analyzed first. The three courses are from a new university. As generally acknowledged, it is found that the over the three to five year period (88-92), the three courses recruited students with diverse entry qualifications (about nine different types of entry qualifications). Nevertheless students with A level entry qualifications (A level group) are still predominant on the three courses for

most of the cohorts. Students with BTEC entry qualification (BTEC group) and Access or Conversion Course (Access group) appeared to be two second major groups. Thus value-added measurements are focused on the major groups of students rather than looking at all different types of entry qualifications. This shows the advantage of the model developed in this study over other existing methods (index methods and the comparative method).

The data analysis has shown that the model is easy to use. The model can be used to calculate value-added for the major groups. The results provide information on value-added for students with particular entry qualifications. For example, by using the model, it is found that with The Engineering courses, over the four year period (1989 -1992), for A level students, the 1992 cohort had less value-added than other cohorts, while for the BTEC students, the 1989 cohort had less value-added than other cohorts. On the basis of value-added for individual groups, the model calculates value-added for the course. For example, with the Business Studies course, it is found that during the period of 1988-1992, the 1990 cohort recruited better qualified entrants than other cohorts, but had fewer students achieving good degrees. Thus the question is raised: why did the 1990 cohort of Business Studies course have less value-added? Were there any problems with quality of teaching of this course in 1990? Therefore the model can be used to measure value-added and raise questions about quality of teaching of a cohort. Nevertheless, whether the lower value-added indicates poor quality of teaching will be discussed in the qualitative study of the thesis.

The main problem found through the tests in this chapter is that the number of students on some of the courses sampled, or the number of students in some groups is very small and therefore this may be a reason that some results are not statistically significant. The data show that the number of students on each course varies. Some courses, for example, the Business Studies course, have a large number of students, over a hundred, while the number of students on other courses such as Engineering courses are less than 20. As a test, the courses with a small number of students and from the same subject area were

combined. Nevertheless this indicates that the model may not be suitable to measure value-added at course level unless courses have a large number of students (over a hundred). This finding has implications for the further test of the method.

It should noted that comparisons of value-added between male and female, and between mature and non-mature students are also made. Nevertheless, the numbers of students for these groups are too small and most of the results are not statistically significant. Therefore the results of these comparison are not reported.

CHAPTER FIVE: MEASURING VALUE-ADDED AT SCHOOL AND INSTITUTIONAL LEVEL

5.1 Introduction

In the last chapter, the framework of the proposed value-added measurement was tried out with actual data at course level. This chapter will apply this method to measure value-added at school and institutional level in order to further explore the strengths and weaknesses of the method developed.

The data used in this chapter were obtained from another new university. Compared with the data used in the last chapter, the data used here are relatively complete. The data were from all of the first degree courses of all schools at the institution from 1988 to 1993 although there is still missing information.

The data were downloaded from the student database of the institution in the form of individual student records. The students include those who studied full-time and sandwich students for the first degree courses. A total of 3339 records was used. Since the number of students in some schools is small, these schools are combined into six schools according to subject. In the following presentation, schools are presented by code. The coding is as follows:

School 1: Construction, Computing, mathematics, and Electrical and electronic engineering
School 2: Urban development and policy, and bakery
School 3: Engineering system and design
School 4: Education, politics and social sciences
School 5: Applied science, Health and social care
School 6: Business

This chapter begins with analyzing the pattern of entry qualifications. The results of the analysis are then used to measure the performance of the

schools and the institution. Value-added at both school and institutional levels are measured in two ways. One way is to enable direct comparison between individual schools or cohorts to be made. Another way is to enable comparisons between the university average and individual schools or cohorts to be made. The comparison between the proposed method and the comparative method is also made. A summary of the findings, and conclusions are presented in the final sections.

It should be noted that the proposed method could be tested at course level with the set of the data used in this chapter again. Nevertheless, it is found that the numbers of students on each course from this set of data are also very small. The finding of the last chapter suggests that the proposed method is suitable to measure value-added for those courses which have a large number of students. Therefore the measurements of value-added at course level are not made in this chapter.

5.2 Pattern of entry qualifications

The frequency and percentage of all entry qualifications for all schools during 1988-1993 are calculated. It is found that there were about 25 different types of entry qualifications over the five year period. Nevertheless, 85.2 per cent of them fall into four major groups: A level, BTEC, Access or conversion, other qualifications (see Table 5.1). Apart from these four entry qualifications, some of the 25 entry qualifications only take 0.1 per cent (just one or two students).

Table 5.1 Pattern of entry qualifications

Entry qualifications	Frequency	Percent
A level	1614	41.18
BTEC	662	16.89
Access or conversion	321	8.19
Other qualifications	742	18.93
Total	3339	85.2

** note: 'other qualifications' is a category defined by FESR, which included students with other qualifications than those qualifications have been included into FESR categories such as BTEC, Access, etc.*

According to the above findings, there is no point in trying to give scores to those entry qualifications that are only a small percentage of all entry qualifications. Therefore the value-added measurement will focus on the major groups because they represent above 85 per cent of the population.

It may be worth noting that A level, BTEC, and Access groups are the major groups of entry qualifications for the data analyzed in last chapter. It might be the case with other new universities since they usually recruit students with these entry qualifications.

5.3 Measuring value-added at school level

5.3.1 Value-added based on direct comparisons between different schools

In the last section, four major groups of entry qualifications were identified. This section will measure value-added for these individual groups first, and then value-added for the schools will be measured by taking into account value-added for all of the four groups.

5.3.1.1 Value-added for major groups

Value-added for A level group:

Value-added for A level group is calculated by

1) ranking the schools by mean A level points to produce an input rank;

2) ranking the schools by their degree scores or percentage of good degrees to produce an output rank;

3) taking input rank and subtracting output rank.

Table 5.2 summarizes value-added for the individual groups from 1988 to 1993 inclusive. Within A level groups, the measurements based on either the percentage of good degrees or degree scores suggest that school 2 had the most value-added. This school admitted students with average A level points of 9.24, which was in the middle of the range of mean A level points for the six schools (7.61-13.61), but the percentage of good degrees (52.8%) and degree

Table 5.2 Value-added for different groups

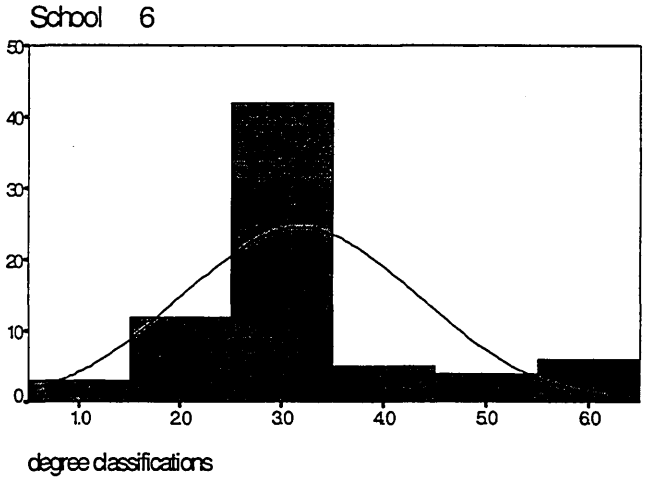
	Mean of A level Points	SD*	Input rank	Good degrees (%)	Output rank	VA good degree	Mean of degree scores	SD	Output rank	VA degree scores	No. of students
A level:											1614
school 1	8.34	3.99	5	39.6	4	1	56.18	9.37	5	0	95
school 2	9.24	4.85	4	52.8	1	3	58.57	8.83	1	3	352
school 3	9.99	5.82	2	51.5	2	0	57.82	9.71	2	0	200
school 4	9.73	4.34	3	36.2	6	-3	56.41	7.56	4	-1	351
school 5	7.61	4.79	6	43.0	3	3	56.10	10.7	6	0	263
school 6	13.61	5.26	1	39.2	5	-4	57.25	6.8	3	-2	352
BTEC:											662
school 1			3.5	25.3	5	-1.5	51.99	10.80	5	-1.5	166
school 2			3.5	44.6	1	2.5	57.42	8.21	1	2.5	83
school 3			3.5	30.0	3	0.5	53.06	10.49	3	0.5	213
school 4			3.5	37.8	2	1.5	55.42	9.67	2	1.5	45
school 5			3.5	26.5	4	-0.5	50.67	11.36	6	-2.5	83
school 6			3.5	20.8	6	-2.5	52.25	11.22	4	-0.5	72
Access:											291
school 3			2.5	22.4	4	-1.5	49.60	12.12	3	-0.5	58
school 4			2.5	25.4	2	0.5	52.32	10.37	2	0.5	134
school 5			2.5	24.2	3	-0.5	48.18	13.04	4	-1.5	33
school 6			2.5	28.8	1	1.5	55.71	7.20	1	1.5	66
Othe qual.:											742
school 1			3.5	41.7	2	1.5	55.38	10.30	2	1.5	96
school 2			3.5	41.2	3	0.5	55.32	11.67	3	0.5	85
school 3			3.5	26.1	5	-1.5	52.75	10.59	5	-1.5	69
school 4			3.5	44.1	1	2.5	56.80	8.91	1	2.5	145
school 5			3.5	25.6	6	-2.5	50.01	12.36	6	-2.5	82
school 6			3.5	33.2	4	-0.5	54.24	10.05	4	-0.5	265

Note: SD = standard deviation

scores (58.57) for this school were the highest in the six schools.

Conversely, school 6 recruited students with the highest mean A level points(13.61), but only 39.2 per cent of the students obtained good degrees. So school 6 had the least value-added. Nevertheless, when degree score is used as a measure of degree quality, the performance of the school is better. This contradictory result may be explained by the distribution of the degree classifications. Fig 5.1 shows the distribution of degree classifications of the school. It indicates that very few students obtained a first class degree, thus, the percentage of good degrees for the school is lower, but on the other hand, there were few students who obtained a degree class lower than lower second, so the degree score for this school is relatively high.

Fig. 5.1 Distribution of degree classifications of the school 6



Note: 1 = first , scored 72; 2 = upper second , scored 64; 3 = lower second , scored 54; 4 = third , scored 44; 5 = unclass / pass, scored 35; 6 = fail, scored 25.

The above results raise the question of why A level students from school 6 appeared to have higher starting points than A level students from school 2, but the proportion of the students who achieved good degrees in school 6 was lower than that in school 2. Is the teaching quality in school 6 poorer than that in school 2? The answers to these questions will be discussed in the second part of the research (chapter 6).

Table 5.2 also shows that the variations in degree scores among the six schools are much smaller than the variations in percentages of good degrees between them. For example, the range of percentage of good degrees between A level students in the six schools is 36.2 - 52.8, while the range of mean of degree score is 56.10 - 58.57

Value-added for BTEC, Access, and Other qualifications groups:

The calculation of value-added for these three groups is based on comparing degree results, and assuming that within the individual groups, the starting points of the students are the same. Thus when we calculate value-added for the individual groups, the comparison of students' academic attainment between different schools can be made directly by comparing degree results.

Nevertheless, we still give an input rank for all the school, since the input rank would be useful when we calculate value-added for the schools taking all the individual groups into account. We use the average rank as an input rank for all the school. The average rank is the average of the rank that the schools would have received if their starting points were distinct. For example, the average rank for the six schools is 3.5 (3.5 = (1+2+3+4+5+6) / 6).

Since the number of students of the Access group from Schools 1 and 2 is small, only comparisons between access students from the other four schools are made.

It appears that some schools' performances vary with individual groups, and variations in value-added for the same school can be dramatic. For example, the Access students of school 6 had much more value-added than other schools, but the value-added for the BTEC students of this school were much less than that for other schools (see Table 5.2). On the other hand, the value-added scores for some Schools were always in the middle of the range

compared with other schools. For example, the value-added scores for the three groups (A level, BTEC, and other qualification groups) of school 1 are all in the middle of the range. School 2, always had positive value-added scores.

Table 5.2 also indicates that regardless of which school they are from, percentages of good degrees between the four groups of students are different. For example, the range of percentages of good degrees for the access students from the six schools is from 22 to 29 percent, while the range for A level students is 36-53 percent, for BTEC students is 20-45 per cent, and for the other qualification group is 25 - 44. Clearly, the range of proportion of good degrees achieved by the A level student group was higher than other groups.

However, the analysis also shows that within some of the schools, the proportion of good degrees achieved by students from BTEC, and 'other qualification' groups are higher than that from the A level group. For example, within school 4 the proportions of good degrees for BTEC (37.8%) and 'other qualifications' (41.7%) groups are higher than that for the A level group (36.2%).

5.3.1.2 Value-added for the Schools

Since the number of students of the Access group from schools 1 and 2 is small, this group is not taken into account when calculating value-added for the schools. The three major groups, A level, BTEC, and other qualification groups can represent over 66 percent of population of each school. Value-added for the whole school is measured by calculating the weighted sum of value-added scores for the three major groups of students. The calculation procedure is the same as that of value-added for the courses (see the section 4.3.2 of last chapter).

Table 5.3 summarizes the results of the value-added for the six schools. It can be seen that the value-added score for school 2 is much higher than the other five schools, while, value-added scores for school 6 are the lowest. Thus it raises the question of why school 6 had much less value-added. Are there any

problems with teaching quality in school 6 ? Is quality of teaching in school 2 better than that in other schools? These questions will be discussed in the next chapter.

Table 5.3 also indicates that two degree measurements suggest different value-added scores for schools 5, 1, 3, and 4. The big difference exists between schools 5 and 4. When degree results are measured by percentages of good degrees, school 5 is second best in value-added and school 4 ranks five. Nevertheless, when degree results are measured by degree scores, the result is completely opposite. This demonstrates that using different methods to measure quality of degree could lead to different value-added results. However the analysis also suggests that compared with other schools, the value-added scores for Schools 3 and 1 are in the middle position, no matter which method is used to measure degree results.

Table 5.3 Value-added for the schools

school code (in ranking order)	value-added (good degrees)	school code (in ranking order)	value-added (degree scores)
School 2	2.08	School 2	2.08
School 5	.80	School 4	.37
School 1	.52	School 3	-.24
School 3	-.24	School 1	-.45
School 4	-.60	School 5	-.57
School 6	-2.56	School 6	-1.19

5.3.1.3 Statistical tests

Statistical tests for A level group

To confirm the above observations, some statistical techniques are used to test the observed results. The mean A level points and the mean degree scores are represented by an interval variable, but the distributions of A level points and degree scores are not Normal (see Fig 2.1, and Fig. 2.2 in Appendix 2). Therefore a non-parametric test, the Kruskal-Wallis test is used to test if the differences in mean A level points between different schools and the

differences in degree scores between different schools are statistically significant. Nevertheless, analysis of variance (ANOVA) is also carried out because ANOVA is very robust.

Table 5.4 summarizes the test results. Both tests suggest that the differences in mean A level points and degree scores between different schools within individual groups (i.e. A level group, BTEC group, Access group, and other qualifications group) are statistically significant. The significant levels are at 1 percent.

Table 5.4 ANOVA and Kruskal-Wallis test results for the four groups

	Kruskal-Wallis A level points		Kruskal-Wallis Degree scores		ANOVA A level points		ANOVA Degree scores	
	value	DF	value	DF	F-value	DF	F-value	DF
A level:	246.17**	5	25.34**	5	54.62**	5	3.69**	5
BTEC:			23.65**	5			4.64**	5
Access:			17.04**	5			4.21**	5
Other qual:			23.20**	5			5.12**	5

Notes: ** = significant at the 1% level ; * = significant at the 5% level

On the basis of above test results, a further statistical test, the Tukey-HSD (honestly significant difference) test is carried out to identify which schools show the significant differences in mean A level points and degree scores. The Tukey-HSD test allows us to compare each pair of the schools to see if their difference is significant.

Table 5.5 (a) and (b) summarise the Tukey-HSD test results. Table 5.5 (a) suggests that the mean A level points score of school 5 is statistically significantly lower than those of school 2, school 4, school 3 and school 6, and the mean A level points score of school 6 is statistically significantly higher than those of school 5, school 2, school 4, and school 3.

These test results may be interpreted as showing that school 6 (Business

School) recruited students with significantly higher academic attainments than other schools did. Conversely, the academic attainments of students at entry in school 5 are significant lower than that in other schools.

Table 5.5 (a) The results of Tukey-HSD test for A level group (mean A level points)

Mean A level points		School 5	School 1	School 2	School 4	School 3	School 6
7.61	School 5						
8.28	School 1						
9.24	School 2	*					
9.73	School 4	*					
9.99	School 3	*					
13.61	School 6	*	*	*	*	*	

*Note: * = significant at the 5% level*

Table 5.5 (b) indicates that the mean degree score of school 5 is statistically significantly lower than that of school 2, and the mean degree score of school 2 is significantly higher than that of school 1. This may be interpreted as showing that the degree quality, in terms of degree scores, of school 2 is statistically significantly higher than those of schools 5 and school 1.

Table 5.5 (b) The results of Tukey-HSD test for A level group (mean degree scores)

Mean degree scores		School 5	School 1	School 4	School 6	School 3	School 2
56.10	School 5						
56.18	School 1						
56.41	School 4						
57.25	School 6						
57.82	School 3						
58.57	School 2	*	*				

*Note: * = significant at the 5% level*

Statistical tests for BTEC, Access, and Other Qualifications groups

Since the distributions of degree scores of BTEC, Access, and 'other qualifications' groups are not Normal (see Fig. 2.3, Fig. 2.4 and Fig. 2.5 in

Appendix 2), the non-parametric test, Kruskal-Wallis statistical test has been used to test whether the differences in mean degree scores between different the schools within the three groups are statistically significant. Nevertheless, the analysis of variance is also used to confirm the results of Kruskal-Wallis test. The results of these tests are summarized in Table 5.4 (on page 117). It is found that there are statistically significant differences in degree scores between the six schools within the three groups (at the 1 per cent level). Therefore, a further statistical test, the Tukey-HSD test is carried out to find out between which pair of schools the significant differences exist. The results of the Tukey-HSD test are summarized in Table 5.6.

Table 5.6 The results of Tukey-HSD test for BTEC, Access, and 'Other qualifications' groups (mean degree scores)

BTEC group							
Mean degree scores		School 5	School 1	School 6	School 3	School 4	School 2
50.67	School 5						
51.99	School 1						
52.25	School 6						
53.06	School 3						
55.42	School 4						
57.42	School 2	*	*	*	*		
Access group							
Mean degree scores		School 5	School 3	School 4	School 6	School 1	School 2
48.18	School 5						
49.60	School 3						
52.32	School 4						
55.71	School 6	*	*				
'Other qualifications' group							
Mean degree scores		School 5	School 3	School 6	School 2	School 1	School 4
50.01	School 5						
52.75	School 3						
54.24	School 6	*					
55.32	School 2	*					
55.38	School 1	*					
56.80	School 4	*					

Note: * = significant at the 5% level

From Table 5.6, it can be seen that the significant differences in mean degree scores only exist in a few pairs of the schools within the three groups (i.e. the BTEC, Access, and other qualifications groups). For example, within BTEC groups, only School 2 had significantly higher degree scores than the other four schools. Within Access group, the significant differences in mean degree scores only exist between two pairs of schools: school 6 and school 5, and school 6 and school 3.

Given the assumption that the starting points of the students within the three groups are the same, these results may be interpreted as showing that within the three groups, there are no big differences in value-added between most of the schools. In general, within the BTEC, Access, and other qualifications groups, the students of the six schools entered the university with similar academic attainment, and graduated with similar degree results.

5.3.2 Value-added based on comparison between the university average and individual schools

As noted earlier, the proposed method can be used to enable comparison between the university average and individual schools to be made. Table 5.7 displays the university average in A level points, degree score, and percentage of good degrees.

Table 5.7 The university average in A level points, degree score, and percentage of good degrees.

	Mean A LP		Degree score		Good degrees (%)
	average	SD	average	SD	average
A level group	10.07	5.31	57.18	8.69	43.7
BTEC group			53.11	10.60	29.8
Access group			52.32	10.67	27.10
Other qual. group			54.41	10.54	35.80

A level group

When the individual schools (using the data in Table 5.2) are compared with

the university average (see Table 5.7), it is found that A level students from schools 2 and 3 entered the university with lower than university average A level points, but they achieved better degree results (both degree scores and percentage of good degrees) than the university average. Therefore A level students from school 2 and school 3 appear to have more added value.

The academic attainments of A level students from school 1, school 4 and school 5 were lower than the university average both at entry and at exit. This may be interpreted as showing that A level students from these three schools achieved what they were expected to achieve in terms of academic attainments compared with the university average.

With A level students from school 6, it is found that they started the university with academic attainment which was much higher than the university average, but their academic attainment at exit was lower than the university average when degree quality is measured by percentage of good degrees. Therefore these students did not achieve what they were expected to achieve. However, when degree quality is measured by degree scores, their academic attainment at exit was higher than the university average. Therefore these students achieved what they were expected to achieve.

BTEC, Access, and 'Other qualification' groups

Within the three groups, when value-added is measured by comparing degree results between individual schools and the university average, it is found that with school 2, the students of the three groups (i.e. BTEC, Access, and 'other qualification' groups) all did better than the university average. In contrast, with school 5, the degree results achieved by the students of the three groups all were lower than the university average. Therefore we may say that compared with the university average, BTEC, Access, and 'other qualification' students in school 2 were added more value, but students of these groups in school 5 are added less value. The other schools had mixed results. Some groups did better than the university average and some groups did worse.

5.3.3 Comparison between the proposed method and the comparative method

To compare the main existing value-added measurement, the comparative method, with the method of calculating value-added developed in this research, the comparative method is used to analyze the same data.

The calculation of the comparative value-added (CVA) involves eight steps: calculating benchmark, actual degree results, expected degree results, differences between actual and expected degree, sum of the differences for each degree classification, multiplying factors, sum of all differences, and average of the differences between actual and expected degree results. For example, the procedure of calculation of a comparative value-added score is as follows:

Benchmark:

Entry qual	1st	2.1	2.2	3rd	Pass	Fail
3A 26-30	.20	.33	.47	.00	.00	.00
3A 18-25	.10	.42	.44	.03	.02	.01
3A 3-17	.05	.39	.46	.06	.04	.01
2A 16-20	.04	.53	.38	.02	.02	.00
2A 10-15	.07	.41	.43	.05	.03	.01
2A 1-9	.06	.35	.45	.07	.06	.01
1A	.02	.37	.47	.07	.07	.01
BTEC	.04	.26	.44	.11	.13	.02
other qual	.04	.31	.43	.09	.11	.02

This benchmark was calculated from the university data over the period of 1988-1993. It means that on average, for example, 7 per cent of those students with 2A (10-15) at the university during 1988-1993 achieved a first class degree. The average percentages are then used to calculate the expected degree results.

The calculation of comparative value-added score for entry qualification of 2 As, 10-15 points:

	1st	2.1	2.2	3rd	Pass	Fail	No.
Observed (N)	2.00	1.00	6.00	1.00	.00	.00	10
Expected (%)	.07	.41	.43	.05	.03	.01	
O-E	1.30	-3.10	1.70	.50	-.30	-.10	
multiplying factor	x 72	x64	x54	x44	x35	x25	
	93.6	-198.4	91.8	22	-10.5	-2.5	
	93.6-198.4+91.8+22-10.5-2.5 = -4						
	CVA = -4/10 = -0.4						

The above calculation procedure has to be repeated for all entry qualifications. At this point, the calculation procedure of the proposed method is simpler than the comparative method.

Further more, from the above example, it can also be seen that the proposed method has an advantage over the comparative method in terms of interpretation. According to the comparative method, the above CVA score of - 0.4 means that the equivalent of 40 per cent of graduates with entry qualifications of 10-15 A level points from 2 A levels achieved one degree class lower than expected. But such a conclusion certainly does not leap out from the above table. The CVA system makes intuitive sense provided students are moving to categories immediately above or below their expected category (Gallagher, 1991). However, it is possible that a student who is expected to fail, but he or she gains a 2.2.

The method for calculating comparative value-added takes into account movement from one category to another (e.g. from 2.2 to 2.1) and distance travelled between categories (e.g. from fail to 2.2). Therefore a CVA score might give the equivalent of percentage graduates achieving one degree class higher/ lower than expected, but it is difficult to know how to interpret this. The same CVA score of -0.40 could be product of very different patterns of gains and losses from different sized cohorts of students(Gallagher, 1991).

The interpretation of value-added scores based the proposed method is relatively straightforward. A value-added score in the proposed method

indicates change of position, in terms of academic attainment, of a school or a cohort. For example, the value-added score for A level group of school 2 is 3 (see Table 5.2). This means that the position of school 2 in academic attainment rank order has moved up 3 rank. This is clearly showed in Table 5.2 where school 2 ranked 4th at entry, but ranked 1st at exit. So the interpretation of a value-added score of the proposed method is straightforward.

The value-added results for each school based on the comparative method are displayed in Table 5.8. These results are compared with the results produced by

Table 5.8 The CVA for the schools

Schools	Value-added
school 2	.83
school 4	.74
school 3	.65
school 1	-.39
school 6	-.65
school 5	-.27

the proposed method. It is found that when the proposed method uses the percentage of good degree as an output measurement, the value-added for the schools suggested by this method is different from that suggested by the comparative method. For example, the proposed method suggested that School 4 performed less well than the other four schools did, being ranked fifth (see Table 5.3), but the comparative method indicated that School 4 performed quite well, being ranked second (see Table 5.8).

However, when the proposed method uses degree scores as a measure of degree quality, the value-added results produced by this method are similar to those suggested by the comparative method. Both methods suggest the same rank for the four schools (i.e. school 2, 4, 3, and 1) although the value-added scores are different. One explanation for this may be that when the proposed method uses degree scores as a measure of degree results, both the comparative method and the proposed method calculate value-added by taking

into account all degree classifications. However when the proposed method uses percentage of good degree as an output measurement, the value-added scores generated from the proposed method are based on good degrees (first class and upper second class degrees). This reinforces the finding, discussed earlier, that using different methods to measure degree results could have different value-added results.

5.4 Measuring value-added at institutional level

In the above sections, we used the proposed method to measure value-added at school level. From this section, we will use the method to identify variations in value-added of the institution between different cohorts (1988-1993)

5.4.1 Value-added based on direct comparisons between different cohorts

We will measure value-added for individual groups (i.e. A level, BTEC, Access, and 'Other qualifications' groups) first, and then value-added for the institution as a whole will be measured by taking into account value-added for these four groups.

5.4.1.1 Value-added for the major groups

Table 5.9 summarises value-added for the four individual groups, A level, BTEC, Access, and Other qualification groups. Within A level groups, there is not a clear upward or downward trend in value-added for all A level students of the institution over the six year period (88-93). Attention may be drawn to the 1991 cohort, because the percentage of good degrees for this cohort was much lower (34.3%) than other cohorts (44%-51%). Nevertheless, the average A level points with this cohort was 9.04, which was also lower than other cohorts. Therefore we may conclude that although there was a smaller proportion of students obtaining good degrees in the 1991 cohort, this may be due to lower starting points. In terms of value-added, this cohort may just achieve what was expected.

Table 5.9 Value-added for different cohorts of the University

Entry cohort	Mean A level points	Input rank	Percentage of Good degrees	Output rank	VA good degrees	Mean degree score	Output rank	VA degree score	No. of students
A level:									1614
1988	10.67	2	44.3	4	-2	57.76	3	-1	228
1989	10.21	4	51.9	1	3	58.43	2	2	266
1990	10.60	3	51.1	2	1	58.63	1	2	186
1991	9.04	6	34.3	6	0	55.83	6	0	362
1992	9.99	5	46.4	3	2	57.25	4	1	318
1993	10.77	1	44.1	5	-4	56.84	5	-4	254
BTEC:									662
1988		3.5	47.62	1	2.5	55.10	2	1.5	21
1989		3.5	36.62	2	1.5	54.45	3	0.5	71
1990		3.5	35.53	3	0.5	55.14	1	2.5	76
1991		3.5	26.24	5	-1.5	53.17	4	-0.5	141
1992		3.5	23.41	6	-2.5	51.87	6	-2.5	205
1993		3.5	33.11	4	-0.5	52.81	5	-1.5	148
Access:									321
1988		3.5	29.79	1	2.5	54.43	1	2.5	47
1989		3.5	27.10	4	-0.5	51.81	4	-0.5	48
1990		3.5	18.52	6	-2.5	50.78	6	-2.5	27
1991		3.5	28.89	2	1.5	53.02	3	0.5	45
1992		3.5	26.79	5	-1.5	53.70	2	1.5	56
1993		3.5	27.55	3	0.5	50.88	5	-1.5	98
Other qual.:									742
1988		3.5	30.77	6	-2.5	58.31	1	2.5	13
1989		3.5	34.78	4	-0.5	54.67	4	-0.5	46
1990		3.5	35.90	3	0.5	55.33	2	1.5	39
1991		3.5	32.79	5	-1.5	53.77	6	-2.5	91
1992		3.5	36.53	1	2.5	53.95	5	-1.5	375
1993		3.5	36.52	2	1.5	55.14	3	0.5	178

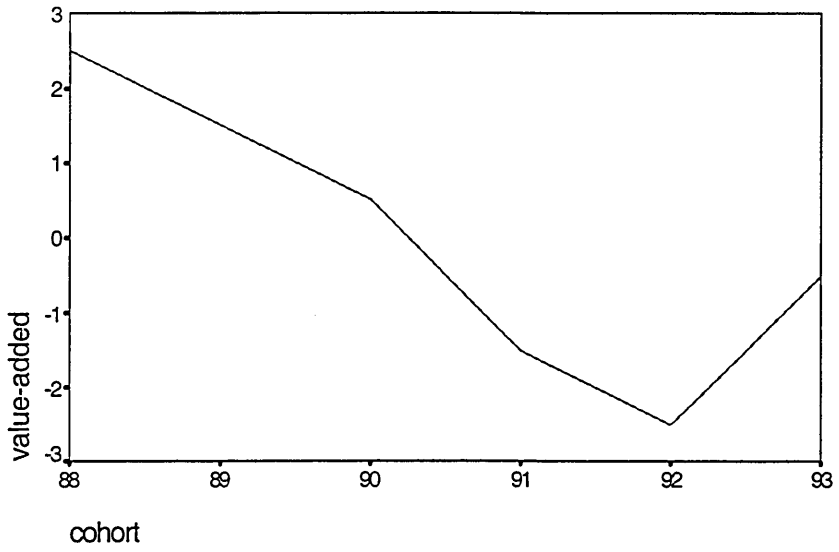
When degree results are measured by average degree score, although the 1991 cohort still achieved a lower degree score than the other five cohorts, the difference in degree results between this cohort (55.83) and the other five cohorts (the range: 56.84-58.63) was smaller than the difference between them when degree result is measured by percentage of good degrees. Nevertheless this was not reflected in the value-added score. This shows the disadvantage of the proposed method.

Another cohort to which the attention may be drawn is the 1993 cohort, because the value-added score for this cohort was -4. This cohort entered the university with the highest average A level point score (10.77) in the six cohorts, but the proportion of good degrees was lower than the other four cohorts apart from the 1991 cohort. Therefore this cohort did not achieve what it was expected to achieve. Nevertheless it will be discussed in chapter 6 whether this value-added result indicates that the quality of teaching for this cohort was poorer than that for other cohorts.

Within the BTEC student group, on the whole, there appears to be a downward trend in value-added from 1988 to 1992, but from 1992 to 1993, value-added for this group increased (see Fig. 5.2). Value-added scores indicate that the 1992 cohort did not perform well. Only about 23 per cent of this cohort achieved good degrees. This is less than a half of the proportion of students who achieved good degrees in the 1988 cohort (47%). So the 1992 cohort had less value-added than other cohorts. This value-added result may raise the question about the quality of teaching of the 1992 cohort. Yet, the value-added score for A level students of this cohort was 2. This indicates that the A level students of this cohort achieved higher than they expected to achieve.

Within the Access student group, on the whole, the value-added decreased from the 1988 to 1990 cohorts. After 1990, there was no clear upward or downward trend in value-added. Although the percentage of good degrees for the 1988 cohort was 29.79 which was much lower than all A level cohorts, nevertheless, in terms of value-added, the value-added score was 2.5.

Fig. 5.2 The trend in value-added for BTEC students at the institution (1988-1993)



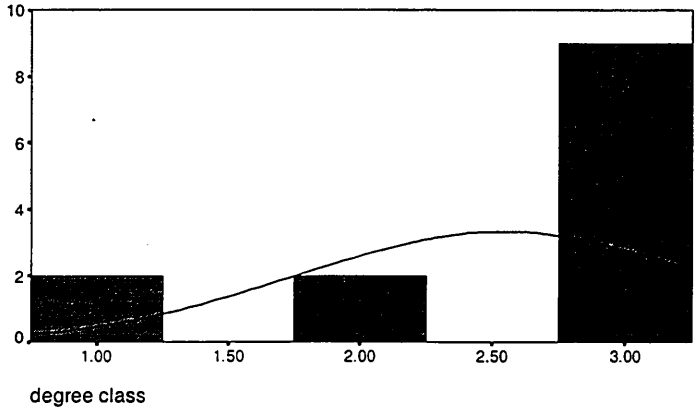
This may be explained by noting that when we take entry standard into account, the 1988 cohort compared with students with a similar starting point from other cohorts had done well. However, for the 1990 cohort, the percentage of good degrees is extremely low, only 18.52%, and the degree score is also the lowest in all groups and all cohorts. Thus we should ask whether there are any problems with the quality of teaching for this cohort.

With the 'other qualifications' group, it appears that there are variations in value-added between the six cohorts. However, the variations both in the proportion of good degrees and degree scores are actually very small. The proportion of good degrees for the six cohorts was in the range of 30.77 - 36.53 and the range of degree scores was 53.77 - 58.31. This may be interpreted as showing that students with 'other qualifications' achieved similar academic attainments over the period of 1988 - 1993. It should be noted that this is not reflected by the value-added scores. It is a weakness of the proposed method. The value-added scores here only provide a very basic information, and they must to be put into context.

It may be also worth noting that degree quality of the 1988 cohort suggested by the two degree measurements was very different. It can be seen from Table 5.1 that the percentage of good degrees for the 1988 cohort (30.77%) was the lowest in the six cohorts, while the degree score for this cohort was 58.31, which was the highest. These results demonstrate that the two measurements of degree quality may lead to very different value-added results.

Again, the distributions of degree classifications seem to explain the contradictory results given by different measurements of degree quality. Figure 5.3 displays the distribution of degree classification of the 1988 cohort of the

Fig. 5.3 Distribution of degree class of students with other qualifications of the 1988 cohort



Note: 1 = 1st , scored 72; 2 = 2. , scored 64; 3 = 2.2 , scored 54; 4 = 3rd , scored 44; 5 = unclass / pass, scored 35; 6 = fail, scored 25.

'other qualifications' group. It can be seen from the figure that since there were no students who obtained degree class lower than 2.2, the degree score for this cohort was high. Nevertheless, there were few students who achieved good degrees in the 1988 cohort, therefore the percentage of good degrees was ranked 6th.

The above result summarise some variations in value-added within the individual groups. There are some interesting findings on value-added or degree results between the four groups. It appears that in general, the

proportion of good degrees achieved by the A level group was higher than the other groups. The range of proportion of good degrees achieved by the four groups are summarized below:

A level group : 51-34 per cent,
BTEC group : 48-23 per cent,
'Other qualifications': 37-30 per cent ,
Access group: 30-19 per cent.

A level, BTEC, Access, and 'Other qualifications' are four different types of entry qualifications. Although A level is the qualification which universities are most favourable towards when recruiting students, there are some arguments about whether A level is a good measure of students' actual level of attainment in areas relevant to their studies as discussed in the literature review. The above finding shows that the proportion of good degrees achieved by the A level group was higher than the other groups.

In contrast, when the degree score is used to measure degree quality, the variations in degree results between the four groups appear to be smaller. Degree scores for the four groups were about between 50-59. Nevertheless, on the whole, degree scores for A level groups were higher than the other groups. The range of degree scores achieved by the four groups are summarized below:

A level group : 56.84 - 58.63,
'Other qualifications': 53.77 - 58.31,
BTEC group : 51.87 - 55.14,
Access group: 50.78 - 54.43.

5.4.1.2. Statistical test

The above observations are tested by a number of statistical tests. With the A level group, since the distributions of the A level points and degree scores are not normal (see Fig. 3.1 and Fig. 3.2 in the Appendix 3), the non-parametric

Kruskal-Wallis test was used to test differences in mean A level points and mean degree scores between different cohorts. It indicates that these differences are statistically significant. A chi-squared test also suggests that there are statistically significant differences in percentage of good degrees (see Table 5.10)

Table 5.10 Kruskal-Wallis and ANOVA tests results

	Kruskal-Wallis mean A level points		Kruskal-Wallis mean degree scores		ANOVA mean A level points		ANOVA mean degree scores		Chi-square Good degrees (%)	
88-93 cohort	value	DF	value	DF	F-value	DF	F-value	DF	value	DF
A level	23.5**	5	27.4**	5	4.69**	5	4.2**	5	25.06**	5
BTEC			8.37	5			1.52	5	11.59*	5
Access			3.85	5			1.08	5	1.26	5
Other quals.			2.25	5			0.80	5	0.61	5

Notes: ** = significant at the 1% level * = significant at the 5% level

Although the distributions of mean A level points and degree scores are not normal, analysis of variance (ANOVA) is used to confirm the above test results. ANOVA indicates the same results (see Table 5.10). A further statistical test, the Tukey-HSD test is then carried out. This test allows us to compare each pair of means to see if their difference is significant. Table 5.11 (a) displays the test results of mean A level points and Table 5.11 (b) indicates test results of degree scores. These test results suggest that mean A level points for the 1990, 1988, and 1993 cohorts are statistically significantly higher than that for the 1991 cohort.

With the BTEC, Access and 'Other qualifications' groups, since the distributions of degree class of the three groups are not normal (see Fig. 3.3, 3.4 and 3.5 in the Appendix 3), the non-parametric test, Kruskal-Wallis statistical test has also been used to test if the differences in mean degree scores between different the cohorts within the three groups are statistically significant. Nevertheless, the analysis of variance is also used because the analysis variance is very robust. The chi-square is used to test the significance of the differences in percentage of good degree. It is found that differences in percentages of good degrees

between BTEC students of different cohorts are statistically significant at 5 % level, but other differences in mean degree scores and percentages of good degrees are not statistically significant (see Table 5.10).

Table 5.11 (a) Tukey-HSD test results (differences in mean A level points between each pair of the cohorts)

Mean ALP		1991	1992	1989	1990	1988	1993
9.04	1991						
9.99	1992						
10.21	1989						
10.60	1990	*					
10.67	1988	*					
10.77	1993	*					

Notes: * = significant at the 5% level

Table 5.11 (b) Tukey-HSD test results (differences in mean degree scores between each pair of the cohorts)

Mean degree scores		1991	1993	1992	1988	1989	1990
55.83	1991						
56.84	1993						
57.25	1992						
57.76	1988						
58.14	1989	*					
58.63	1990	*					

Notes: * = significant at the 5% level

The interpretation of the test results may be that when we take random variations in entry qualifications and degree results into account, in general, over the six year period (1988-1993), the university had recruited students with similar A level points and these students achieved similar degree results. Therefore there were not big differences in value-added between most of these cohorts. The mean A level points and mean degree score for the 1991 cohort were significantly lower than other cohorts, but this cohort achieved degree

results which were expected, in terms of value-added.

5.4.1.3 Value-added for the institution

Since the four major groups of students can represent over 80 percent of students of each cohort, the value-added for the whole university can be measured by the weighted sum of the value-added for the four groups of students. The calculation procedure of value-added for an institution is the same as that of value-added for a course and for a school. The results are summarised in Table 5.12, although these results can be presented more effectively by a figure (see Fig 5.4)

Table 5.12 value-added for the institution

Cohort	value-added (degree scores)	value-added (good degrees)
1988	.59	-.80
1989	.92	1.60
1990	1.56	.46
1991	-.60	-.48
1992	-.19	.89
1993	-2.28	-1.67

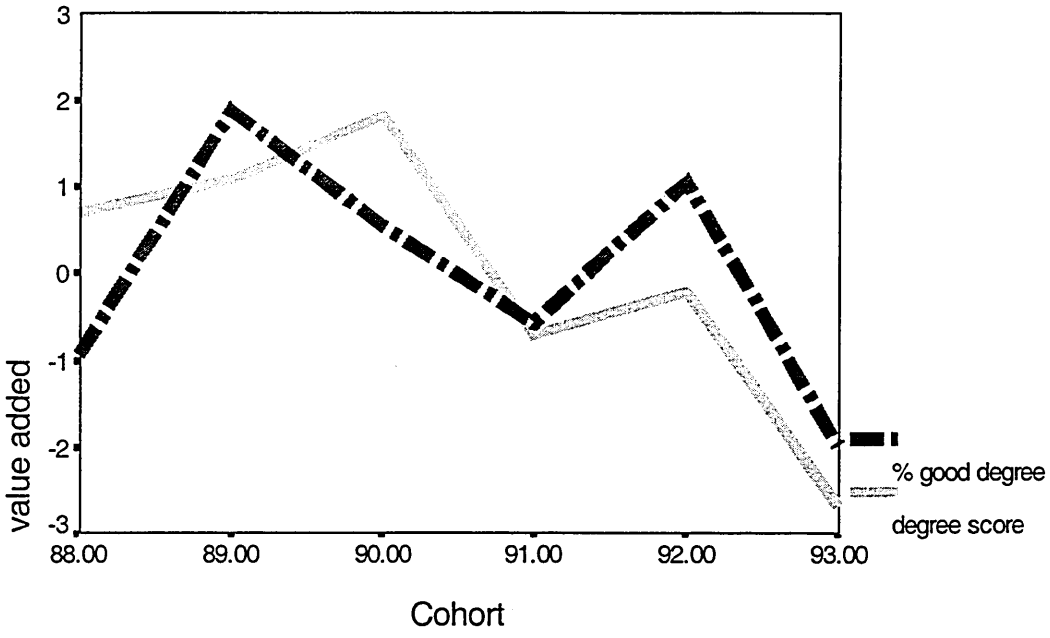
From 1988 to 1993, on the whole, there was no clear upward or downward trend in value-added for the institution. Nevertheless, the measurement based on degree scores suggests a downward trend after 1990.

It can be seen from Fig. 5.4 that compared with the other five cohorts, the 1993 cohort of the university appeared to have less value-added. This was the case when degree quality was measured by either percentage of good degree results or degree score. Therefore the attention may be drawn to this cohort. Why did this cohort have less value-added than the others? Was the quality of teaching with this cohort poor?

Figure 5.4 also shows that the percentage of good degrees for the 1988 cohort was low, but degree score for this cohort was relatively high. This may be interpreted as showing that there should have been more students achieving

good degrees. Is this because teaching had focused on helping all students to pass to obtain a degree rather than helping students to achieve good degrees?

Fig. 5.4 Variation in value-added over the six years



5.4.2 Value-added based on comparison between average previous cohorts and current cohort

In the above, value-added was measured by direct comparison between different cohorts. Value-added for a particular cohort, for example a current cohort, may also be measured by enabling comparison to be made between this cohort and average of previous cohorts. In this section, this method is employed to measure value-added for the four major groups of students.

Table 5.13 indicates average scores of input and output for the 88-92 cohorts and the scores for the 1993 cohort (96/97 graduates). Within the A level group, mean A level points for the 1993 cohort was higher than that for the previous five years (88-92). This may imply that in 1993, the university attracted students with higher starting points than before. Therefore it is expected that the degree results for this cohort should be better than the average of the previous five cohorts.

Table 5.13 Comparison between the 1993 cohort and the average of the previous cohorts

	88-92 cohort	93 cohort
A level :		
Mean ALP	9.96	10.77
Std Dev	5.17	5.93
Good degrees	44.6%	44.1%
Degree scores	57.37	56.84
SD	8.53	9.34
BTEC:		
Good degrees	28.8%	33.1%
Degree scores	53.2	52.81
SD	10.28	11.68
Access:		
Good degrees	26.9%	27.6%
Degree scores	52.96	50.88
SD	10.21	11.58
Other qualifications:		
Good degrees		
Degree scores	35.6%	36.5%
SD	54.17	55.14
	10.88	9.41

Nevertheless, this did not appear to be the case. When the degree result was measured by either percentage of good degrees or degree score, degree results for the 1993 cohort were lower than the average of the previous five years. This result is consistent with the result when the 1993 cohort was compared with the five individual cohorts. Therefore, again, the attention should be drawn to the 1993 cohort. The question should be asked about why the 1993 cohort did not achieve better degree results than the average of the previous five years. The quality of teaching for the 1993 cohort should be particularly examined.

Within the BTEC group, there were more students from the 1993 cohort achieving good degrees than expected, nevertheless, the degree score for this cohort was lower than the average of previous five years. This was the case with the access group. The 'Other qualifications' group appeared to be the best among the other groups of the 1993 cohorts. It is found that this group had achieved better degree results than the previous five cohorts when the degree quality was measured by both measurements.

The significance of the difference in mean A level points and the difference in degree scores between the pre-1993 cohort and the 1993 cohort is tested by

ANOVA. A chi-square is used to test the differences in percentage of good degree between them. The test results show that only the difference in mean A level points between pre-1993 and 1993 cohorts is statistically significant at the 5 % level (F value = 4.87, DF= 1), and the other differences displayed in Table 5.13 are not statistically significant.

5.4.3 Comparison between the proposed method and the comparative method

The same data are also analyzed by the comparative value-added method (CVA). The calculation procedure of the CVA is the same as that of CVA at the school level described earlier. The CVA results are displayed in Table 5.14.

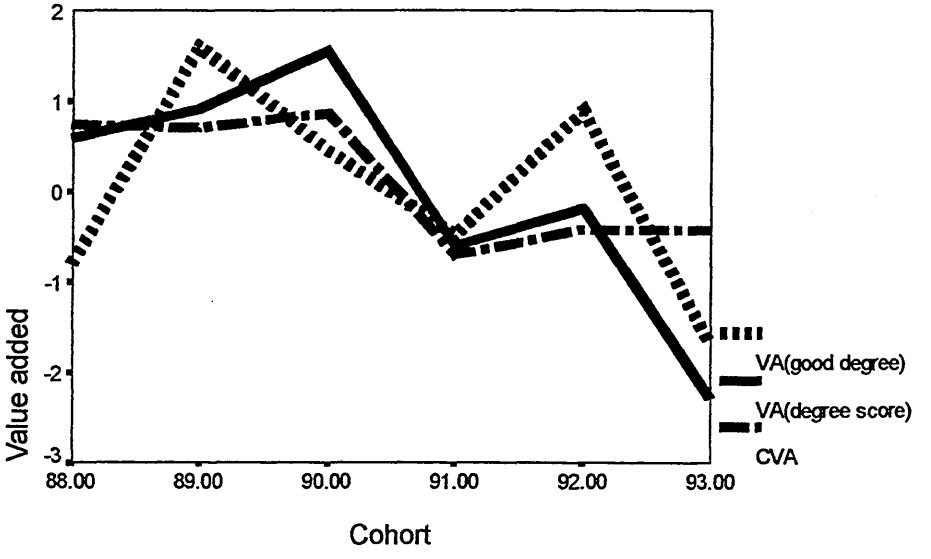
Table 5.14 The CVA for the six cohorts

Entry cohort	CVA
1988	0.74
1989	0.71
1990	0.87
1991	-0.70
1992	-0.42
1993	-0.42

These results are compared with the results derived from the proposed method in this study. It is found that although the actual value-added scores themselves generated from the two methods are very different (this may be due to the fact that the two methods are different), in terms of which cohort had more value-added than others, the results generated from the comparative method are similar to the results derived from the method developed in this study, when degree scores are used. For example, both methods suggested that the 1988 cohort had less value-added than the 1990 cohorts (see Table 5.12).

When the proposed method uses percentage of good degrees as an output measure, there appears to be more difference in value-added results between the two methods. When comparing the value-added results among the following measurements: the value-added based on degree scores, value-

Fig. 5. 5 Value-added results based on different methods



added based on percentage of good degrees and the comparative method (see Fig. 5.5), it is found that the differences in value-added results between the two method are even smaller than differences in value-added results from the same method of calculating value-added, but using different methods to measure quality of degree (one based on degree scores, the other based on percentage of good degrees). This again reinforces the point discussed earlier that using degree scores, and percentage of good degrees to measure the same data could generate different results.

The comparative method attempts to calculate CVA for each entry qualification. With the data in this study, it has been found that there were 25 different types of entry qualifications involved during 1988-1993 in the university. Therefore, apart from the four major groups identified (A level, BTEC, Access, and 'other qualifications'), the comparative method needs to calculate CVA for another 21 types of entry qualifications, where some of them only have two or three students. The value-added results based on such a small number of students may ring false alarms (Straw and Kaye, 1995).

5.5 Conclusions

This chapter uses the value-added method developed in the methodology chapter to measure value-added of the institution over the period 1988-1993. The method is used both at school and institutional level. At school level, the method is used to measure value-added for the six schools over the period of 1988-1993. At the institutional level, the method is used to measure the value-added for the institution in the six years (1988-1993).

It is found that both at school and institutional levels, the method could be used in two ways. One way is to apply the method to measure value-added by making direct comparisons between different schools and between different cohorts. Another way is to apply the method to measure value-added by making comparisons between the university average and individual schools or between the average of the previous cohorts and current cohort. In the first way, the value-added for the major groups of a variety of entry qualifications is measured first, then the value-added for the whole schools or the whole university (cohort) is measured by calculating the weighted sum of value-added for the individual groups.

A pattern of entry qualifications of students at the institution is identified in that 85.2 per cent of the students' entry qualifications fall into four major groups: A level, BTEC, Access or conversion, and other qualifications. The measurement of value-added is then focused on these groups. By doing so, the proposed method avoids the difficulties with attributing scores to diverse entry qualifications, and measures value-added by comparing like with like (the comparisons of input and output are made within the same type of entry qualifications), Nevertheless, when total number of students of a school or a cohort is small, this may lead to a problem that the number of students is so small that it is difficult to draw any statistical conclusion, since the total number of students is divided into four different groups.

Degree results are measured by the two methods, percentage of good degrees, and degree scores. It is found that both at institutional and at school level, on the whole, the variations in percentages of good degrees between the four major groups were bigger than variations in degree scores between them. The analysis indicates that the proportion of good degrees achieved by A level students (e.g. 51-34% at institutional level) are the highest in the four major groups, while the proportion of good degrees obtained by access students (e.g. 30-19% at institutional level) is the lowest. This finding may indicate that students with A level entry qualifications are better-qualified than students with other qualifications. However, when using degree scores as a measure of degree quality, it is found that degree scores for the four groups are similar. For example, at institutional level, over the six years (1988-1993), the mean degree scores for the three groups, Access, BTEC, and 'other qualifications' are all around 51-55, and for the A level group is 56-59.

Furthermore, it is also found that that using different methods to measure degree results can have different value-added results. For example, when degree results are measured by percentages of good degrees, school 5 is second best in value-added and school 4 ranks five among the six schools, but when degree results are measured by degree scores, the result is completely opposite. The reason for this may be associated with distributions of degree classifications.

The comparisons between the method developed in this research and the main existing method - the comparative method, are also made. It is found that in general, although the two methods are different, the value-added results generated from the two methods are similar when the proposed method uses degree scores to measure quality of degree. However, when the proposed method uses percentage of good degree as a measure of degree results, the value -added results from the two methods are different. Compared with the Comparative method, the method developed in this research is much easier to use to calculate value-added.

A number of statistical tests were used to test the significance of the differences measured. The Kruskal-Wallis test, the analysis of variance (ANOVA) and the Tukey-HSD test are used to test the differences in mean A level points between different schools / cohorts and differences in degree scores between them. The chi-square test was used to test differences in percentage of good degrees.

By using the value-added method developed, the variations in value-added for major groups, schools and the institution are identified. These variations in value-added results raise a series of questions about quality of teaching of the institution during 1988-1993. At school level, with the A level group, for example, over the six year period, why did school 2 admit students with an average A level point score of 9.24 , and about half of them (52%) achieved a good degree, while school 6 recruited students with the highest mean A level points (13.61) in all the schools, but only 39.2 per cent of them obtained good degrees? It is also found that some schools' value-added varies with individual groups. For example, why did the Access students in School 6 achieve much more value-added than in other schools, but the value-added for the BTEC students of this School is much less than that for other schools?

This chapter focuses on measuring value-added. Answers to the above questions will be discussed in the next chapter. The next chapter will focus on investigating whether the value-added results can be used to indicate the quality of teaching.

CHAPTER SIX: VALUE-ADDED RESULTS AND THE QUALITY OF TEACHING

6.1 Introduction

Chapter 4 and chapter 5 focused on measuring value-added and used the model developed to measure the value-added. Some variations in value-added at the institution have been identified. This chapter will step back to critically assess the method developed and focus on investigating whether these value-added results can reflect the quality of teaching.

In this chapter, I evaluate the value-added measurement proposed by using a qualitative technique (semi-structured interviews) to investigate academic staff views. This is because it is academic staff who deliver teaching and ultimately are responsible for degree classification decisions. Therefore it is important to investigate their views on whether those value-added results can reflect the quality of their teaching. Thus this chapter lays emphasis on investigating acceptability and legitimacy of the value-added measurement as a performance indicator of teaching in the eyes of lecturers.

This chapter will look at what academic staff said about whether the value-added measurement can be used as a performance indicator. The data to be analyzed were the transcripts of interviews with 18 members of academic staff. It was central to the analysis of these data to identify problems or difficulties with using the value-added measurement as a performance indicator of teaching. This was done by categorizing the academic staff's views through noting elements in common, and then developing some theoretical abstractions.

The views of the members of academic staff expressed in the interview may be summarized in terms of the following seven themes, which indicate their concerns over the use of the proposed value-added measurement as a performance indicator of teaching:

- Conception of value-added
- Value-added and comparability
- Value-added, progress, and level of achievement
- Factors which have influence on students' achievements
- Factors which have impact on degree classifications
- Possibility of use of value-added as a performance indicator

This chapter will present the findings on each theme. Each presentation of a finding will be followed by discussion.

6.2 Conception of value-added

6.2.1 Findings

As explained in the methodology chapter, most attempts to measure value-added in higher education in the UK have been focused on comparing academic attainment of students entering the institution with their attainment on graduation. The value-added measurement developed in this study falls into this category. Therefore the concept of value-added here only refers to value reflected by degree results. The academic staff argue that degree results can not represent all value-added by their work. Students have other valuable assets as a consequence of their degree such as skills, quality of personal maturity, life experience, salary, employability. Some value-added may not be easily quantified, but it actually is important, and it is real addition in value.

> *I would argue that our students get as much value-added, but it is qualitative rather than a tangible measure of first, upper-second, and lower second...*

> *It is difficult if all this is being done in a quantitative way and you need things that you can measure, but I think there is a lot that students can gain from their course that you can't measure easily.*

Many members of academic staff also emphasized value-added through students' placements. They suggested that students developed personal maturity, obtained experience of working life, and skills in their placement.

> *where (placement) they get opportunity to apply what they have learnt in year one and two and they actually come back, quite a bit more mature, having gained a wider appreciation of not academically in the subject but experience of working life and area they are working for.*

> *... real value-added that our students get, is not quantifiable in terms of percentages, 61 percent (good degree) for this course. It is probably more to do with skills gained e.g. in the placements, experience of working and living abroad... They are not students that come out with the best marks. So we are adding things to these students' education experiences, skills which you can't quantify in terms of upper - second, lower-second etc.*

Some members of staff from a professional course questioned the value reflected by degree results. They argued that a good degree is not enough to make a professional (e.g. social worker). Professional expertise isn't entirely linked to degree level. Therefore, a performance indicator which is only based on degree results (at finishing point) would distort the reality.

> *...you could have people with what you wouldn't class as a good degree, who might actually be a very good professional social worker...it may well be that somebody with a upper-second may also have performed well as a professional social worker, but it is equally possible that somebody who'd got a lower-second or third may also have performed well as a social worker. Their professional expertise isn't entirely linked to degree level. So that would be the only worry I would have about that as a marker for something with a professional component.*

Some members of staff also pointed out the danger which may arise if the value-added measurement is only based on degree results.

> *...If you measure anything, the danger is that that becomes the thing people focus on , and they ignore other things.*

6.2.2 Discussion

The academic staff view, in that the concept of value-added based on entry qualifications and degree results is narrow, is consistent with the literature where some broad concepts of value-added have been suggested (see literature review). Nevertheless, what academic staff emphasized in terms of the concept of value-added, to some extent, is different from what is suggested in the literature. In the literature, the emphasis is placed on some comprehensive and theoretical concept of value-added. For example, Cave et al (1997) suggest that the value of higher education should include both private benefit and social benefit. They emphasized difficulties in measuring social benefit and concluded that in practice, we lack the capacity to measure all these benefits, In contrast, some of the academic staff 's suggestions about value-added are focused on the practical side of private benefits such as skills, personal maturity, or what students gained through placement. Therefore it seems that it may be difficult to realize a comprehensive concept of value-added as set out by Cave et al, but the concept of value-added need not be as narrow as it currently is.

There are two implications arising from the above finding.

First of all, the above finding suggests that there are some potential areas which a value-added approach may be able to take into account. Skill is an area which deserves attention. The findings of the interview suggest that academic staff regard skills as an important value-added to students, although some members of staff believe that some skills may not be quantifiable. More importantly, current higher education policy places more emphasis on skills. A recent government-commissioned report has recommended that employability elements should be built into the courses and that students' aptitudes in a

range of skills employers expect to find in graduates (such as problem-solving, critical thinking and analysis) should be assessed (Tysome, 1998). A large amount of money has been invested in more than 100 key skills projects that run across higher education (The Times Higher Education Supplement, March 1998). These developments would provide a basis on which measurement of value-added may be explored.

Students' performance in placement is another important area to which the value-added approach should give more attention. In theory, to perform well in their placement, students need subject knowledge, ability to apply their knowledge gained from teaching, quality of personal maturity, and skills such as interpersonal skills, communication skills, and problem-solving skills. In other words, students' performance in their placements can reflect what they have gained from teaching at the university, in terms of not only their subject knowledge, but also their skills and personal development -and these are not reflected by current degree classifications. Furthermore, the placement itself also adds significant value (knowledge, skills, and personal development) to students. Additionally, Sir Ron Dearing's committee recommends that higher education institutions, government and employers should expand dramatically the number of work experience opportunities. The placement becomes an important element of teaching.

In practice, some institutions have actually made the decision that the placement assessment will contribute to the degree results (10%) from 1997. This may be a way to take placement into account.

Nevertheless, taking placement into account can be a complicated issue. For example, it is known that in practice, there are huge variations in what students actually do in their placements. Some placements may not be directly relevant to the subjects which students study.

The second implication is that the above findings raise the question of defining specific value that higher education intends to add to students. Although the

first implication suggested two areas of value-added which the value-added approach should examine, it is still necessary to define what specific value higher education is intended to add to students. The academic staff emphasized that there are some values which may not be quantifiable. One would accept this view. It seems difficult to say that we can quantify all value-added to students. Nevertheless, we should be clear about what value higher education intends to add to students, and what value is measurable and what value is not quantifiable. For example, if academic staff argue that degree results can not represent all value they add to students, then what value does the degree result reflect?

6.3 Value-added and comparability

6.3.1 Finding

The academic staff argued that the diversity and flexibility existing under the current higher education system make it seem impossible to 'compare like with like' even at course level. The differences are summarized below.

6.3.1.1 The differences between students with BTEC entry qualifications

As explained earlier, since the only information obtained for BTEC students is that these students entered the university with the BTEC entry qualification, the starting points in terms of academic attainment are, therefore, assumed to be the same for all BTEC students. However there actually are grades for the BTEC entry qualification in terms of pass, merit, and distinction. Many members of academic staff suggested that the differences between BTEC students should be taken into account.

> *...so there would be a way of separating out the goodness of one BTEC student compared with another, and again a bit like you did with the A level points it could be calculated on a proportion of distinctions, to merits combination or a ranking of distinctions, merits, pass and therefore working out an average score based on that. So the BTEC students could be differentiated year on year.*

With a transfer policy, the members of academic staff from the School of Engineering reported a complicated picture with some engineering courses. They suggested that students who are at the end of the year on some courses, will not all have come in with the same cohort. For some courses, A level students come in, start year one and then they carry on, while

> *BTEC students could take one of the two routes. Firstly, they start year one on the BTEC HND course, at the end of year one some students who are very good will be transferred into the second year of a degree course and they would then carry on. Now the other BTEC students would complete their second year of BTEC HND course and some of them would then join the second year the of degree course*

The staff suggest that it would be interesting if we looked at these groups separately and there might be a change in the statistics because those who do a full HND first are obviously thought to be less able at the beginning.

6.3.1.2 The differences between students who come from access courses

It has been found that a major group of students, who study on the courses of BA Applied Social Studies and BA Applied Social Studies with diploma in Social Work courses, are from Access courses. These students usually take an Access course in a Further Education college first, then if they pass the Access course, they are offered a place to study a degree course at the university. Although access students are not graded, the members of academic staff argued that within the access student group, there are huge differences which are not reflected by statistics.

> *...I think what these figures don't show is a good deal of internal differentiation between those access students, so while you know the numbers of access students each year as a proportion of the total, you've got no way of actually discriminating between them.*

An access student could have been on a variety of courses and start a degree course at university with different credit levels. These differences would have some impact on how well they do at university.

> *...Within the access group, there are huge differences in what students come in with. They could have been on a variety of courses- full-time, part-time. But in terms of credit levels for entry, the higher education certificate is given to an access student with 16 credits... Now a lot of students on access courses do actually come out with far more than 16 credits. They may come out with as many as 36. So there is a huge difference within that group in terms of how many credits that they have got, and you would have thought that would have some impact then on how well they do here. Students who've got experience of coping with a lot of work, all at the same time, come out with a lot of credits, because the credits relate to the volume of work, rather than the level. Then you would expect those students to possibly be better prepared...*

The differences within the Access student group were also reflected in the fact that some Access students' starting points were so low that they were not capable of studying for a degree course at the university. This is due to a particular admission policy. There are some arrangements between the university and a local Access course, which guarantees students who pass their Access course a place at the university.

> *We had one year... where some of the Access students did very, very poorly indeed and we think they, we know they came from one particular access course in one particular college and we had to go back to that college and say that we were not at all happy, with the way they were operating.*

> *...There are a number of students (access) every year.. who have been supported by a course, an Access Course, who said we are going to get*

these students in, regardless of outcome...

The staff felt that

> *as a teacher here, as a lecturer here, I am therefore dependent on those students' teachers in the FE college making a good judgment about whether they are going to cope with my degree course.*

6.3.1.3 Hidden differences behind degree classifications

It has been argued that degree classifications may not be comparable between institutions and between sectors. Here, the academic staff also argued that there are differences between students with the same degree classification even when they are in the same course. This difference is that when two students achieved the same degree classifications, some of them may achieve it straight away, while some may have taken resit examinations. This kind of difference is not reflected by degree classifications:

> *For example, in year one a student may take two resits, in year two he might have one resit, but in his final year he might get a 2-1. Now you are just looking at 2-1, you are not looking how he got there and another student goes straight to 2-1 .*

> *...a lot of students do not pass all the first year exams all in one go, neither in second year, but the degree classification does not at the moment reflect that.*

The staff argued that although two students may achieve the same degree classification, the value-added for students who take resits is more than for a student who does not have resits.

> *... there are certain things that need looking at which in a way show the value-added for a student that takes resits is a lot more than a student who doesn't have resits.*

The academic staff emphasized that resit examination policy could change a picture completely, when comparisons are made between years or institutions which have a different examination policy .

The other thing that has changed is, originally if they (students) failed the first year course that was it, they were out. The university changed that policy, so the high failure rate went down because students were able to progress through.

...in some places if for example you were in year 2 and they did not allow any resit examinations, the students might then have to be left with no other situation than to repeat the year, which then changes the picture completely... in our set up, if a student fails 2/3 subjects in year one then we give him the opportunity to resit, the same happens in year two... Are we assessing the value-added as we are?

6.3.1.4 The differences that exist within the A level student group

Within the A level group, some students come in with relevant A levels to the subject they are studying while some of them do not. Therefore, those students starting with non relevant A level have to start off from scratch with the subject.

...some students enter with relevant A levels or GNVQ qualifications and others come in with non-relevant qualifications. If they come in with relevant A levels and that has been built on, then the value-added will be less, perhaps, than a student coming in with non relevant subjects, because they have to start off from scratch with the subject.

... students who come in with non relevant A levels have a much harder task on the course than those that have done the A levels before.

6.3.2 Discussion

One implication arising from the above findings is that it provides qualitative evidence that demonstrates more diversity and complexity exists in entry qualifications and degree classifications in higher education in the UK than is described in the literature. Thus it reinforces the problem about comparability of entry qualifications and comparability of degree results.

In the literature (e.g. Cave *et al*, 1997), it is argued that although the external examiner system is designed to maintain standards of degree across institutions, the credibility of the external examiner system is questioned (Silver et al 1995), thus it is questionable whether degree classifications are comparable between institutions and between sectors.

The above findings suggest that the problem about the comparability of degree results exists not only at a high level (such as sector level or institutional level), but also at a low level (departmental level or course level). It is found that even within an institution or within a course, the same degree classifications may still not indicate equivalence of achievement. For example, two students both achieve 2-1 degree class, but one may have taken a resit to gain a 2-1, while another has not. The academic staff argue that in this case, the same degree class does not indicate equivalence of achievement. They also argue that BTEC students could take different routes to study on the same degree course.

As far as comparability of entry qualifications is concerned, the discussions in the literature (e.g. Barnett, 1988) have concentrated on the comparability of A level grades between examination boards and over time. The finding of our research suggests that we also need to be concerned about the comparability problems which exist between other entry qualifications such as between BTEC students and between access students.

In short, the above findings indicate that in practice, the reality of entry qualifications and degree results is not that tidy or clear cut but is very diverse and complicated. More importantly, such diversity or complexity may be inevitable under the current higher education system in the UK since the current

higher education policies encourage diversity and flexibility such as the credit accumulation and transfer scheme and widening of access to higher education. Therefore one can't expect a tidy reality, and one can't expect to reach a high level of accuracy when measuring teaching performance. We should bear in mind this diversity and complexity of entry qualifications and degree results when we construct and interpret not only a value-added performance indicator, but also other performance indicators of teaching such as the progression rate.

Another implication of the above findings is that the recommendation should be made that the information on BTEC grades should be included in student data. According to the member of staff from the University Student Records Office at the institution, the information on grades of the BTEC has never been required for students records by either the Further Education Statistical Record (FESR) or the Higher Education Statistical Agency (HESA). However, as noted above, the grades of BTEC students can be used when measuring the starting point for BTEC students and this measurement would help new universities to monitor their students' progress. It should be emphasised that 'entry qualifications' do not equal 'A level entry qualifications'.

6.4 Value-added, progress, and level of achievement

6.4.1 Findings

One of the academic staff's concerns about the value-added measurement is that some courses generally recruit students with lower academic attainments than other courses, but these courses would generate better value-added results. This, they felt, would distort reality.

> *...the entry requirements between different degrees vary a lot, students need about 18 points at A level to get onto this degree whereas there might be other degrees and they might need a much lower standard of entry. So I think that a distortion would occur in that courses which accepted students with much lower entry requirements would generate some good results, because that would be translated in terms of a bigger value-added, assuming that they came out with the same kind of spread*

of degree qualifications at the end of the degree, So I think there could be quite a lot of distortion if you compare between different courses and degrees.

Some members of staff worried that schools which recruited students with higher academic attainment would feel discouraged.

For example, a school with 24 input and another school which has 8 points input. You look at the degree that they get with the average, the numbers getting good degrees may be similar, if they turned out to be very similar then can you say that you have more value-added ? ... That might be a dodgy one, if you say something like that, it might upset people in the business school for instance

Some members of the staff gave their own explanations. For example, one member of staff said

you have to look at it in a way that if the final goal is to produce a graduate, who can work in industry and we train and they work there and are successful, the other universities produce it and they are successful, our students don't in anyway show that they are not properly equipped to do that job. So we have achieved our objective, even though our starting point is lower.

Another member of staff suggested that at individual student level, students with different academic attainment at entry actually study at a different intellectual level. Students with higher academic attainment made 'very definite intellectual progression', while students with lower academic attainment 'still operate in too simple-minded a fashion' even though they struggled through study and obtained a degree. Students, whose academic attainments were in the middle of range, lack the capacity for the more penetrating intellectual thought.

I think we see some students, particularly ***at the top end of the range*** (students with higher A level points)*, where we are conscious of very definite intellectual progression here, that they really do make academic progress, that they are ready, you know, when they've graduated to go on to post-graduate study, and looking at their performance we can see* ***a clear improvement over time.*** *That's absolutely the top end of the range.*

At the very lowest point of the ability range (e.g. students with lower A level points)*, there are students who* ***struggle through*** *each year with referrals, and who are going to come out with a very low 2:2, possibly a Third, where* ***one feels that they're not really operating at the sort of level you would expect intellectually of a university graduate.*** *I mean these are a very small percentage I'm talking about, but they stay in the memory because one is very worried about them, because they just never seem to get it intellectually, in terms of understanding the more sophisticated nature of questions, the more discursive elements or the finer points of theory. They still operate in too simple-minded a fashion really, and it's like, however hard you try with them, they really don't get it.*

Then there are a number of students, there are a lot of students one gets in ***the middle of that range.*** *There are some* ***who always get 58%,*** *they always get a good 2:2 but never a 2:1, and sometimes these are very conscientious and hard-working students, and you do the best you can to help them but* ***they just haven't got the intellectual sparkle and flair. They just haven't got the capacity for the more penetrating intellectual thought.*** *I know that there are some who work steadily and do come out with a 2:1, so a lot of our students obtain satisfactory degrees, but not brilliant degrees. These are the people who perhaps feel they need to get a degree to get a job, they're not*

necessarily sparked by the love of learning or anything like that! And so with a lot of those, ***there perhaps isn't a lot of value-added*** *but they plod through it at a fairly pedestrian way and come out with a high 2:2 or a low 2:1.*

However, the academic staff also argue that they had spent more time and made more effort to help students with lower academic attainment. Their teaching process had to be specially designed to help those students to catch up.

Now we would try to pictch our teaching on the average of these students so that we then have established a base from which we can then go onto teach other subjects or subjects at a higher level. In order for the students to follow certain things, we would do the basic things so that they can understand what is being taught...

... I will go back 2 or 3 steps to pick it up, which would mean that the students those who have done it and forgotten, will have their memories jogged and then understand and those who haven't done it will then learn it and they will not be totally lost.

In order to enable the students with a lower starting point to achieve similar degree results to students with higher starting points, academic staff have to spend more time with them to give more help. They argued that to them, more help means giving those students more value. The students themselves have to make more effort as well.

...The students (with lower starting points)have to be given more help and to me more help means I am having to give them ***more value*** *because he gets more tuition, I have to give him more, he has to put in more, so it is more effort.*

Many of them (access students) do not necessarily have the skills in essay writing, some of them may not have the skills in say computer literacy, those kinds of things, they need more time, you need to give them more time.

One member of staff made a comparison in tutorial time between two different institutions.

If you look at it from a time point of view, my son who falls into that category for example, he's got 24 A level points and in the 2nd year of his degree at the university of A and he probably gets tutorial time of 8 to 9 hours per week at the most, and he probably gets that for 20 weeks, but the engineers down here who are coming in with this (average 10 points at A level and less) are getting 15-18 hours per week ,over 24 weeks, big difference in time...

He also argued that his students often did not prepare for their tutorials.

He (his son) has to read a lot more. They are expected to do a lot more of their own reading and their own research and their own studies , so he will spend a good half day preparing for his tutorials, where of most of this lot (students here)don't come and prepare for the tutorials.

6.4.2 Discussion

The above analysis suggests that on the one hand, the academic staff concerned about the value-added measurement may be in favour of courses (schools or institutions) which recruited students with lower academic attainments; on the other hand, they recognized they had spent more time and made more effort to help students with lower academic attainments and they have given those students more value.

The implication of the findings is that it suggests an alternative explanation of the statement *'value-added scores are negatively correlated with initial scores, therefore a low initial score will generate a high value-added score'* , as

discussed in the literature review.

The image now portrayed is that academic staff have made more effort to help students with a lower starting point catch up their study and students themselves have made more effort too. The effort enables students who enter university with lower starting point to reach a similar level (i.e. degree class) as students who have a higher starting point. Therefore the students with a lower starting point have more added value compared with those who start from a higher level. Value-added reflect these efforts.

In my view, we need to distinguish between **progress** in academic attainment and **level** of academic attainment. Value-added reflects the progress made in academic attainment, but may not reflect the eventual level of academic achievements. Thus, for example, if student A had a higher value-added score than student B, this may be interpreted that student A made more **progress** in his / her academic attainment than student B, but the **level** of his / her academic attainment may not be necessarily higher than student B. It is, as one member of academic staff said in the interview, students with different academic attainment at entry actually study at a different intellectual level.

This interpretation may make it easy to understand that courses or institutions which recruit better-qualified entrants may have lower value-added scores.

6.5 Factors which have influence on students' achievements

'You can take a horse to water, but you can't make it drink'. The academic staff argued that there wasn't a one to one relationship between the quality of teaching and what students achieve, because there were other factors that come in to it. There are things to do with the students' ability, there are things to do with motivation, there are things to do with financial situation, family influence, mental health...etc. Their students could achieve better academically if they had more money, and didn't work three or four evenings a week in a supermarket. But, individual academic staff can't personally do anything about that. On the other hand, when value-added for a course improved, that might

be nothing to do with teaching, but might be to do with other things that were happening, for example, with the aura of class.

The factors mentioned by the academic staff may be summarized in terms of the following categories: motivation, the aura of the class, some attainments which are not measured by entry qualifications, the domestic situations of the students, increasing amount of time devoted to paid employment in term time, social class, and management of teaching. Most of these factors are student - related factors, but the last factor is related to management of teaching.

6.5.1 Motivation

Students' motivations are seen as an important factor which has great impact on students' achievement by most of members of staff. The staff perception is that whatever students started with, if they are highly motivated, they often obtained a good result.

> *One of the largest things; I mean obviously ability is a key factor in terms of how fast you progress; but I think possibly equally as important is motivation. And whatever students come in with, if they're highly motivated they generally come out with a decent performance at the end.*

The members of staff found that mature students usually were highly motivated. They themselves really wanted to study at university. They want to study for themselves and may be their children, while non-mature students might not be personally motivated although they had a lot of advantages to study at university. Motivation is a special advantage for mature students in studying at university compared with many non-mature students.

> *'In many ways I think mature students; and many of the access students are mature students; in many ways I do believe that they do have a certain advantage. The advantage that they have is that most of the time they really want to be here and in a sense they are a joy to teach - because they want it for themselves and they want it for their children. Whereas many of the A-level students, it's expected that they will be*

here. They may not necessarily want to be here right now... But they are not perhaps motivated within themselves - they've gotten everything else going for them, but they may not want it very badly.

6.5.2 The aura of the class

Highly motivated students on a course may influence other students on the course and the whole class can start to work hard. In such case the aura of class changes . A good aura of class could lead to improvement of value-added for the course. 'Aura' here refers to group dynamics or a shared academic culture.

I think it is all down to the, something about the aura of the class, that if you get a group of students who are clearly motivated and working hard, that tends to rub off on other students, and then the whole group starts to work hard.

'We have got in particular last year, a significant number of Chinese / Malaysian students on our degree course... they potentially will raise the profile of the course, they will raise the work rate of the course, they are very dedicated, they will have an influence upon the UK students, and the UK students, performance and work level because they will maybe change the ethos of the course a little bit . Now that will not reflect in your results, will it... When you have got other things like that are impacting upon the course, then you could say well this has improved the value-added because of this other external influence but that might be nothing to do with the teaching, that might be to do with other things that are happening'

Equally, a small disruptive element could destroy the entire class and that could cause a problem, where a few mature students could transform the atmosphere in the classroom.

A lot of it has to do with the atmosphere in the classroom and that

depends on the personalities of that particular group of students. We tend to find that perhaps there is 10% of mature students in the class, it completely transforms the class because the mature students are bold enough to tell the younger ones to shut up, if they are being disruptive - it is quite helpful actually. So I think that there are a number of influences and this is probably one.

You sometimes find that you get years of students that just don't seem to gel, the groups don't fit together very well, the students don't seem to be integrating very well - and that seems to happen some years doesn't it?

6.5.3 Some potential attainments which are not measured by entry qualifications

Some students may start a course with less than the ideal academic attainment, but they may actually have some other attainments, experience or social science aptitude which are not measured or reflected by entry qualifications and have a positive impact on students' achievements. In this sense, these students' starting point is higher than their entry qualifications indicate. The academic staff found that reading widely in the area of their subjects and being involved in local political parties, or environmental movements could benefit students on social science courses:

...as a part of conscious policy, I know, not just for under-graduates, but when I was responsible for the post graduate course, accepted people onto the degree course with less than the required academic attainment... and I can vividly remember one particular person who didn't even have a first degree, and I think had two A-levels and that was it, and some experience as a teacher, but who clearly had read very widely in the area of the Masters, and I'm pleased to say, touch wood, he came out with a distinction. Now, if we were looking at it in terms of your figures, that would represent an enormous value-added - what I'm really saying I suppose is, that, well, am I really measuring value-added there

or am I measuring a mismatch between measures of attainment, because in a sense his attainment was much more than was indicated by the formal measures in the first place'

...particularly with mature students, often with mature women... They're actually quite knowledgeable, let me put it that way. Either through reading or, in the case of social science, on the basis of their experiential learning. ...quite often it's the case that if somebody's involved in, say in, one of the local political parties or in one of the environmental movements or in a voluntary organization, maybe to do with children, if they have children and so on; that they quite often have, what we would measure in terms of social science aptitude, quite a lot of skill, quite a lot of competence. That isn't measured in any sense in terms of A-level attainment, because in some cases they may not even have taken A-levels. Now this would often be the case with Access students, because they're coming in out of there. But it just occurs to me it's less likely to be the case I think with students coming straight from school...

Foreign travel can benefit students who do courses which have international elements.

'Another one is actually foreign travel. People that have been abroad, I know a few over the years have been abroad, worked abroad and so on and so forth, who come in with, as a result of that, quite considerable expertise in terms of... if they're then going to do something on the more European and internationally focused elements of our units, do quite well on those... So, it strikes me that there is again an area where you need to take into account at least, or think about and discuss, the implications; particularly on the social science course... about the input measure needing to be at very least set alongside the potential for their being other attainment which is not measured on that, and what the implications of that would be in terms of your input and output measure.

That in a sense the value-added maybe either over or under estimated.'

6.5.4 Domestic situations of the students

The domestic situation of students may be difficult to measure, but it has great impact on mature students' achievements. Students' academic lives tend to interact with their whole domestic situation. Problems with a mature students' marriage or children could seriously influence a student's progress and can often lead to a student dropping out.

> *What you couldn't measure is equally important, and that's things like the domestic situations of the students; whether they've got broken marriages; children; no children; financial difficulties or whatever; ...what tends to happen when they get here is that there's an interaction between the whole domestic situation and the academic life here, which often leads to many students dropping out. And that's something that is very difficult to measure.*

> *We have known of cases of students where the marriage is breaking up because the partner can't cope with the new person they've become. I mean in some ways if that's happening it's better if the break-up comes before they start the degree programme...*

> *And sometimes, too, people have very real problems with teenage children, who are perhaps getting into trouble with the law or being tearaways - and that's very, very difficult. Or indeed even sometimes possibly grandchildren...*

Their domestic situation particularly has serious influence on women mature students, especially those from a working class backgrounds. Woman mature students often play the traditional role of woman in the family and try to hold the family together. They take too much family responsibility and have little time to

do their coursework.

But in some cases, particularly I think for mature women students, there can be even active hostility from the family... I think one of the problems for a lot of working class students is if they are the first person in their family to go into higher education, the family may not appreciate the demands of higher education, they may think it's a bit of a skive being a student, not see it as a full-time job. So, women mature students in their 40s, for example, may be asked by their daughters to look after their grandchildren or something like this, when they should be doing their college work.

I knew of one student who was refused money to repair the washing machine when it had broken down, or refused money to repair the cooker, so she was expected to cook and wash without a functioning cooker or washing machine...

6.5.5 Amount of time of working in term time

Many members of staff expressed their worries that there was an increasing amount of time students are working in paid employment to support their studies.

I think one of the things many people in higher education are worried about now is the increasing amount of time students are working in paid employment to support themselves through their studies, and we don't see that getting better as a situation, but it is detrimental for the students. When I was a student, people may have worked in the holidays, but they didn't work much in term times...

...many of them are working for money, for employment, many of them are forced to develop increasing amount of their time to work, mature students who have families and a range of responsibilities outside the course and need to work, not just in the vacations as they used to but

during term time, often for very large amount of time..

Not only are part-time students working, but also full-time students and students who have a grant are working. Students are not only working at weekends or in the holidays, but also working in term time. This has seriously influenced students' study. They leave class early and do not have time to read broadly and discuss questions with each other.

> *...not just in the vacations as they used to, but during term time, often for very large amount of time that must in someway effect their performance... How that could be measured in value-added terms, one assumes that younger students who don't have conpound responsibilities and students who are from better off family backgrounds or social class would be less effected than others, one assumes that because they would have, likely to have more money and therefore less likely be working, having to work during term time*

> *... whereas now because students are working two or three evenings a week, or they're leaving classes early or whatever, they don't have the time for the broader reading, they don't have the time for discussing ideas with other students...*

Sometimes financial difficulty causes stress. So financial difficulty is coupled with illness.

> *...financial worry is an area that's on the increase, and by the time we look at this year, 1997, I would expect that value-added, probably, is definitely going to be going down because there is increased financial worry on the part of students. ...We also have an increase in the number of students who are on medication. So their financial worry is coupled with illness quite often - stress, stress-related illness. So there's a great deal of that that's going on and I don't know who's picking that up, but I get extenuating circumstances, and the extenuating circumstances - they come with doctor's notes, and the doctor says well so-and-so is on this*

kind of medication, which is problematic in terms of their performance. So there's a lot more of that.

6.5.6 Social class

The social class backgrounds of students could have some impact on their studies. Some students coming from a working class or less scholarly background lack some of the skills, knowledge or study habits which the university system expects that they already have; for example, the skills to collect information and writing skills; the habit of reading newspapers or watching news. These students have some disadvantages in studying a social science course, while those students from other particular backgrounds (e.g. middle class) are already prepared with these skills, knowledge, and habits.

I believe that the university system is geared for people who come straight in at 18 after having done A levels and those whose parents have been to university or in some way have a particular kind of class background which means that they know how to get information, which means that they know how to write in different sorts of styles, which means that they're familiar with doing things like reading the newspaper every day, and watching the news and discussing politics, discussing social policy - those students have an advantage in a social science course.

So all too often, people coming from an access background, coming from a background where they have been miseducated or perhaps have had problems in education, they have not had that kind of practice at discussing social issues. They've not had, they've not developed that discipline, perhaps around the dinner table of discussing the news or the elections or what have you.

Students from a middle class background usually expect to go to university, but students from a working class background may never have had the expectation

that they would come to university. That expectation is very powerful.

> *I believe that that is something that in general, if you have a middle class background, if you've been comfortable doing A-levels in this country, if your parents have been to university or at least have done A-levels and are professionals, you come in already prepared that way. And then, icing on the cake, it is expected that you will go to university. And that expectation is very powerful. Those people who are on access courses, mature students, there was never really an expectation that they would come to university... So they have a lot more that they have to do, and they have to have a lot more inside to keep them going through the difficulties*

6.5.7 Management of teaching

Some factors which may have an impact on students' learning and achievements are related to the management of teaching. These factors include time arrangement of the first year of course, flexibility in scheduling teaching time, course administration and management.

6.5.7.1 First year of course

The first year is important for new students. Students need more time because they need time to know the system and mature students often have a lot of problems to sort out.

Nevertheless some of the university policies, for example, the norm of the 'twenty credit unit', may fail to appreciate what students' need and therefore affect their progression.

> *I think the first year is important because students have to find friends in the institution. Not only friends of their own age, but they need to know who's a counselor that they could go to, they need to know who their year tutor is, they need to know, are there some sort of informal systems*

of monitoring around that they can benefit from. So all of that is very important. I think that they need to learn how the system works, the formal system as well as the informal system

' In a twenty credit unit, ...So they (students) start in September, they're examined in January and the whole of a year's work for xx(subject A) is fitted into that. So we're now expecting them to reach if you like end of level 1 standard by January whereas we used to expect them to reach it by July. So they've had a much shorter time in which to develop, you know get familiar with the institution and so on, and that I think is quite ... might have an effect on peoples progression because you can appreciate that, particularly with mature students with lots of problems that they need to sort out, they start in September and it may well take them until October to get their travel plans sorted out and their childcare problems - by that time, a third of the course has gone.

6.5.7.2 Flexibility of schedule of teaching time

A flexible schedule of teaching time can help mature students to combine their study and their family responsibilities. Therefore it can make mature students' academic life easier.

...in the past we used to make attempts to accommodate mature students far more than we are able to do now. We used to be able to timetable them between 10 and 3 in the afternoon so that they could do their education around the children being at school, we used to take account of half terms, we used to give them reading weeks when the children were at home for half term holidays. What's happened increasingly over the last few years is that the university, the centre as it's called, has imposed a more and more rigorous schedule on terms, semesters, assessment and also the teaching day. We're now technically teaching 9 in the morning till 9 at night. It's become increasingly difficult for us to accommodate mature students and take account of their special needs. So although I think individual teaching

staff do spend more time with mature students when they need it, there's now a limit to what we can do to make their academic life smooth for them.

6.5.7.3 Course administration and management

Good practice in course administration, such as a course directory, or a student advisor, can help arrange student study in terms of time and choose their study units. It can make both students' and academic staff's lives easier.

I believe that we need to have people who are advisors to students about their programmes... Someone who knows the system, who knows the courses, who knows how to get the information... Instead of the madness that we have now! So in other words what I'm saying is that you need better administration, and that the administration is not just about dealing with the marks, but there's administration in dealing with the students...

I think also that we need to have a course directory... This course generally runs, this unit, Unit A runs generally Friday from 10 till 12. So that students know from the very beginning what their timetable is, so they could put together a timetable in their head, they know what's on, they know what day it generally runs. I don't know why we don't do that. It would make life so much easier for the combined studies students for instance! And for any student who wants to exercise their choice.

One of the most embarrassing things that happens of course is that students will come along and they're lost, they're knocking on the door and they go, "I'm looking for so and so's class" and you don't even have a clue about how to begin to help them, really, not unless you're lucky enough to belong to that programme... So then you have to bother the people at reception, and they won't know, and that's too bad. But if you had a directory which had all the courses and you say, "oh well, so and so"... And then also when you're sitting down with students who want to

follow on, say for instance do unit A at level two, they want to do something at level three, then you're able to give them the information that at level three there's this - it's in another school, but here it is. I think then you'll find that value-added would even get better.

6.5.8 Discussion

The academic staff argued that quality of teaching could not be measured by what students achieved, because there are other factors which could have an influence on students' achievements. Highly motivated students may often achieve good degree results, no matter what they come in with. A few hard working students could influence other students on a course and change the aura of a class. Value-added for the course could be improved because the whole class starts to work hard. Some students start the course with less than the required academic attainments, but they actually have other potential attainments which are not measured by entry qualifications, but which have a positive impact on their study. Mature students' domestic situations such as broken marriages or problems with their children, interact with their academic life and often lead them dropping out. Some students work three or four evenings a week, leave class early, do not have time to do broad reading and discuss questions with others. Individual members of academic staff personally have limited control over that. Factors such as social class and management of teaching can also have an influence on students' learning and affect value-added. Therefore when a figure shows that value-added has improved, that might not be to do with teaching, but it might be to do with, for example, the aura of the class; and when a score indicates less value-added to students, that might be because students are working an increasing amount of time in term time, and it might not indicate poor quality of teaching.

There are three implications arising from the above findings.

Firstly, the above findings provide qualitative evidence which that the value-added results can not be directly used to indicate the quality of teaching. When we attempt to interpret a value-added score, we should be aware that it could be that these potential factors rather than the quality of teaching have impact

an on student achievements. Thus the findings reinforce the view in the literature where Johnes and Taylor's research (1990) suggests quantitative evidence that inter-university variations in degree results cannot be attributed to inter-university differences in quality of teaching, because over 80 percent of the variation in degree results is explained by six student-related and university-related factors.

Secondly, the findings above also suggest some factors which may have an impact on student academic achievement, but which are different from those suggested by previous research. The most recent research in this area is Johnes and Taylor's research. They found that two student-related factors and four university-related factors can explain the variation in degree results. These factors are:

- *the mean A level score of entrants;*
- *the percentage of students who live at home;*
- *library expenditure as a percentage of total expenditure;*
- *whether or not a university is ex-college of advanced technology;*
- *whether or not a university is one of the new greenfield universities established in 1964/65;*
- *and whether or not a university is located in Scotland.*

Nevertheless, as Johnes and Taylor emphasized, only those variables for which data can be readily obtained are considered in their study. Their study is carried out at institutional level, while the present research is based on individual course or unit level.

There is other research in this area. Table 6.1 summarises the findings of the research. Comparing the factors suggested by previous research with the factors by the present study, it can be seen that some factors are overlapping. For example, the factors 'motivation' and 'social class' (or parents' education) which are mentioned by academic staff in this research have been studied in previous quantitative research. However, the present study also suggests some

new factors which have not previously been examined. These factors are

- *the 'aura' of the class,*
- *some attainments which are not measured by entry qualifications,*
- *the domestic situation of the students,*
- *amount of time of working in paid employment in term time,*
- *and management of teaching.*

Clearly, the two factors, 'the domestic situations of the students' and 'working in term time', are associated with recent changes in higher education in the UK, such as widening access to higher education. Nevertheless, it should be stressed that the findings in this research are based on qualitative evidence and the effect of the factors suggested in this research need to be further examined through quantitative methods. Although it may be difficult to directly use a sophisticated model to construct a performance indicator, the sophisticated statistical techniques, (e.g. multilevel modeling) may be used to investigate these factors in order to obtain some insight into the factors which have an impact on student achievements.

Table 6.1 A summary of the factors studied in the literature

Student-related factors:

- School qualifications *(Sear, 1983, Entwistle and Wilson, 1977 and etc.)*
- Type of school attended *(Barnett and Lewis, 1963)*
- Age *(Eaton and West, 1980, and etc.)*
- Sex *(Rudd, 1984)*
- Motivation, and career aspirations *(Bailey et al., 1986, Entwistle and Wilson, 1977)*
- Personality and study methods *(Entwistle and Wilson, 1977)*
- Place of home residence *(Entwistle and Wilson, 1977)*
- Parents' education *(Entwistle and Wilson, 1977)*
- Subject of degree *(Nevin, 1972)*

University-related factors:

- Type of accommodation *(Williams and Reilley, 1972)*
- Student-faculty relationship *(Centra and Rock, 1971)*
- Flexibility of the curriculum *(Centra and Rock, 1971)*

Thirdly, another implication of the findings above is that the interpretation of value-added results should be linked to the process of learning. Table 6.2 shows that we can classify most of the factors suggested in this research into different aspects of learning. For example, the factors, domestic situation of the students; and the amount of time of working in term time are especially associated with the time students spend studying. If we were to have the measurements of all aspects of learning, the value-added measurement could be used together with these measurements of learning. For example, were both the value-added scores and students' study time increase, we might say that the improvement in value-added scores may be due to students working harder.

Another point which is worth noting is that the findings suggest that academic staff made a distinction between performance of administration / management

Table 6.2 The relationship between the factors and learning

Factors ⇒	**Learning** ⇒	**achievement**
the aura of the class	Motivation	
some attainments which are not measured by entry qualifications	academic abilities	
social class	study method / skill	
domestic situations of the students; amount of time of working in term time	study time	
social class	study habits	

of teaching and performance of teaching. They suggested that there are some factors related to management of teaching which have an impact on students' achievements. This may suggest that we should distinguish between performance of teaching and performance of management / administration of

teaching.

6.6 Factors which have an impact on degree classifications

6.6.1 Findings

From the interviews, it has been found that some members of staff have a strong negative attitude towards using value-added as a performance indicator. They argue that degree classifications do not necessarily indicate how good the students are or how good teaching is. There are many factors which interact to influence degree classifications. Therefore when a figure / score shows the improvement of degree results, it could be a result of the improvement of teaching or be a combination of the improvement of teaching with all the other factors or variables.

The factors suggested by staff may be summarized into the following categories: 'academic community', 'the way the examination boards behaved', 'attitudes of external examiners', 'how we assessed students', 'standard of marking', and 'coaching students for examinations'.

6.6.1.1 'Academic community'

The members of academic staff believe that what degree class you get is dependent on what subject you study because different academic communities assess their subject differently. Some academic communities are reluctant to use the full range of marks (e.g. Business Studies), while others may use a higher mark for a good paper.

> *...it almost becomes irrelevant as to what degree classification you get, and as to whether that is value-added, it is surely dependent on what degree you study for and therefore if you go to science, and you see all these people who had 2 points to get onto their degrees and then there will be lots of people getting firsts and 2-1s because of the nature of the subject and because of the academic community thought on it.*

I mean how come nobody gets more than 70 on an assignment in business and management and it is nothing to do with the quality of the paper relative to someone in science who may get 100%. It is a difference in approach that the academic community take to their subject and therefore if somebody gets 75%, you are probably saying that it is about as good as it will ever get but there I am still reluctant to use the full stretch of marks.

6.6.1.2 'The way the examination boards behaved'

Examination boards have the right to decide what makes a first a first, and a 2-1 a 2-1. The same examination board may have different requirements for the same degree classification for different years.

I do not think that there were any differences in the cohorts, I think it was different in how we assessed students, the way the examination boards behaved. For example an examination board one year, may decide that someone who is 58 should be given a 2-1 whereas the next year the same examination board with either the same or different external examiners said no. It is 60 or 59 . Same *thing with extenuating circumstances. Measuring value-added in terms of degree classifications is too deterministic; I think it is too dependent on the statistics. It has been driven by the statistics of classifications rather than what is actually being taught to the students.*

The members of staff reported that the examination board could even normalize the average mark for each unit and thus control the proportion of each degree classification.

...between 1989 - 90 there has been a step down in the percentage of good degrees and I can tell you why that has occurred.

Basically, the largest proportion of students on our degrees do marketing as their option, the Marketing external examiner was unhappy about the

proportion of 2-1s on the degree and therefore the program management at the time made the decision that all units should attempt to normalize around a 57% mean, so that then, the proportions of 2-1s from marketing would be the same as the proportion of 2-1s from HR and so on and so forth.

Now, I am not convinced that 57 was the right point to normalize around, but the consequence was that very significant step change between 89 and 90 and the consequence is, if you like, it has become part of the organizational learning that now you expect your mean to become 57% and therefore the proportion of 2-2s is significantly higher...

6.6.1.3 'Attitudes of external examiners'

The attitudes of external examiners could change the proportion of good degrees dramatically. The same course may have different external examiners in different years. Different external examiners may have different attitudes.

I am not sure how you would assess value-added in that process, when there will be wild fluctuations, from year 61% getting 2-1s to the next year less than 30% simply as a product of attitudes of external examiners. And then on top of that changes in course work, changes in structures so students can't choose the easy option. With so much change taking place over a five year period for example from 1992 - 1997. I do not know how you would separate out those variables and come to some kind of judgment about the impact/effect of entry qualifications, value-added etc., on final degree results

6.6.1.4 'How we assessed students'

Unit leaders and members of academic staff can decide what should be taught and what should be examined. The difficulty of examinations will influence the proportion of good degrees.

...for our students they find it terribly difficult to do quantitative things, if we had the course redesigned and input more quantitative elements, the degree classification would come down over night.

The degree results also could be influenced by the balance between coursework and examinations.

the balance between coursework and examinations for example, the kinds of subjects the students can choose, have had a big impact upon the distribution of marks at final degree level

6.6.1.5 'Standard of marking'

The members of academic staff believe that marking is a qualitative judgment. Marking leniently or tightly can change the proportion of good degrees.

if you were comparing value-added on one degree with value-added on another degree, you wouldn't necessarily know that they were operating exactly the same standard in terms of degree classification. I know the regulations are the same, but marking is a qualitative judgment I understand, so it may be that one course would give a 2:1 and another one would give a 2:2.

... the average mark for 'Unit A' was somewhere in the region of 53 %, the average mark for 'Unit B' was 60 something % . So you may have one student, taking into account that a student doing 'Course A' is likely to do better than a student doing 'Course B', this is primarily a problem because they have less access to the subjects that we know are marked more leniently because they are doing languages, the subjects that they do are notorious, that the teachers pride themselves on maintaining standards of academic rigor and so the mark that these students come out with in this subject would be 53/54 %, whereas an average mark in the other subject would be 64%...

6.6.1.6 'Coaching students for examinations'

The members of staff suggested that some members of staff overcoached students for examinations by giving students a large handout and telling them what was going to be on the examination paper. More importantly, it is difficult for external examiners to discover this. The only persons who would know this are tutors themselves and students. Students will not complain about it.

> *Anecdotal story, I discovered subsequently after the enquiry why I thought that we had got so many 2-1s and why all these 2-1s seemed to be grouped in one subject and it turned out that one of the major reasons was, a certain member of staff, was giving in coursework during classes, a case study on x x x* (the title of the case study*) and then lo and behold the same case study came up in the exam. Now the only person that would know that that was the case was the tutor and the students. Now the students won't complain and say we've got an easy question and the member of staff appears to be an excellent teacher because the students are coming out with excellent scores and the external examiner, who gets the paper, does not know that the students have done this in class, the rest of the examination board in other subject areas don't know that this is the case. So the external examiner will come to the exam board and say I don't understand why these marks are so out of line with other courses because the paper is rigorous, the marking is rigorous but the students just did better. So the point I am making, is that that external examiner, who was doing his job, did not know that the students were being coached and passed the examination by having exactly the same case studies in the examination as they had in class. Staff under pressure might do the same thing.*

6.6.1.7 The reasons for manipulations

The above has shown the idea that members of staff believe that degree classifications can be manipulated in different ways even though there is an external examiner system. The academic staff also express the view that if

value-added is to be used as a performance indicator - with formal incentives or penalties, or it affect funding, recruitment or salaries - manipulation would undoubtedly happen and manipulation could become very common and serious.

> *If you introduce performance indicators of that nature (use value-added as a performance indicator with formal incentives or penalties), then what will happen is that people will make their courses easier, make their assessment easier, and give higher marks.*

> *If it became a measurement that people took some notice of and it affected funding, recruitment, or salaries or things of that nature, then of course it could be manipulated, yes. Undoubtedly. And would be.*

The manipulation also could happen when academic staff are under pressure to do more research or to do external consultancy for 'income generation'. Since they have to spend more time doing other work than teaching, they may manipulate their students' degree results to cover own inadequacies of their teaching.

> *I know the pressure that staff are under, I know the enormous work load that people have got, I know for example, a large number of staff here, in this institution here, who see teaching as a minor irritation. They see themselves as researchers, the policy research centre.... and teaching is an inconvenience. You give them (students) as many handouts as you can to keep them quiet and when the examinations come in, you mark nothing less than 40, so you don't have to set another exam paper, consciously or sub consciously*

> *...a teacher with outside interest is making more money from consultancy than he is in the School A, but the last things they want are students and other staff complaining. The way you stop students complaining, is by giving them lots of handouts and telling them what is going to be on the*

examination paper.

Even the pressure from university quality assessment itself could lead the staff to manipulate degree results.

> *It is possible for people to feel under pressure in an institution like this, for example, when I was course leader of two degrees, I was constantly being asked by the centre why we had so many failures. And so the staff as a result of that, felt under pressure that we were doing something wrong if people are failing, now we may well have been. But it could also have been that the students weren't good enough, as that we had a standard that the students had to reach , and if the students didn't reach it, then they failed, or they went elsewhere or whatever. Now if people feel under pressure not to fail students in that kind of environment, so they say oh well my head tells me 36 but 40 we don't want too much fuss.*

6.6.2 Discussion

The output measurement of the value-added is an approach based on degree results. The academic staff argued that there were many factors which could have an impact on students' degree results, therefore the degree classifications did not necessarily indicate how good students were or how good teaching was. If a figure shows an improvement in degree results, it could be an improvement in amount of knowledge that teaching staff impart to students, and it also could be due to a decline in staff input, but staff compensate for this by rewarding students with a high mark. The academic staff have given many factors, and these factors have been summarized into six categories: 'academic community', 'the way the examination boards behaved', 'attitudes of external examiners', 'how we assessed students', 'standard of marking', and 'coaching students for examinations'.

Furthermore, methods of assessments vary hugely within as well as between universities and disciplines (HEQC, 1997). The higher education assessment

process rests on individual assessors' or examiners' decisions. Assessment could be traditional unseen papers taken by students at the end of their courses, smaller or larger course work projects, oral presentations, practicals etc. The balance between them, and the number of individual assessments which count towards the final degree are different between disciplines and institutions. All these variations in methods of assessment could have an impact on variability of degree results.

Some implications arise from the findings above.
Firstly, the findings have provided some qualitative evidence which shows that the picture of degree classifications of a course or an institution can change dramatically without any improvement in students' actual achievements, and the degree results can be changed by 'the way the examination boards behaved', 'attitudes of external examiners', 'how we assessed students', and 'standard of marking'. At least, this is the view of members of staff.

The factors discussed in this section are those which have an impact on whether degree classifications can accurately reflect what students actually have achieved. These factors are different from the factors in the last section. The factors discussed in last section are those which have an impact on what students can actually achieve. When we attempt to interpret value-added results, we need not only to be aware that some factors other than quality of teaching can have some impact on student achievement, but also to question the way 'the examination boards behaved', 'attitudes of external examiners', 'how we assessed students', and 'standard of marking'.

One may argue that the external examination system may have a control effect on some of the factors, such as standard of marking. However, it is an individual member of academic staff who is doing the teaching, preparing the handouts, setting the examination paper, and marking. External examiners cannot sit in practically every lesson or being involved in all the processes of teaching and assessment, thus they cannot know what academic staff have actually done in a particular process.

Furthermore, the main reason why these factors have influenced degree classifications is the pressure from quality assessment itself. When the purpose of quality assessment is focused on the accountability, the factors are more likely to have an impact on degree classifications. So, unless academics accept the performance indicator, and academics themselves want to evaluate their teaching and want to find the problems with their teaching for the purpose of improvement, the influence of these factors on degree results may be inevitable.

Another implication of the findings is that value-added results should be linked to the process of assessment, because all of the factors which have an impact on degree classification are associated with the processes of assessment (see Table 6.3). For example, 'how we assess students' is associated with setting an exam. Therefore value-added results should be linked to the processes of assessment. For example, when value-added scores have increased, we should
examine whether the standard of marking has decreased. To do so, we need to develop some performance indicators which measure specific aspects of the processes of assessment as shown in the table.

Table 6. 3 The relationship between the factors and the assessment

Factors ⇒	**Process of assessment ⇒**	**degree classifications**
how we assessed students	setting exam.	
coaching students for examinations'	Tutorial	
'academic community'; 'attitudes of external examiners'; 'standard of marking'	Marking	
'the way the examination boards behaved'	decision making about criteria for each degree classification	

6.7 Possibility of use of value-added as a performance indicator of teaching

6.7.1 Findings

Although academic staff have a lot of concerns about using value-added as a performance indicator, some members of staff believe that it is important to examine whether a university has added value to students, though value-added results should be used with certain conditions. Some members of staff emphasise that it only could be used along with other performance indicators and it has to be clear what is being measured.

> *Well it's one way, I mean, I don't object to it in principle and to be able to show that one is adding value because impressionistically a lot of what we do is based upon the assumption we are being valued. To be able to confirm that in more rigorously would be important, if we are not, if we can show that we are not then that in term means that there are issues here that need to be investigated and rectified. So I've no objection to the process in principle, it's just safeguards and reservations and warnings about the difficulty of it and could be controlled over the process for those of us who are responsible for these things.*

> *...I would not use it (value-added) as the sole indicator, I would be confident in it being one indicator.*

> *I think it is a good tool, it wouldn't worry me to use it, as long as I have in mind what it is really measuring*

Some members of staff found it acceptable were value-added measurement to be used to identify problems. The value-added result may be treated as a symptom of an 'illness' which needs to be further investigated. Nevertheless, how to interpret the differences in value-added is very difficult. The interpretation of value-added results may lead to the question of blame and responsibility, which is a very sensitive issue.

> *'Yes, I think it is acceptable... I think it is about identifying problems and*

important issue then becomes try to understand and interpret and respond to that problem,... Yes, it is a symptom you see. It is something that needs to be followed throughout a bit further.'

I think as a start as indicator to draw your attention to changes, fine, it's then I think the issue of how you actually interpret the changes but it's not easy, it's not an easy sense...

Yes, if it (value-added results) showed a major deterioration against, the trend over a number of years, it would clearly be alarming, it's about, it's how you interpret it and the problem is if interpretation leads to question of responsibility and in terms of problems of blame, you know, the conclusions that are drawn from this and people are increasingly sensitive, right, because of the political climate onto these kinds of crude, and it's the crude interpretations of performance indicators.

6.7.2 Discussion

The above findings suggest that some members of academic staff believed that it is important to measure value-added and the value-added measurement may be used to identify the problems. Some members of academic staff found it acceptable if the value-added results are treated as a symptom which needs to be further investigated.

Therefore it is clear now that although introduction of value-added as a performance indicator is highly controversial, there is one possibility of using value-added as a performance indicator, which can be accepted by the academics: using value-added results to identify problems.

The implication is that within an institution, the quality assurance office of a university or head of a department could use value-added results to raise questions: why did your school or your course this year appear to have less value-added than others? Are there any problems with the quality of teaching in your schools or courses? Next other methods of evaluating quality of teaching

may be used to directly assess processes of teaching, assessment, and learning. For example, observing classroom teaching and learning, assessing handout, course work or exam paper, and investigating problems with students' learning could be used. Using value-added results in this way could help schools or courses to locate problems with their teaching and to seek improvements.

6.8 Conclusion

This chapter has investigated whether the value-added measurement developed in the quantitative study can be used as a performance indicator of teaching. The technique of semi-structured interviewing is used to investigate academic staff views on this issue. The problems, in the eyes of academic staff, of using the value-added measurement as a performance indicator of teaching have been revealed.

Firstly, academic staff are concerned that the concept of value-added based on entry qualifications and degree results was narrow. They argued that degree results could not reflect all sources of value-added to their students and some values might not be easily quantified. They suggested such things as skills and the value-added through placement should be measured.

Secondly, academic staff revealed that there were problems with the comparability of entry qualifications and the comparability of degrees even at course or school level. They argued that BTEC students could take different routes to study on a degree course, there were huge differences within the access student group in terms of their starting points, and some students took two resits to gain an upper second degree, while other students achieve the same degree class without taking any resits.

Thirdly, the academic staff argued that there was not a one to one relationship between the quality of teaching and what students achieve, because there were other factors that come into it. The factors suggested by the academic staff may be summarized into the following categories: motivation, the aura of the class,

some attainments which are not measured by entry qualifications, domestic situations of the students, increasing amount of time of working in term time, social class, and management of teaching.

Finally, the academic staff also argued that degree classifications did not necessarily indicate how good the students were or how good teaching was. There are factors that interact to influence degree classifications. The factors suggested by academic staff are summarized into following categories: 'academic community', 'the way the examination boards behaved', 'attitudes of external examiners', 'how we assessed students', 'standard of marking', and 'coaching students for examinations'. Therefore when a figure / score shows an improvement of degree results, it could be a result of an improvement in teaching or could be a combination of an improvement in teaching with these factors.

It should be emphasised that the above factors are different from those factors which have an impact on student academic achievement, and which are associated with learning. These factors are ones which have an impact on how accurately degree classifications reflect students' true achievements. They are associated with teaching and assessment.

On the basis of the above findings, it is concluded that the value-added measurement may be used to draw attention to the quality of teaching of those courses or departments which appear to have less value-added, but it cannot directly indicate quality of teaching. Therefore the value-added measurement may be used to identify problems of teaching, but can not be directly used to produce any league table for resource allocation.

The main implications of the findings of the qualitative study for the development of a value-added performance indicator of teaching are as follows:

Firstly, it may be difficult to develop a comprehensive concept of value-added (e.g. the concept of value-added set out by Cave et al 1997); nevertheless, the

concept of value-added need not be as narrow as it is. Skills which are currently emphasised in higher education in the UK and students' performance in placement are two important areas to which the value-added approach should give more attention. Furthermore, it is also required to define what value-added higher education intends for students. What are the objectives of higher education in terms of value ?

Secondly, a value-added performance indicator measures the input and the output of teaching, but it does not take into account the processes which we suppose lies behind that outcome, yet this is central to the phenomena for which it is supposed to be a performance indicator. The findings of this research suggest that value-added results can be influenced by some factors which are associated with the processes of learning and teaching. Therefore, the significance of the value-added results would derive from the link between them and processes of teaching and learning.

Finally, the findings of this research suggest an alternative explanation of the idea that *'value-added scores are negatively correlated with initial scores, therefore a low initial score will generate a high value-added score'* , as discussed in the literature review. We need to distinguish between **progress** in academic attainment and the **level** of academic attainment. Value-added reflects the progress made in academic attainment, but may not reflect the final level of academic achievements. This interpretation may make it easier to understand that a course or an institution which recruits better-qualified entrants may have lower value-added scores.

CHAPTER SEVEN: CONCLUSIONS

This final chapter includes: in 7.1, a brief restatement of the research problem, a description of the main features of the method, and a summary of the main findings; in 7.3, a discussion of the main implications of the findings; in 7.4, a discussion of the study's limitations; and in 7.5, a personal view on the value of value-added.

7.1 Conclusions

This thesis aims to explore the feasibility of constructing and using a value-added performance indicator of teaching in higher education in the UK.

The thesis contains two main parts. The first part focuses on the issue of measuring value-added. It has developed a method of calculating value-added which is based on entry qualifications and degree results. This method is then used to measure value-added at course, school and institutional level with actual data in order to examine the feasibility of the method in actual use. The second part of the study focuses on evaluating whether the value-added measurement can be used to indicate the quality of teaching. It has used the value-added results obtained from the first part of the study to interview academic staff, who have taught on the courses or in the schools measured, in order to assess acceptability and perceived legitimacy of the value-added method developed as a performance indicator of teaching. The main feature of this research is that unlike most research on performance indicators in higher education which are quantitative and carried out at highly aggregated level (e.g. between institutions), this research has been conducted at the lower levels of aggregation (within an institution), and has also employed a qualitative research method. It is a very specific study.

There is evidence from the test of the proposed method against actual data that the method can be used to identify variations in apparent value-added at course (courses with large number of students) school and institutional level (within an institution by comparing different cohorts).

The comparison between the comparative value-added method (the CVA), and the method developed in this research found that in general, although the two methods are different, the value-added results generated from the two methods are similar when the proposed methods uses degree scores to measure quality of degree. However, the proposed method has the following advantages to main existing methods.

Firstly, it is unlike Index methods, which combine scores measuring two different things, entry qualifications and degree results. The method developed in this research compares like with like. It links entry qualifications to degree results by comparing **differences** (or positions) in academic attainments between departments or cohorts at entry, with **differences** (or positions) in academic attainments between them at exit.

Secondly, this research has identified a pattern in the entry qualifications, thus the measurement of value-added is focused on major groups among students with a variety of entry qualifications. By doing so, the proposed method avoids the difficulties with attributing scores to diverse entry qualifications, and measures value-added by comparing like with like (the comparisons of input and output are made within the same type of entry qualifications). Hence it retains the strength claimed by the comparative method that the CVA creates a level playing field for assessing educational value-added. By focusing on major groups of entry qualifications, the proposed method also overcomes the difficulty of the CVA that the number of students with some entry qualifications is so small (some of them only have two or three students) that the value-added results based on such a small number of students may ring false alarms.

However, it should be noted that even so, the method developed in this research only applies to courses with large numbers of students. One way to solve this problem in future research may be classify entry qualifications into two major categories: A level and non-A level, though this introduces difficulties of its own.

Thirdly, the method developed has advantages over the comparative method in terms of both calculation and interpretation. With the comparative method, one has to accept that one student moving five classes of degree is the same as five students moving one class of degree and there are some confusions in interpretation of the CVA scores. The calculation and interpretation of value-added scores in this research are straightforward and avoid those assumptions and confusions of the CVA.

Nevertheless, the proposed method still is a crude measure of value-added. Strictly speaking, value-added should be measured based on the results of a standardized pre-test (at start of a study period) and post-test (at end of a study period) if such tests can be made feasible. The proposed method uses ranking to measure value-added. It should be emphasized that ranking could hide the information although it makes the measurement of value-added very easy.

It may be worth noting an interesting finding. The proposed method uses both percentage of good degrees, and degree scores to measure degree results. It is found that on the whole, the variations in percentages of good degrees between the different groups were bigger than variations in degree scores between them, and using the two methods to measure the same degree results can have different results. The reason for this may be associated with distributions of degree classifications.

The evidence from the interviews with academic staff suggested that the academics found it acceptable for the value-added measurement developed to be used to identify problems, and for the value-added results to be treated as a symptom for further investigation. However, it would not be acceptable to them to use the value-added results to indicate the quality of teaching for the purpose of accountability. The difficulties of doing so from the points of views of the academics are as the follows:

Firstly, the academic staff argued that degree results could not reflect all the values they added to their students, so some benefits of university are not measured. They added other values to their students such as skills, employability and the effects of placement - which are highly emphasized by current higher education policies, but are not measured by degree result. They were also concerned that there are some values which cannot be classified into an upper second, lower second or other degree class.

Secondly, the diversity and flexibility which exists under the current higher education system makes it seems impossible to 'compare like with like' even at course level. The academic staff argued that even students who have the same entry qualifications and study on the same degree course should not be treated the same when we measure value-added, because they may take different routes in their university courses. For example, students with BTEC entry qualifications could take two different routes to study on the same engineering degree course: some BTEC students start year one on the BTEC HND course, and at the end of year one, transfer into the second year of the degree course, while other BTEC students would complete their second year of the BTEC HND course and then would be joint the second year of degree course. There are also diversities within access student groups and the A level student group. The policy of allowing a 'resit exam' could have the result that the same degree class might not necessarily indicate the same academic achievement.

Thirdly, the academic staff argued that there was not a one to one relationship between the quality of teaching and what students achieve, because there were other factors that come into it. The factors suggested by the academic staff may be summarized into the following categories: motivation, the aura of the class, some attainments which are not measured by entry qualifications, domestic situations of the students, increasing amount of time of working in paid employment in term time, social class, and the management of teaching.

Fourthly, the academic staff suggested that degree classifications did not necessarily indicate how good the students were or how good the teaching

was. There are factors that interact to influence degree classifications. The factors they suggested are summarized into following categories: 'academic community', 'the way the examination boards behave', 'attitudes of external examiners', 'how we assess students', 'standard of marking', and 'coaching students for examinations'. Therefore when a figure / score shows an improvement in degree results, it could be a result of an improvement in teaching or a combination of the improvement of teaching with these other factors.

7.2 Implications

The development of a value-added performance indicator of teaching is in its infancy. This research has suggested a method to measure value-added. The method has some advantages over the existing value-added measurements, and can be used to identify the problems of teaching, but it cannot be directly used to indicate quality of teaching. It should be emphasized that those difficulties with using the proposed value-added measurement to indicate quality of teaching, suggested by the academic staff, apply not only to the particular value-added measurement developed in this research, but also to value-added as a performance indicator as a whole. In fact, some of them apply to all performance indicators that are based on degree results such as progression rate.

The main implications arising from the findings of this research for the development of a value-added performance indicator in higher education are as follows:

7.2.1 Defining objectives of value-added

The value-added approach is intended to measure the value-added of students who study in higher education. Nevertheless it seems not to be clear what kind of value higher education intends to add to students. The academics in this research argue that degree results can not represent all values they added to students and they added other values such as skills, some value-added through placement, personal maturity, employability, and life experience... Thus

the questions are raised: what kind of value do degree results represent? What are 'other value' higher education added to students? Are all these 'other values' suggested by the academics, higher education or an institution aims to added to students, in short, what are the objectives of higher education in terms of value?

Pollitt (1990) noted:

> *Strictly speaking, performance indicators can only acquire their full meaning and significance when they are deployed within a context where the overall objectives of the system are defined and known*

Pollitt pointed out that in the UK, setting up indicators has run well ahead of the specification of objectives and emphasized that the objectives of the higher education system and objectives of individual institutions need to be well defined. This research reinforces this view, and further suggests that development of a value-added performance indicator also needs well defined the objectives first. A value-added performance indicator should measure the achievement against the objectives in terms of value-added. The well defined objectives of value-added would help to decide what should be measured, to prioritise those aspects that need to be measured and decide relative weight that should be applied to those aspects measured.

In the literature, the discussions about the concept of value-added have concentrated on what the value of higher education is. This is different from what value higher education aims to added to student, because value of higher education can be broad, but not all values of higher education are necessarily to become the objectives of higher education. For example, the academics in this study believed that students can gain a lot from their courses in terms of value such as knowledge, skills, employability, life experience, and personal maturity. Among them, the first three values, 'knowledge, skills, and employability', may be more important than others, as far as objectives of a course are concerned.

Furthermore discussions about the value of higher education in the literature are abstract and needs to be more specific to become an objective. For example, Cave et al (1997) note the value of higher education included benefits to the individual and society. The benefits to the individuals are increased earning potential and personal development, but what is the objective of personal development of higher education?

7.2.2 Linking value-added results to the process of teaching and learning.

The difficulties about using value-added results to indicate the quality of teaching suggested by the academics indicate that value-added results can not be directly used to indicate quality of teaching. Because a value-added measurement only takes into account input (entry qualifications) and output (degree results), we do not know what process lie behind the value-added results) and nor do we know the link between the value-added results and process. While during the process of input transferring output there are many factors could have an impact on value-added results.

Nevertheless, value-added results could derive significance by linking to the process of teaching and learning, because the further analysis of those factors that may have an impact on value-added results suggests that all those factors can be linked to the processes of learning and teaching. Those factors that may have impact on what students achieve are linked to process of learning (see Table 6.2 in Chapter 6). For example, the factor of 'the aura of the class' is linked to motivation of learning. while the factors, 'domestic situations of the students'; 'amount of time of working in term time' are associated with the time students spent studying.

Therefore, value-added results should be linked to those aspects of learning: motivation, academic abilities, study skills, studies time, and study habits. So, for example, when we see an improvement in value-added results this year compared with last year, we should ask: did this year 's students work harder

than last year? Did this year's students take less part-time work than last year's and spend more time doing their course work?

Similarly, those factors, that have impact on the extent to which degree classifications reflect students' true academic achievements, can be linked to different aspects of the assessment process (see Table 6.3 in Chapter 6). The factor, 'how we assessed students' is associated with setting examinations, while the factors, 'academic community', 'attitudes of external examiner' and 'standards of marking', all have something to do with marking.

Therefore value-added results should be considered in the context of assessment. When we see an improvement in value-added, we should first examine the assessment of the achievement-- is this year's examination easier than last year's? Have the staff marked leniently this year?

Linking the value-added results to the process of learning and teaching could help to improve value-added results, because by doing so, we can identify possible reasons for the variations in the value-added results, and then suggest how to improve the value-added. As Banta and Borden (1994) noted, to improve outcome, performance indicators need to link to process, because measuring an outcome will not, in and of itself, result in improvement. We need to examine carefully the processes that lead to outcomes if we hope to improve them.

However, it should be emphasized that linking value-added results to the process of teaching and learning does not solve all the problems with using value-added to indicate the quality of teaching suggested by the academic staff in this research. It is still difficult to use the value-added results for the purpose of accountability, because if a value-added indicator is to be used to indicate quality of teaching for the purpose of accountability, it should be able accurately to take all the factors into account. A value-added performance indicator based on entry qualifications and degree results is still a crude measure of value-added. The diversity and flexibility which exists under the current higher

education system makes it difficult to compare 'like with like' not only at a high level (sector level or institutional level), but also at low level (department and course level).

7.3 Limitations

Although the method of calculating value-added developed has some advantages over the main existing methods, it still is a crude measurement (The methodology chapter has discussed why such a crude measurement is developed). This method has to be used with caution. The value-added results can only be treated as a symptom in evaluating quality of teaching. The value-added results should be linked to the process of teaching and learning.

Although sophisticated statistical techniques may not be suitable to be used directly to construct performance indicators, it is important to carry out some research in this area if the data are available.

Because of availability of the data, the method developed in this research has only tested with data from two new universities. Some practical problems may appear when the method is used with other universities, particularly old university data since old universities have a different recruiting profile from new universities. The method may need to be refined when some questions arise. On the other hand, with present data, some differences in entry qualification and degree results are found not statistically significant, this may be due to the number of students being small. The insignificant results could change when the method is applied to a different data set.

The sensitivity of the research topic has affected availability of the data for the present research. The attitudes of universities or academics towards performance indicators have affected my access to the data on student entry qualifications and degree results. Universities are not willing to release the data that are going to be used to measure their performance even for the purpose of research. This has limited me to obtaining the data directly from the individual institutions and also obtaining their permissions to use their data collected by

USR (University Statistics Record), FESR (Further Education Statistics Record), and HESA (Higher Education Statistics Agency).

The availability of the data for the present research has also been affected by the status of the research project. While some research project on performance indicators could obtain the data from USR, FESR and HESA and even publish research results with the actual name of institutions, the present research project has to obtain individual institution permission to use HESA data, even though the present research has planned to present results anonymously. This is because the present research project is not an official project. The research project was not proposed by any official organizations or bodies such as Higher Education Funding Council. The research topic is chosen by an individual institution. In the first report of Performance Indicators Steering Group (1999 p35), it is noted that:

> *The HEFCE therefore intends to make the longitudinal database it has constructed available to academic researchers (Such arrangements would involve the restrictions, security arrangements, and so on, that currently apply to consultants carrying out work for the HEFC)*

Therefore, it appears that research on this area has to be 'official' research, otherwise, one would have difficulties in accessing to the data to conduct research with a large data set. This means that individual institutions have been excluded from conducting research on this area, but is this right?

Ideally, the method should be tested by comparing value-added results produced by it with results generated from other teaching and learning assessments. In the present study, efforts have been made to enable comparisons between the value-added results of the measurement developed and results of the University Annual Teaching Evaluation which is operated by the institution studied. However, this is found not feasible because of the availability and quality of the data on the University Annual Teaching Evaluation. It is found that either the data are not available, or the period or

levels assessed by the University Annual Teaching Evaluation are not matched with that of measurements made in this study. One has to appreciate that the University Annual Teaching Evaluation and the value-added measurement are not designed to enable comparisons to be made between them and they are designed for different purposes.

This study has suggested a number of factors that may have impact on the value-added. The effects of these factors on value-added need to be tested through a quantitative study. Nevertheless, collecting data on these factors is going to be a big issue.

The qualitative evidence in this research is based on interviews with 18 members of academic staff. These members of staff are from three subject areas: Business Studies, Health and Community Studies, and Engineering of a new university. (The methodology chapter has explained why these participants were chosen). Therefore some of the findings of the present research need to be generalized cautiously to other subjects and other institutions particularly old universities. For example, some of the factors that may have an impact on student academic achievements found in this research are associated with a particular subject and particular institution.

7.4 Value of value-added

Politicians tend to place undue reliance on performance indicators. Academics in general dislike performance indicators and see them as a threat to the autonomy previously enjoyed.

In my view, on the one hand, value-added and performance indicators as a whole are not as powerful as politicians think. It may be very difficult to develop a value-added performance indicator that is feasible in practice and can be used d*irectly* to allocate resource or used *directly* to indicate quality of teaching. In short, it may be difficult to develop a value-added performance indicator for the purpose of accountability. However, on the other hand, it is possible to develop a value-added performance indicator can be used to identify problems

of teaching for the purpose of improvement. A value-added performance indicator can derive its significance by linking with the process of teaching and learning. It needs to be used with other measures of quality of teaching. The value-added performance indicator is just one of many tools that are aimed to evaluate and improve the quality of teaching and the outcome of teaching.

REFERENCES

Aitken, A. (1991) 'The subtle link between teaching and research.' Symposium on the Quality of Teaching in Higher Education Institutions, Canberra

Astin, A. (1965) Who goes where to college? In Astin, A (Ed.) Who Goes Where to College? Chicago: Science Research Associates.

Astin, A. (1982a) 'Why not try some new ways of measuring quality?' Education Record 63, spring, 10-15

Astin, A (1982b) Minorities in American higher education: Recent trends, Current Prospects, and Recommendations. San Francisco; Jossey-Bass.

Ashworth, P (1997) 'Qualitative research methods' Fifth Improving Student Learning Symposium, Strathclyde University

Baker, K. (1987) Report in The Times Higher Education Supplement. 6 February 1987, 1

Ball, R. And Wilkinson, R. (1994) 'The use and abuse of performance indicators in UK higher education' Higher Education 27, pp417-427

Barnett, R., (1988) 'Entry and exit performance indicators in higher education: some policy and research issues', Assessment & Evaluation in Higher Education, Vol.13, NO.1, Spring 16-30

Baird, L. L., (1987) Value added : using student gains as yards of learning. In C. Adelman (Ed.), Performance and Judgment: Essay on Principles and Practice in the Assessment of College Student learning Washington DC: US Government Printing Office

Bauer, M., (1986), 'A commentary on the Northeast Missouri and Tennessee Evaluation Models' ' International Journal of Institutional Management in Higher Education, Vol.10, NO.3

Bekhradnia, B. (1999) 'Tables for our times' in The Times Higher Education Supplement, 3rd Dec. 1999

Bennett, R. E. (1993) On the meanings of constructed responses. In R. E. Bennett and W. C. Ward (Eds.) Construction versus Choice in Cognitive Measurement Hillsdale, NJ: Lawrence Erbaum

Birch, D., Calvert, J., and Sizer, J., (1977) 'A case study of some performance in higher education in the United Kingdom' International Journal of Institutional Management in Higher Education, Vol.1, No.2, October, 133-142

Borden, V. And Bottrill, K. (1994) 'Performance indicators; history, definitions, and methods.' In V.Borden and T. Banta (eds.) Using Performance Indicators to Guide Strategic Decision Making. New Directions for Institutional research 82, 5-21

Bauer, M., (1986), 'A Commentary on the Northeast Missouri and Tennessee Evaluation Models' ' International Journal of Institutional Management in Higher Education, Vol.10, NO.3

Bourke, P., (1986), Quality Measures in Universities, The Commonwealth Tertiary Education Commission Australia

Bourner, T., and Hamed, M., (1987): Entry Qualifications and Degree Performance. CNNA Development Services Publication 10, London

Boyer, E., (1985), 'Olive branches out with five year tenure', The Times Higher Education Supplement, 9.2.89.p.17

Boys, C. J. and KirKland, J., (1988), Degree of Success: Career Aspirations and Destinations of College, University and Polytechnic Graduates, Jessica Kingsley.

Brennan, J. And McGeevor, P. (1988) Graduates at Work. Degree Courses and the Labour Market. London: Jessica Kingsley Publishers.

Brennan, J . *et al* (1994) 'The experiences and views of graduates: message from recent surveys.' Higher Education Management 6, 3, pp275-304.

Bryman, A., Haslam,C., Webb, A., (1994), 'Performance appraisal in UK universities: a case of procedural compliance?' Assessment & Evaluation in Higher Education, Vol.19, NO.3

Chapman, K. (1996) ' Entry qualifications, degree results and value-added in UK universities' Oxford Review of Education, 22, 3, p 252-264

Cave, M., Hanney, S., Kogan, M. and Trevett, G. (1991) The Use of Performance Indicators in Higher Education, London: Jessica Kingsley:

Cave, M., Hanney, S., Henkel, M., and Kogan, M. (1997) The Use of Performance Indicators in Higher Education, The Challenge of the Quality Movement . London: Jessica Kingsley

Clift, J., Hall, C., and Turner, I., (1989), 'Establishing the Validity of a set of Summative Teaching Performance Scales' Assessment and Evaluation in Higher Education Vol. 14, No.3 Autumn, 193-206.

Cohen, L., and L. Manion, (1989), Research Methods in Education, Routledge.

Committee of Vice-Chancellors and Principals and University Grants Committee (1986) Performance Indicators in Universities: A First Statement by a Joint CVCP/UGC Working Group. London: CVCP

Committee of Vice-Chancellors and Principals and University Grant Committee (1987) A Second Statement by a Joint CVCP/UGC Working Group. London: CVCP

Cuenin, S. (1987) 'The use of performance indicators in universities: An international survey' International Journal of Institutional Management in Higher Education 11, 2

Cutter, P. (1987) The Effectiveness of Schooling Ph.D. thesis, Edinburgh University

CVCP (1987) University Management Statistics and Performance Indicators, UK Universities, London

Dehayes, D. and Lovrinic, J. (1994) 'Activity-based costing model for assessing economic performance.' In Borden, V and Banta, T. (Eds) Using Performance Indicators to Guide Strategic Decision Making. New Directions for Institutional Research 82, 81-93

DES (1985) Green Paper The Development of Higher Education into the 1990s Cmnd. 9524. London: HMSO

DES (1987a) White Paper Higher Education: Meeting the Challenge, Cm. 114. London: HMSO.

DES (1987b) Accounting and Auditing in Higher Education. London : DES.

DES (1991) White Paper Higher Education: A New Framework London HMSO

Egan, J., (1986), Value-added Testing and Measurement: Will It Save Higher Education From Its Critics? Paper Presented at the Value-added Learning Conference, Empire College, State University of New York, 4-6 June

Elton, L., (1987) ' Warning signs' The Times Higher Education Supplement 11 september

Entwistle, N.J. and Wilson, J.D. (1977) Degree of Excellence: The Academic Achievement Game. Hodder and Stoughton, London

Evans, T. And Clift, R. (1987) 'Assessing quality' The Times Higher Education Supplement, 18 September .

Ewell, P. (1991) 'Assessment and public accountability: back to the future.' Change 23, 6,12-17.

Ewell, P. And Jones,D. (1994) 'Pointing the way: indicators as policy tools in higher education.' In S.Rupper (1994) Charting Higher Education Accountability: A Sourcebook on State-Level Performance Indicators. Denver: Education Commission of the States

Filstead, W. J. (1970) Qualitative Methodology. Chigo, IL: Markham

Fincher, A.,(1985)' What is value-added education?' Research in Higher Education, 22, 395-398

Fitz-Gibbon, C.T. (1997)The Value Added National Project Final Report London: School Curriculum and Assessment Authority

Fontana, A. And Frey, J. H. (1994) 'Interviewing: the art of science' in N.K. Denzin and Y.S. Lincoln (eds), Handbook of Qualitative Research. Thousand Oaks, CA: Sage. pp361-76

Frank, R. F. (1984) Are workers paid their marginal product? American Economic Review, 74, pp519-571

Frackmann, E. (1987) 'Lessons to be learnt from a decade of discussions on performance indicators' International Journal of Institutional Management in Higher Education 11,2, 149-162

Frey, J. H. (1993) 'Risk perception associated with a high-level nuclear waste repository', Sociological Spectrum

Galtung, J. (1967) Theory and Methods of Social Research London: Allen and Unwin

Gallagher, A., (1991) ' Comparative value added as a performance indicator' Higher Education Review 23 pp19-29 summer 1991

Goddard, Thomson, and Wojtas, (1999) ' How did you perform?' The Times Higher Education Supplement, 3 December 1999

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

Greenwald, D., and Associates (1983). The McGraw-Hill Dictionary of Modern Economics (Third ed.) McGraw-Hill, New York

Grint, K., (1993) What's wrong with performance appraisal? a critique and suggestion. Human Resource Management Journal, 3, pp61-77

Gulck, S. B. (1991) ' Advocacy oral history: palestinian women in resistance', in S.B. Gluck and D. Patai (eds), Women's Word: The Feminist Practice of Oral History. London: Routledge,

Hadley,T., and Winn, S.,(1992) 'Measuring value added at course level: an exploratory study. Higher Education Review V25 pp7-30 Fall 1992

Harris, M. (1986) 'Judgments of quality in higher education: the role of the University Grants Committee' Paper presented to Anglo-American Seminar on Quality Judgments in Higher Education at Templeton College.

Hampden-Turner, C., (1977) Education, Nihilism and Survival Darton, Longman and Todd, London

Higher Education Management Statistics (HEMS) Group (1995) Higher Education Management Statistics: Publication in 1996. London: CVCP/SCOP/COSHEP.

Higher Education Funding Council for England, (1999) Performance Indicators in Higher Education. Bristol: HEFCE 99/66

Higher Education Quality Council (1997) Assessment in Higher Education and the Role of 'Graduateness' London: HEQC

Ions, E., (1977) Against Behaviouralism: A Critique of Behavioural Science, Basil Blackwell, Oxford

Jarratt Report (1985) Report of the Steering Committee for Efficiency Studies in Universities. London:CVCP.

Jesson, D. And Mayston, D. (1990) 'Information, accountability and educational performance indicators.' In C. T.Fitz-Gibbon (ed) Performance Indicators, British Educational Research Association Dialogues No. 2. Clevedon: Multilingual Matters

Joint Performance Indicators Working Group (1994) Management Statistics and Performance Indicators in Higher Education: Consultative Report. Bristol: HEFCE.

Johnes, J. (1990) 'Determinants of student wastage in higher education', Studies in Higher Education, 15, 1, pp87-99.

Johnes, G. (1995) 'Scale and technical efficiency in the production of economic research.' Applied Economic Letters 2, 7-11

Johnes, J. And Taylor, J. (1987) 'Degree quality: an investigation into differences between UK universities' Higher Education 15,3, 209-225.

Johnes, J. And Taylor, J. (1990) Performance Indicators in Higher Education The Society for Research into Higher Education and Open university Press

Karabel, J., & Astin, A. (1975). 'Social class, academic ability, and college quality'. Social Forces, 53, 381-398.

Kitwood, T.M. (1980) Disclosures to A Stranger: Adolescent Values in An Advanced Industrial Society. Routledge and Kegan Paul, London

Lather, P. (1986) 'Issues of validity in openly ideological research: between rock and a soft place.' Interchange, 17, 4, 63-84

Lather, P. (1993) 'Fertile obsession: validity after post - structuralism' The Sociological Quarterly, 34, 673-683

Laurillard, D.M. (1980) 'Validity of indicators of performance' In D. Billing (ed) Indicators of Performance. Guildford: The Society for Research into Higher Education

Layard, R. and Glaister, S. (Eds) (1994) Cost-benefit Analysis. Cambridge: Cambridge University Press.

Le Grand, J. And Robinson, R. (1979) The Economics of Society Problems. London: Macmillan

Lincoln Y. S. and Guba, 1985, Naturalistic Inquiry, California Sage

Lord, F. M. (1956), The measurement of growth. Educational and Psychological Measurement, 16, pp. 421-437

Maxwell, J.A. (1992) 'Understanding and validity in qualitative research' Harvard Educational Review, 62, 279-300

Metcalfe, L. and Richards, S. (1990) Improving Public Management, 2nd edition London Sage

McClain, C. J., Krueger, D.W. and Taylor,T., (1986), 'Northeast Missouri State University's Value Added Assessment Program: A model for education accountability' Institutional Journal of Institutional Management in Higher Education Vol.10 NO.3

McPherson A and Patterson L (1990) 'Undergraduate non-completion rates: a comment', Higher Education 19 pp377-383

Miller, R. I. (1986) 'A ten year perspective on faculty evaluation.' International Journal of Institutional Management in Higher Education 10, 2, July, 162-168

Moravcsik, M. (1986) ' Assessing the methodology for finding a methodology for assessment.' Social Studies of Science 26, 534-39

Morris Report (1990) Performance Indicators: Report of a Committee of Enquiry Chaired by Mr. Alfred Mossis. London PCFC

Morrison, H. G., Magennis, S. P. and Carey, L. J. (1995) 'Performance indicators and league tables: a call for standards' Higher Education Quarterly, vol.49, No. 92, April 1995

Moses, I., (1989), 'Role of problem of heads of departments in performance appraisal' Assessment and Evaluation in Higher Education, Vol. 14, No. 2, Summer 95-96

Nedwek, B. P., and Neal, J. E. (1994) 'Performance indicators and rational management tools: a comparative assessment of projects in North America and Europe.' Research in Higher Education 35, 1, pp75-103

Nesfield-Cookson, B., William Blake: Prophet of Universal Brotherhood, Crucible

Nevin, E. (1985) 'The finance of university academic departments' Applied Economic, 17, pp761-79.

Page, E. (1987) Letter to The Times Higher Education Supplement, 25 September 1987

Pascarella, E. and Terenzini, P. (1991) How College Affects Students Jossey-Bass Oxford

Performance Indicators Steering Group (1999) Performance Indicators in Higher Education HEFCE Bristol

Peters, M., (1992) 'Performance and Accountability in 'Post-industrial Society': the crisis of British Universities' Studies in Higher Education Vol. 17, No.2,

Pollitt, C., (1987), The politics of performance assessment: lessons for higher education? Studies in Higher Education 12 pp87-98

Polytechnics and Colleges Funding Council and Council for National Academic Awards, (1990), The Measurement of Value Added in Higher Education.

Pratt, J. (1982) 'Resource allocation within the public sector' in A. Morris and J. Sizer (Eds.) Resources and Higher Education SRHE, Surrey.

Punch, K. (1998) Introduction to Social Research, Quantitative and Qualitative Approaches SAGE Publications London

Raudenbush, S.W. and Willms, J.D. (1991) Schools, Classrooms, and Pupils , International Studies of Schooling from a Multilevel Perspective . Academic Press, London and New York

Reynolds, D. (1981) 'The naturalistic method of educational and social research' Interchange, 11, 77-89

Richardson, R. (1994) 'Effectiveness in undergraduate education: an analysis of state education indicators' In S. Ruppert (ed.) Charting Higher Education Accountability: A Sourcebook on State-Level Performance Indicators. Denver: Education Commission of the States

Rimson, J. (1991) Activity Accounting: An Activity Based Costing Approach. John Wiley and Sons Inc., New York

Rogers, A. And Scratcherd, T. (1986) 'Dissecting the UGC rankings' Letter to The Times Higher Education Supplement, 7 November 1986.

Sear, K. (1983) The correlation between A level grade and degree results in England and Wales, Higher Education 12 pp609-619

Schwartz, H. and Jacobs, J. (1979) Qualitative Sociology: A Method to the Madness New York: The Free Press

Silver, H., Stennett, A., and Williams, R. (1995) The External Examiner System: Possible Futures, Higher Education Quality Council, London

Sizer, J. (1979) 'Assessing institutional performance : an overview.' International Journal of Institutional Management in Higher Education 3, 1, 49-77

Sizer, J. (1981) 'Performance assessment in institutions of higher education under conditions of financial stringency, contraction and changing needs: a management accounting perspective' Accounting and Business Research 11, 43, 227-242.

Sizer, J. (1988), 'In Search of excellence -performance assessment in the UK', Higher Education Quarterly, 42: 152-61

Sizer, J., (1989) Performance Indicators and Quality Control in Higher Education, Keynote Address to an International Conference Institution of Education, London

Sizer, J. (1989) 'Performance Indicators and Control quality in Higher Education.' Keynote Address to an International Conference, Institute of Education, London.

Sizer, J. (1990) 'Performance Indicators and management of universities in the UK: a summary of developments with commentary.' In F. Dochy, M. Segers and W. Wijnen (eds.) Management Information and Performance Indicators in Higher Education : An International Issue. Assen/Maastricht: Van Gorcum

Sizer, J. (1992) 'Performance indicators in government institutional relationships: lessons from government' Higher Education Management 4, 156-163

Sizer, J., Spee, A. And Bormans, R. (1992) 'The role of performance indicators in higher education' Higher Education 24, 133-155.

Straw,C., and Kaye, M., (1995), 'Measuring Comparative Value Added: a Modified Methodology Higher Education Review Vol. 27, No.2 Spring

Stake, R. (1977) 'Some alternative presumptions' Education News, 18-19

Taylor, J. (1995) 'A statistical analysis of the 1992 research assessment exercise.' Journal of Royal Statistical Society, A, 158, 241-261

Taylor, T., (1985), 'A Value Added Student Assessment Model: Northeast Missouri State University', Assessment and Evaluation in Higher Education, Vol. 10, NO.3, 190-202

Taylor, W. (1987) University Under Scrutiny, OECD, Paris

Thorndike, E.I. (1924), The influence of chance imperfection of measures upon the relationship of initial score to gain or loss. Journal of Experimental Psychology, 7, pp. 225-232.

The Times Higher Education Supplement (THES) (1987) 20th February.

Tomkins, C and Green, R. (1988) 'An experiment in the use of data envelopment analysis for evaluating the efficiency of UK university departments of accounting.' Financial Accountability and Management 4, 2, 147-164

Trower, P., and Vincent, L. (1996) Value Added National Project: Experimental Study of Different kinds of Data Presentation London: School Curriculum and Assessment Authority

Tysome, T. (1998) 'Job skills may be added on' , The Times, Higher Education Supplement, March 1998.

UGC (1982) Annual Survey, Academic Year 1980/81, Cmnd. 8663, London, HMSO.

UGC (1984) A Strategy for Higher Education into the 1990s, London, HMSO.

University Grants Committee (1985), Circular Letter 22/85 Planning for the Late 1980s: The Resource Allocation Process.

Weale, M. (1992) 'The benefits of higher education' Oxford Review of Economic Policy 8, No. 2, pp 35-47

Wolf, F. A. (1981) Taking the Quantum Leap . San Francisco: Harper and Row

Woodley, A. (1984) The older the better? A study of mature students Research in Education, no 32, 35-50

Yorke, M. (1991) Performance Indicators: Observations on Their Use in the Assurance of Course Quality CNAA project Report 30. London: Council for National Academic Awards.

Yorke, M. (1996) Indicators of Programme Quality London: Higher Education Quality Council

APPENDICES

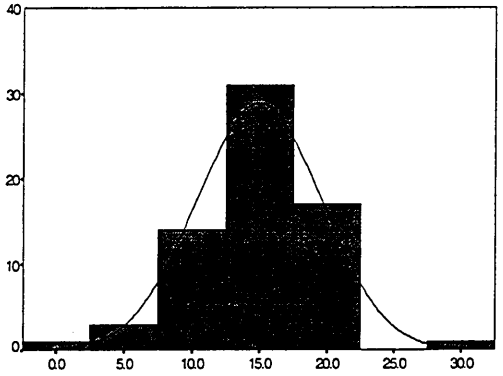
Appendix 1-3 present the histograms which show the distributions of A level points and degree scores. The purpose of examining the distributions is to choose appropriate statistical tests to test the significance of the differences in mean A level points between different cohorts and differences in mean degree scores between them. Because the mean A level point scores, and degree scores are represented by interval variables, and the samples of the data (the four cohorts) are independent, the analysis of variance (ANOVA) is a potential test that can be used to test the differences. Nevertheless, the ANOVA test is based on the assumption that the population of scores is normally distributed. Therefore we need to test whether the distribution of A level points and the distribution of degree scores are normal.

A histogram is a clear and informative chart, providing us with a way of showing how A level point scores and degree scores are distributed across the range of A level points and the range of degree scores. The histograms below show a number of cases of A level points distributions, and the corresponding normal curves with the same means and standard deviations. Therefore using the histograms, we can see how the A level points and degree classifications of each cohort or school are distributed and whether the distributions are normal, by comparing with the normal curves.

Appendix 1 The distributions of A level points of the courses

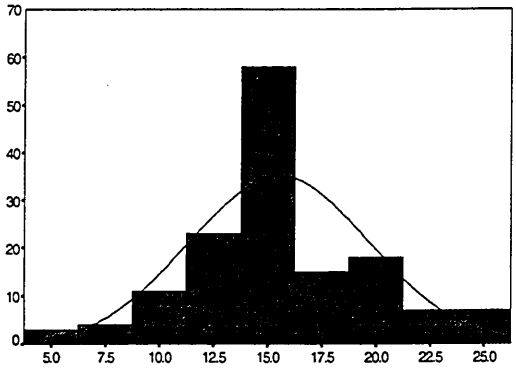
Fig. 1.1 Distribution of A level points of the five cohorts of the Business Studies Course

1988 cohort



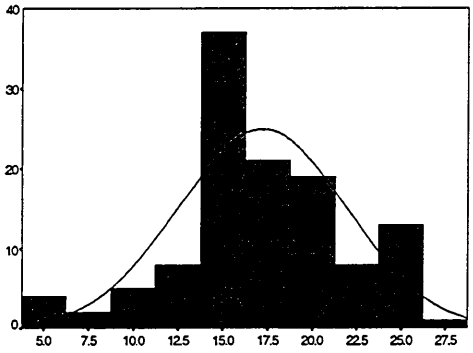
A level points

1989 cohort



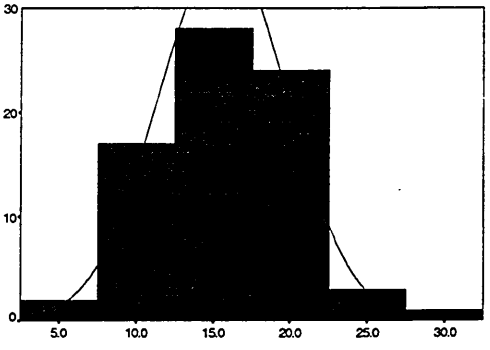
A level points

1990 cohort



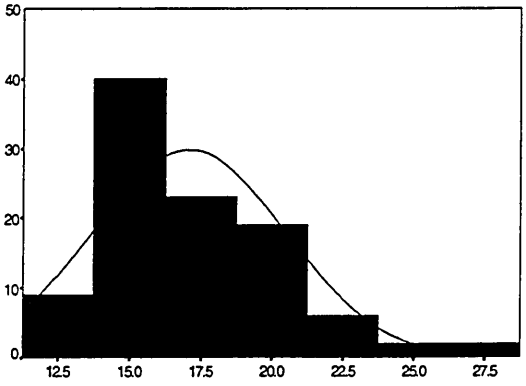
A level points

1991 cohort



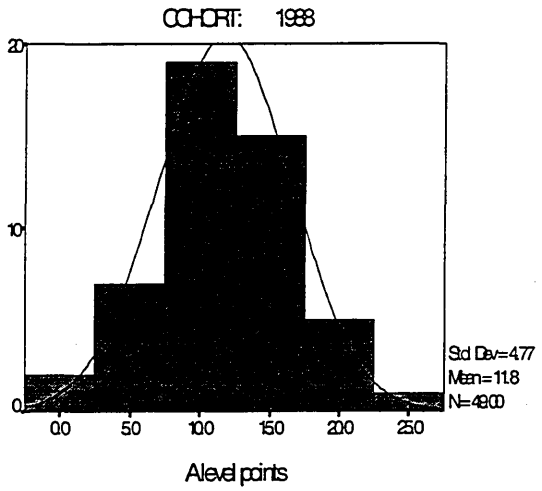
A level points

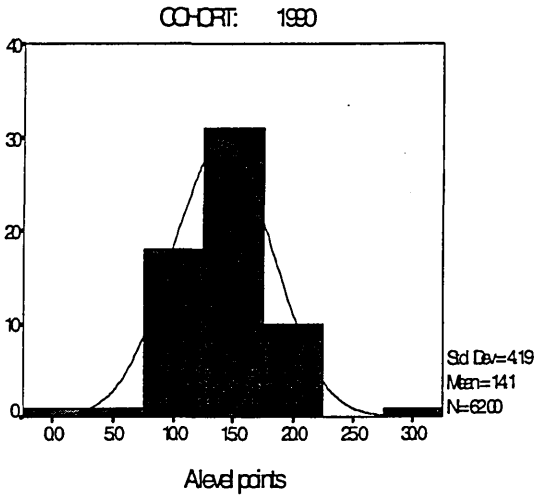
1992 cohort



A level points

Fig. 1.2 Distribution of A level points of the three cohorts (88,90,91) of the Applied Social Studies Course





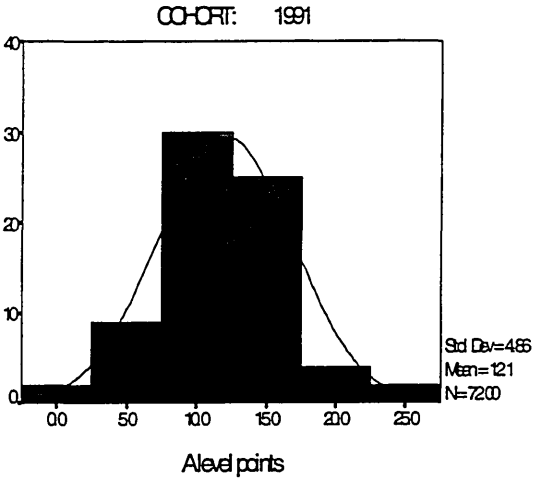
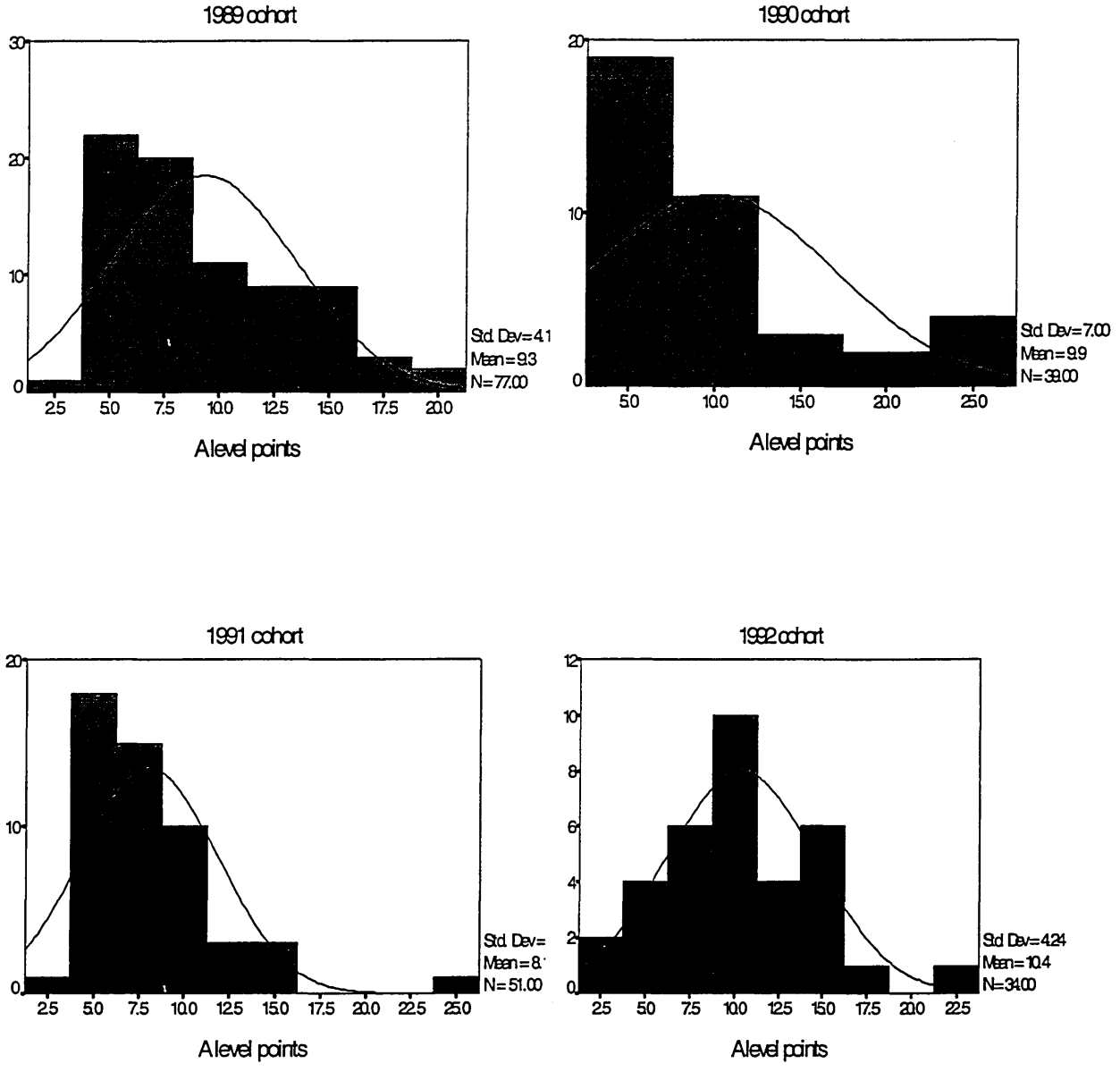


Fig. 1.3: Distribution of A level points of the four cohorts (89-92) of the Engineering Studies Course



Appendix 2 The distributions of A level points and degree results of the schools

Fig. 2.1 Distribution of A level points of the six schools

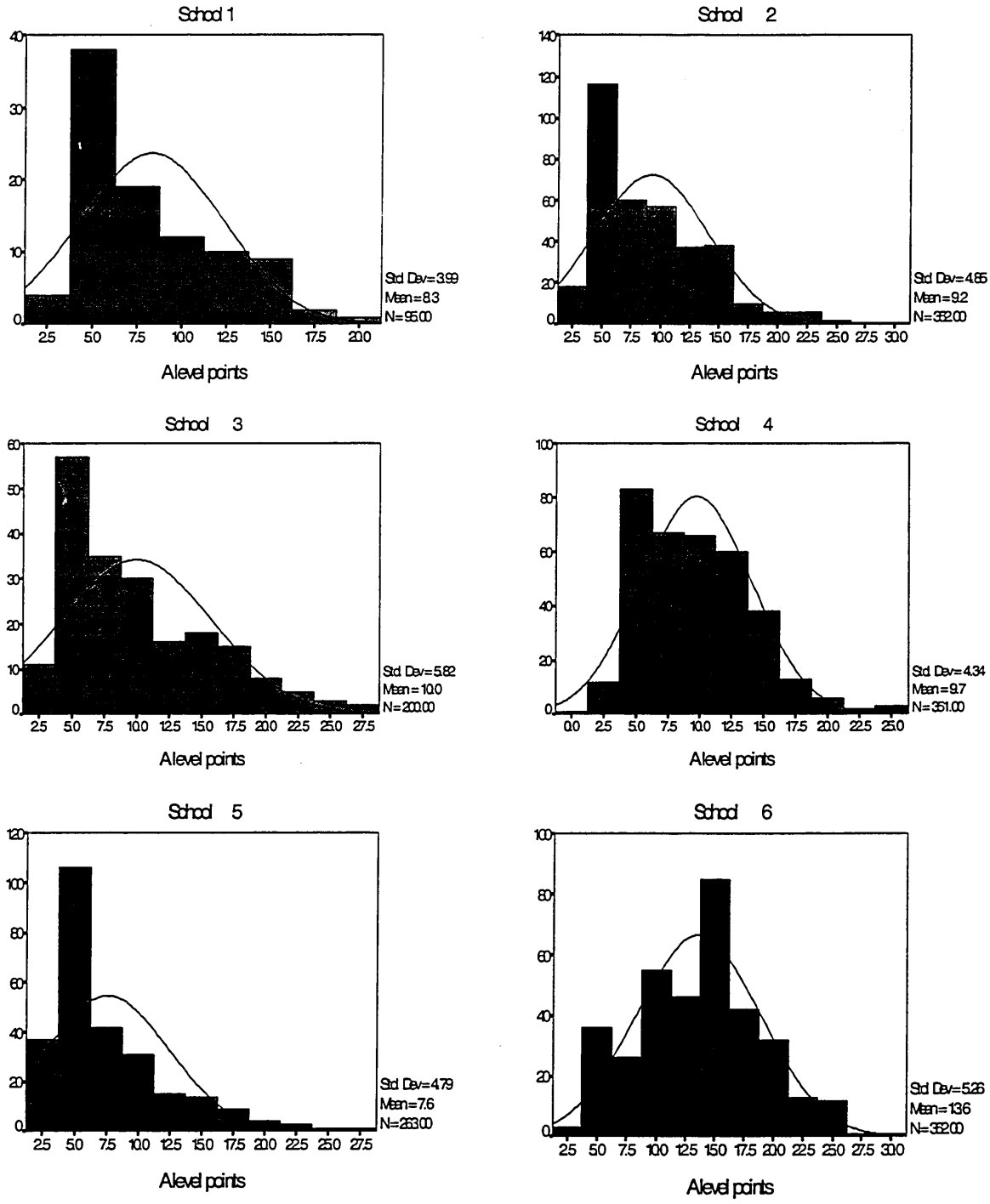
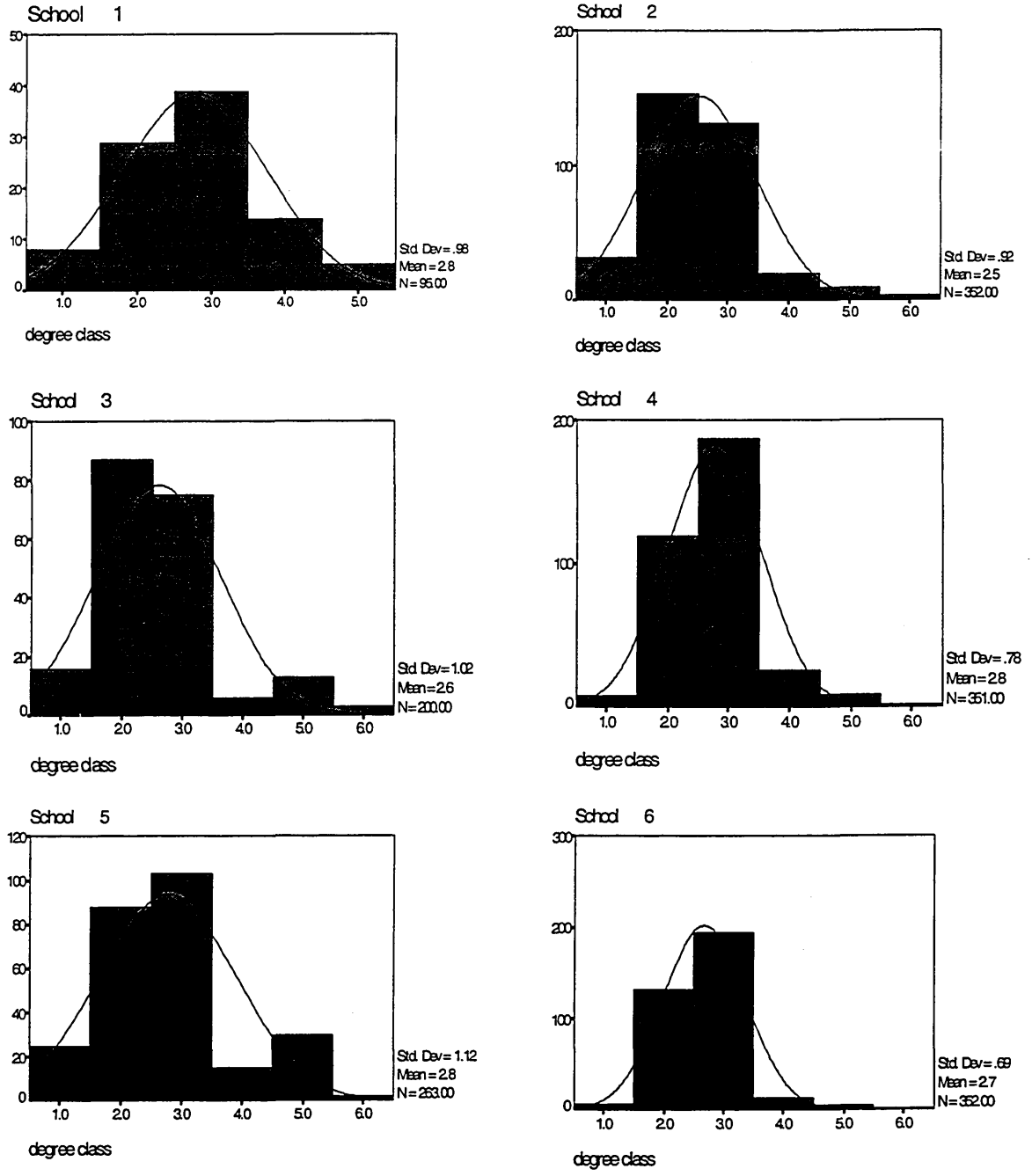


Fig. 2.2 Distribution of degree classifications of the six schools

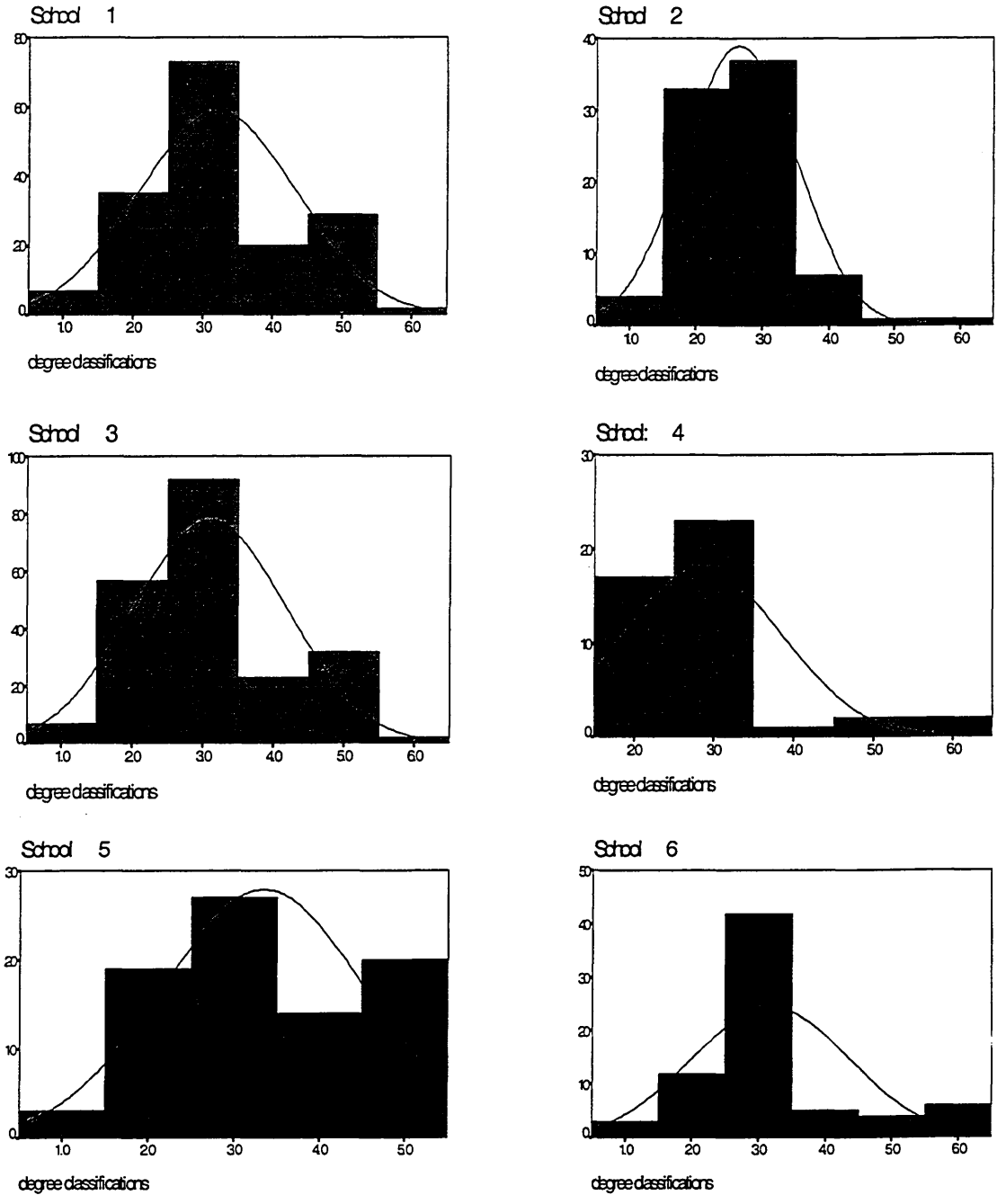
A level group



Note 1 =1st, scored 72; 2 = 2.1, scored 66; 3 = 2.2, scored 54; 4 = 3rd, scored 44; 5= pass, scored 35; 6= fail, scored 25

Fig. 2.3 Distributions of degree classifications of the six schools

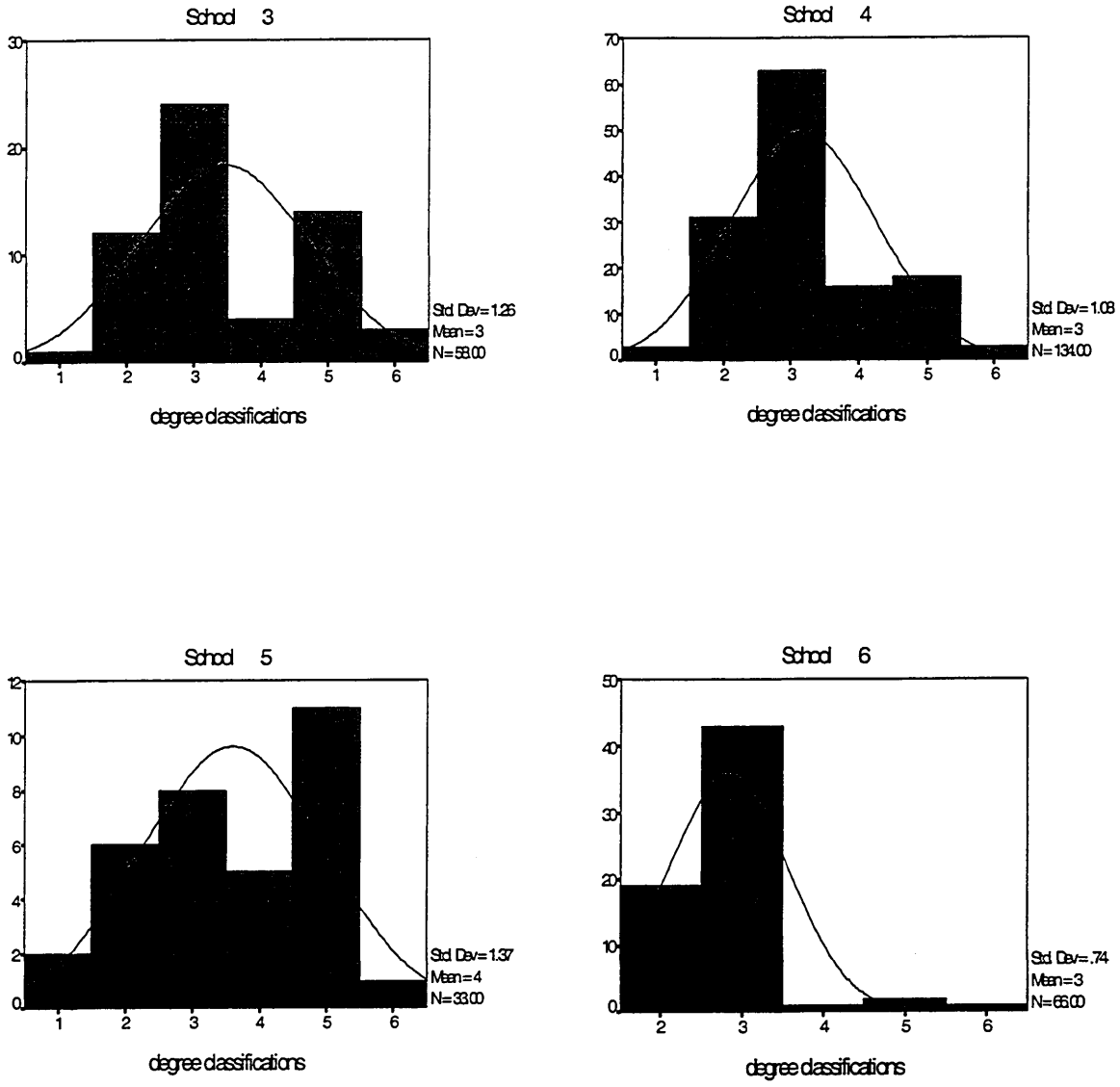
BTEC group



Note 1 = 1st, scored 72; 2 = 2.1, scored 66; 3 = 2.2, scored 54; 4 = 3rd, scored 44; 5= pass, scored 35; 6= fail, scored 25

Fig. 2.4 Distributions of degree classifications of the six schools

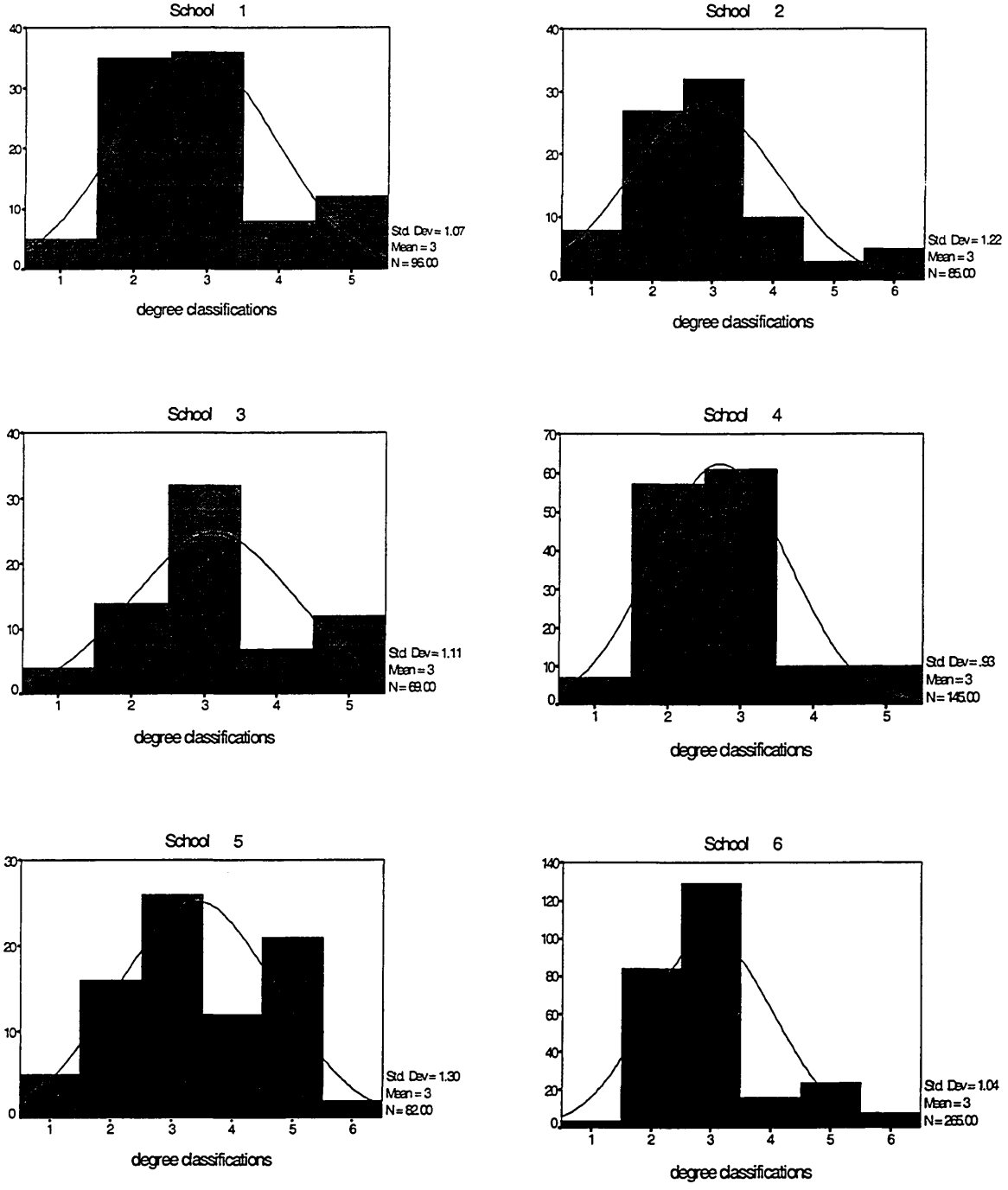
Access group



Note 1 = 1st, scored 72; 2 = 2.1, scored 66; 3 = 2.2, scored 54; 4 = 3rd, scored 44; 5= pass, scored 35; 6= fail, scored 25

Fig. 2.5 Distributions of degree classifications of the six schools

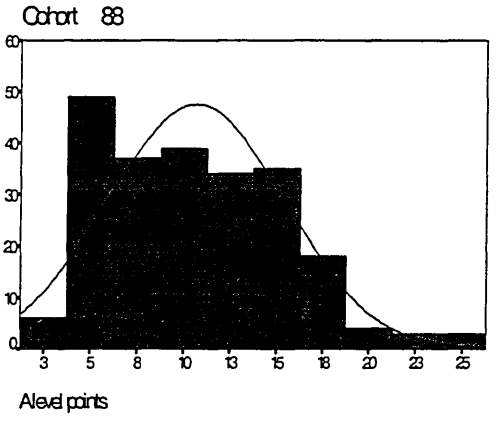
'Other qualifications' group

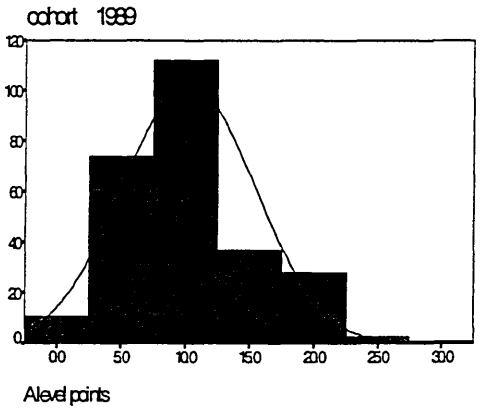


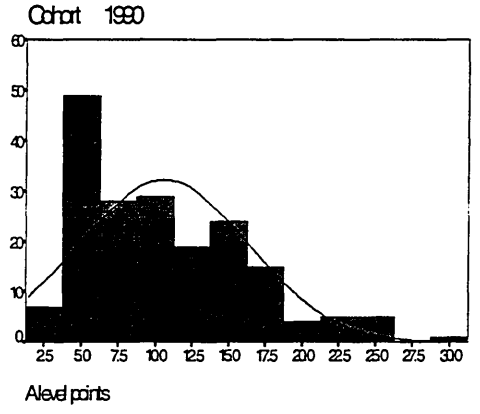
Note 1 = 1st, scored 72; 2 = 2.1, scored 66; 3 = 2.2, scored 54; 4 = 3rd, scored 44; 5= pass, scored 35; 6= fail, scored 25

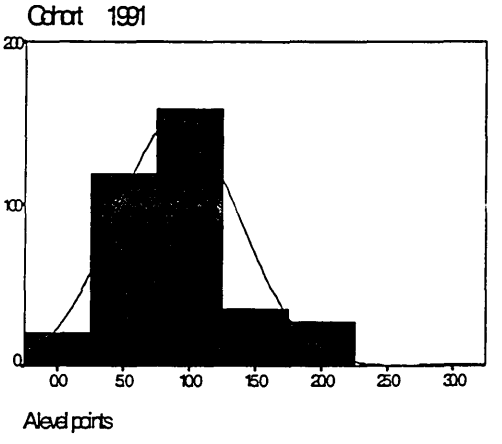
Appendix 3 The distributions of A level points and degree results of the cohorts

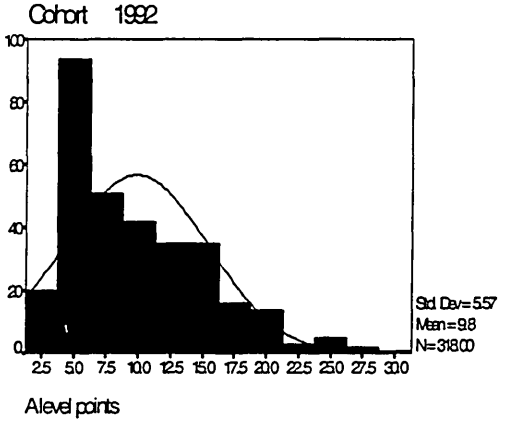
Fig. 3.1 Distribution of A level points of the six cohorts











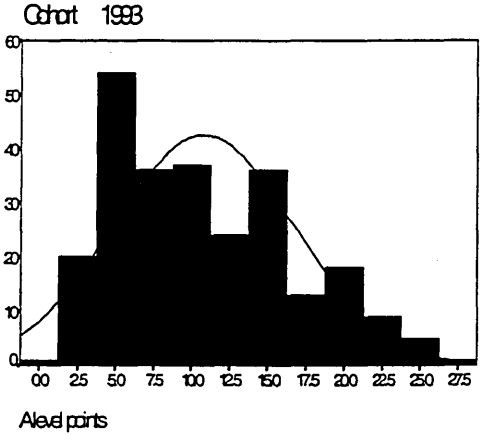
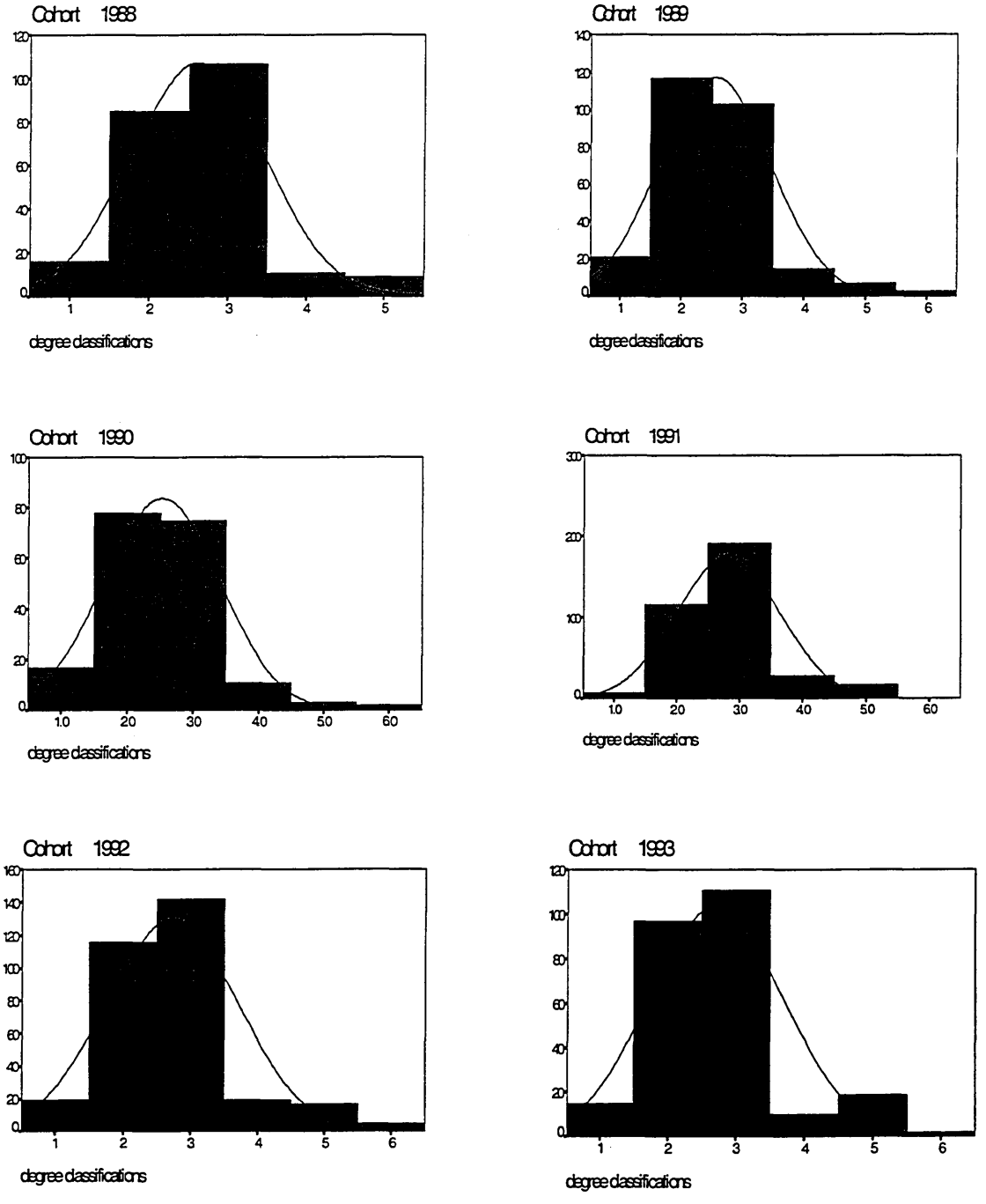


Fig. 3.2 Distributions of degree classifications of the six cohorts

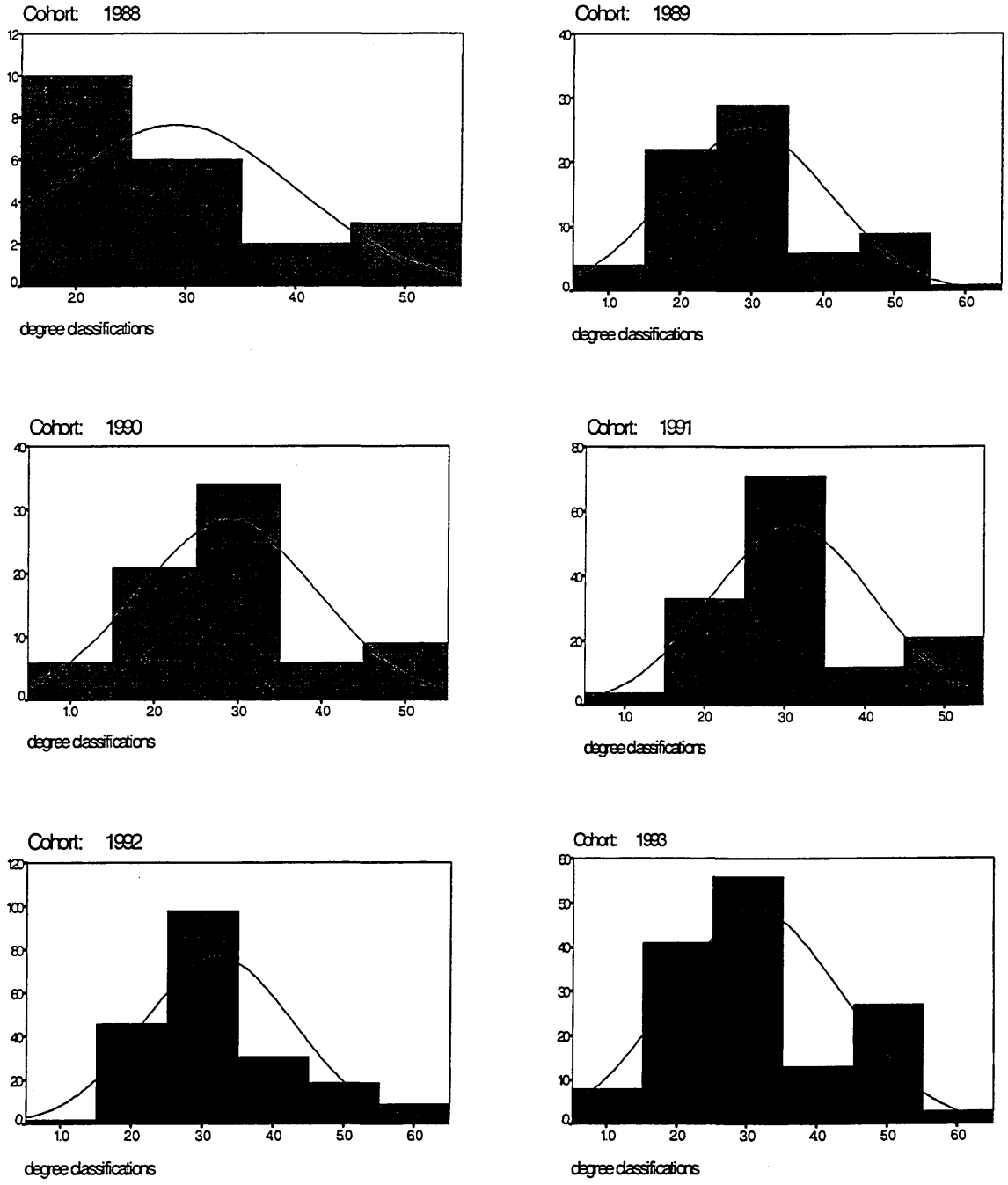
A level group



Note 1 = 1st, scored 72; 2 = 2.1, scored 66; 3 = 2.2, scored 54; 4 = 3rd, scored 44; 5= pass, scored 35; 6= fail, scored 25

Fig. 3.3 Distributions of degree classifications of the six cohorts

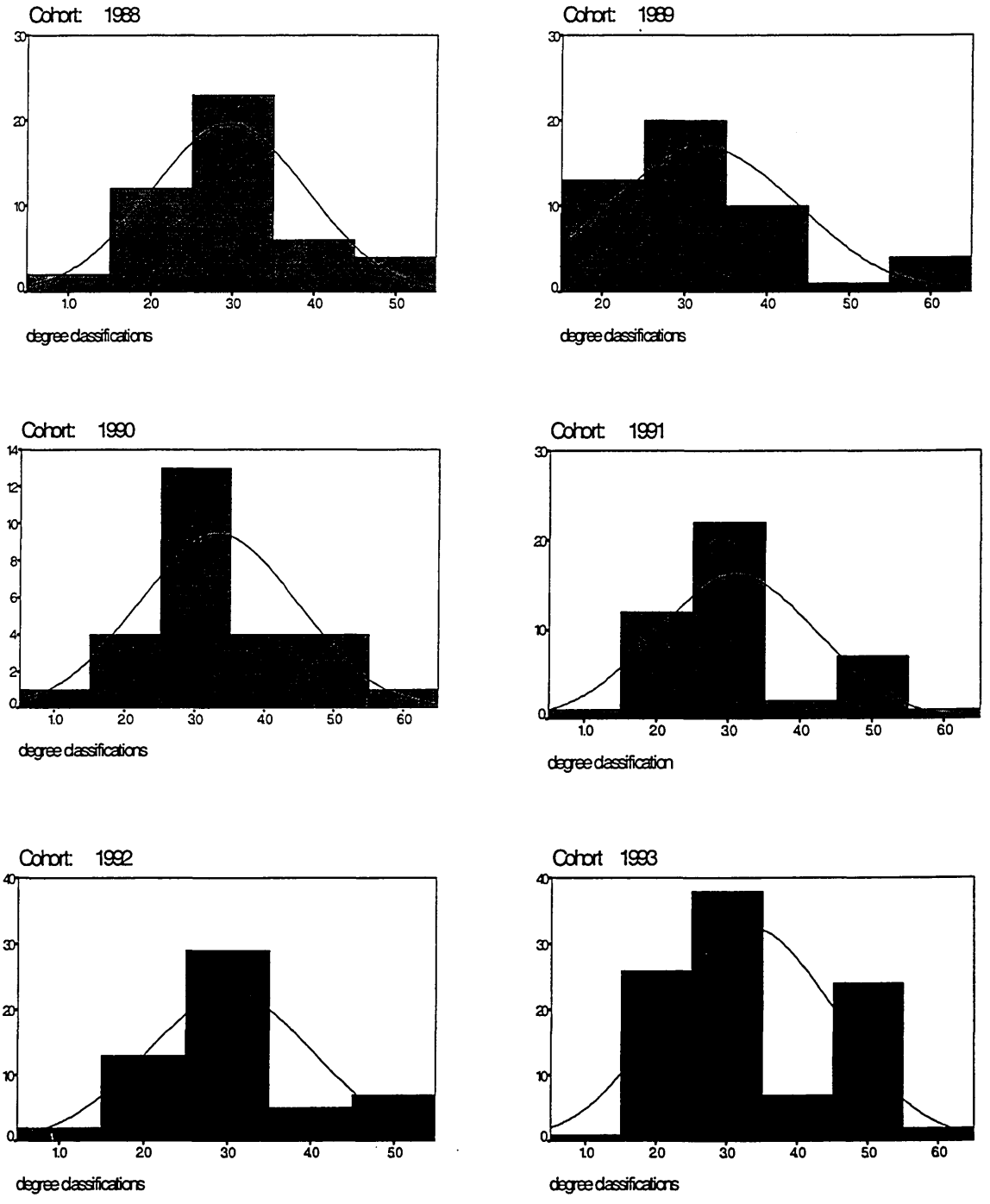
BTEC group



Note: 1 = 1st, scored 72; 2 = 2.1, scored 66; 3 = 2.2, scored 54; 4 = 3rd, scored 44; 5= pass, scored 35; 6= fail, scored 25

Fig. 3.4 Distributions of degree classifications of the six cohorts

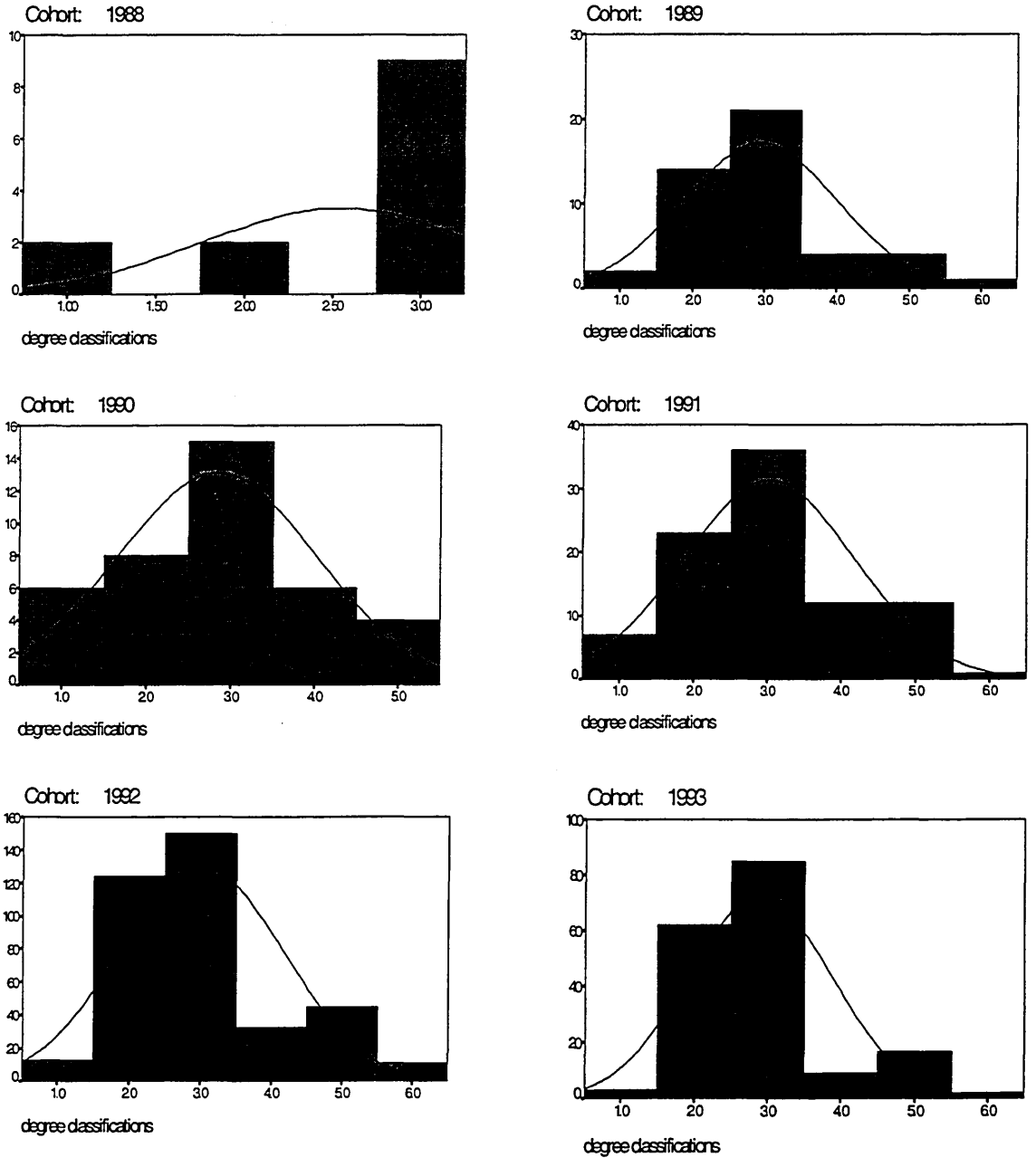
Access group



Note: 1 = 1st, scored 72; 2 = 2.1, scored 66; 3 = 2.2, scored 54; 4 = 3rd, scored 44; 5= pass, scored 35; 6= fail, scored 25

Fig. 3.5 Distributions of degree classifications of the six cohorts

'Other qualifications' group



Note: 1 = 1st, scored 72; 2 = 2.1, scored 66; 3 = 2.2, scored 54; 4 = 3rd, scored 44; 5= pass, scored 35; 6= fail, scored 25

Appendix 4 The information sent to the interviewees before the interview

From: Ms Li Wang
Learning and Teaching Institute
Addsetts Centre
City Campus
Sheffield Hallam University
Sheffield S1 1WB

Telephone 0114 2534740
E-mail L. Wang @ SHU.AC.UK

Dear

I am writing to ask if I could interview you for my research.

I am a research student, under Prof. Peter Ashworth's supervision, in the Learning and Teaching Institute at Sheffield Hallam University. I am undertaking a project on value-added as a performance indicator of teaching in HE. I have developed a method of calculating value-added and have applied it to measure the value-added of a group of students from your school.

As a member of teaching staff, you play a very important role in teaching and learning. Therefore, your comments on the value-added measurement proposed and whether you accept it as a performance indicator of teaching are very important for the research. It would be very helpful if I could interview you to have your comments.

I am enclosing a copy of my paper which explains how the model works and reports the results of the measurements. If necessary, I could explain more about the method when we meet (It would be better if you could read it before the meeting). I will contact you by telephone.

Yours sincerely

Li Wang

Value-added as a performance indicator of teaching in higher education in the UK

Background of the research

- In recent years there has been increasing emphasis on the accountability of higher education. The concept of 'value added' as a performance indicator has received growing attention.

- Proponents of value added suggest that to evaluate performance of teaching, the measurements of students achievement(i.e. degree results) must be related to measures of entry standards (entry qualifications)

- The idea of value added is attractive. However, to quantify the relationship between inputs and output is a formidable problem. The measurements proposed to measure value added (e.g. the index methods and the comparative method) and the suggestion that value added results should be used as a performance indicator of teaching, have met with considerable criticism

The method of calculating value-added

The method is split into three steps:

- The first step uses entry qualifications to produce an input rank order. This enables us to find the difference in academic attainment of students at entry between different cohorts.

- The second step is concerned with producing an output rank by using degree results in order to show the cohort in which the subject /course produces graduates who achieve more highly.

- The final step takes input rank and subtracts output rank. The results of this provides the value added score.

The population

Table 1 Business Studies Course

	1988 cohort		1989 cohort		1990 cohort		1991 cohort	
Number of students	108		191		162		127	
No of students with								
A level	67	62%	146	76%	118	73%	77	61%
BTEC	11	10%	18	9%	24	15%	24	22%
Access / conversion	4	4%	4	2%	0	0	3	0.2%
O level	1	1%	0	0	1	0.1%	3	0.2%
No formal qualific	0	0	0	0	2	0.1%	0	0
Other qualification	0	0	0	0	1	0.1%	0	0
European Bacalaur	0	0	0	0	1	0.1%	0	0
Degree ordinary	0	0	0	0	0	0	1	0.1%
Missing data	25	23%	22	12%	15	9%	19	15%
No of students with								
1st	1	1%	0	0	1	0.1%	0	0
2.1	46	43%	108	57%	44	27%	30	24%
2.2	15	14%	38	20%	65	40%	44	35%
Referred	0	0	7	4%	8	5%	10	8%
Deferred	0	0	4	2%	4	3%	7	65
Missing data	46	43%	34	18%	40	25%	36	28%

- A Level and BTEC were two major groups of entry qualifications

The results

The value-added for A level and BTEC students (Table 2)

- The 1988 cohort of A level students had most value added in the four cohorts.
- The 1990 cohort of the A level students appeared to have least value-added
- The 1989 cohort of BTEC students appeared to have most value added
- The 1990 cohort of BTEC students had least value-added in the four cohorts

Table 2 Value added for students with A level and BTEC

cohort	1988 cohort	1989 cohort	1990 cohort	1991 cohort
Number of students with A level	52	117	82	45
Mean of A level points	15.69	15.93	17.32	15.02
Rank of input	3	2	1	4
Percentage of good degree	79%	75%	43%	45%
Rank of output	1	2	4	3
Value added	2	0	-3	1
Performance	1	3	4	2
Number of students with BTEC	9	11	17	13
Percentage of good degree	67%	82%	35%	46%
Performance	2	1	4	3

The Interview Schedule

Topic: whether the current value-added measurement can be used as a performance indicator of teaching

Suggested questions:

1 Is this method of measuring value-added readily understandable?

2 Do you have any questions about the method and the information sent to you?

3 What are the problems with the measurement?

4 Do the results of the measurement reflect quality of teaching and why?

5 What information or reality was missed?

6 Can a high value-added score be interpreted as a sign of a better performing course?

7 How do you explain the results of value-added (that is, a particular value-added result)?

8 Can the measurement be manipulated by individual members of teaching staff or examination boards?

9 What are your main concerns if the measurement is to be used to measure the quality of your teaching?

10 Is it acceptable for the value-added measurement to be used for the purpose of diagnosing problems with teaching, and why?

11 Is it acceptable for the value-added measurement to be used as a performance indicator of teaching with related formal incentives / penalties, or used to allocate resources for the purpose of accountability, and why?

Transcription of interview with Carol
in Business School

L = Li, C = Carol (The name of the interviewee is not real)

L 1: There are two things that I hope to find out from this interview, one is validity, whether these value-added results reflect the quality of teaching, and another is the acceptability of the measurement. And I have some questions.

C 1: Right

L 2: The first thing is, do you feel that this method is readily understandable?

C 2: It would have to be very readily understandable. Firstly let's make sure that I have understood what you are effectively saying. You are looking at different cohorts entry qualifications coming in and looking at the difference in their performance on the final degree.

L 3: Yes, here, I looked at the population first, this is one course and I found the two major groups of entry qualifications - A level and BTEC and then I looked at the two major group's performance. I made a separate measurement, the first is for A level students, so I measured the mean of A level points for each cohort, and then ranked the mean, and then I looked at the percentage of good degrees and ranked the percentage and then I put the two rankings together to produce these scores and I call it the value added scores, but they are relative scores because it is the result of comparison of these cohorts.

C 3: Then with BTEC students have you done that ?

L 4: The BTEC students, the information that I got for them was only they came in with BTEC, I do not know the grades.

C 4: (1):Because they should have recorded against them their number of merits and distinctions that they came with, so there would be a way of separating out the goodness of one BTEC student compared with another, and again a bit like you did with the A level points it could be calculated on a proportion of distinctions, to merits combination or a ranking of distinctions, merits, pass and therefore working out an average score based on that. So the BTEC students could be differentiated year on year.

(2): But I tell you where you hit a very obvious problem, when you know what has happened on our degrees because as you can see here, between 1989 - 90 there has step down in the percentage of good degrees and I can tell you why that has occurred.

L 5: Oh Why ?

C 5 : Basically, the largest proportion of students on our degrees do_Unit A as their option, the Unit A external examiner was unhappy about the proportion of 2-1s on the degree and therefore the programme management at the time made the decision that all units should attempt to normalise around a 57% mean, so that then , the proportions of 2-1s from unit A would be the same as the proportion of 2-1s from HR and so on and so forth.

Now, I am not convinced that 57 was the right point to normalise around, but the consequence was, that very significant step change between 89 and 90 and the consequence is, if you like, it has become part of the organisational learning that now you expect your mean to become 57% and therefore the proportion of 2-2s is significantly higher, we very rarely now get any firsts because the problem is if all units normalise around that and they have to do so many units and in order to get a 2-1, they're aggregate needs to be over 60, that if they drop in one subject, then they are not likely to get over 57 in other subjects, the proportion of 2-1s suddenly comes tumbling down. Whereas there was no attempt to normalise prior to that and there was no course view about what average that we should be working too.

The problem is degree classifications does not necessarily indicate how good the students are, and this is an argument that I have had since. In 1990 the course team at that time, made this decision, it had an impact immediately as can be seen and that programme team continued for the next few years and it is now almost part of unspoken law, that this is how we mark so, effectively it has become part of the group thing of the people teaching on the programme.

And I personally do not think it has anything to do with the calibres of student that is being output from the course.

L 6: Yes, the degree classification cannot tell you everything about a student, but if we don't have any other tangible measures of student attainments.

C 6: This is part of the problem though isn't it, if you go to financial studies, as if you go to science. Science as we know has a far greater proportion of firsts than any other subject, financial studies always has a far less proportion of firsts. I think back, if my memory serves me correct, someone was doing some research on this, out of all the subject areas, business and management are the third from bottom in proportion of firsts that they give, so it almost becomes irrelevant as to what degree classification you

get, and as to whether that is value added, it is surely dependant on what degree you study for and therefore if you go to science, and you see all these people who had 2 points to get onto their degrees and then there will be lots of people getting firsts and 2-1s because of the nature of the subject and because of the academic community thought on it. I would argue that our students get as much value added but it is qualitative rather than a tangible measure of first, 2-1s, 2-2s, which then undermines your hypothesis.

L 7: Was the change of the policy for degree classifications from 89 ?

C 7: It came into effect in 1990, because I remember the course leader asking me at the time how I would normalise around a mean of 57% and so I said this is how you do it, so that is how I know how the policy came about.

L 8: So from 1990, the degree classifications should all be under the same policy, if I compare, for example. 92 and 93

C 8: So if you take that out of it, if you like, there has been a change, but like I say if you look at value added in one group relative to another, that is one thing. But then if you start comparing value added from one programme to another, this is where the problems arise because there are variability in the group thing for the academic community, so you get the situation with engineering, with science Vs those areas like management and business and finance

L 9: You mean the criteria of the degree classification is different. Is that difference very big ?

C 9: Well I was talking to someone who was doing a PhD on this subject, I don't know if they still are because they went away for two years, and they were looking at business and management because that is where they came from and they were looking at assessment generally,

I mean how come nobody gets more than 70 on an assignment in business and management and it is nothing to do with the quality of the paper relative to someone in science who may get 100%, It is a difference in approach that the academic community take to their subject, and therefore if somebody gets 75%, you are probably saying that it is about as good as it will ever get but there I am still reluctant to use the full stretch of marks.

L 10: Are there any ways to compare the criteria between different degree classifications.

C 10: I mean there are qualitative statements about what makes a first a first, a 2-1 a 2-1 and a 2-2 a 2-2, and I was writing some notes for guidance on dissertations and asked the centre for, if there was any university policy other than if you get an aggregate of 70% or more form your part one and part 2 averages then you will get a first. Is there something that says, this is what differentiates, and they didn't have anything. Now I do not know if they have since come up with one, but I think the comment at the time was 'No, but if you can think of some way, then that would be very interesting.'

L 11: But does the degree classification mean something ? If you say that it is all non- comparable between different cohorts, even within the course.

C 11: Within a course and within a programme, yes maybe, like I say if you take 1990 onwards then that will probably have more validity than comparing 88-89 with the following years. So those were taken out because there was a change in the policy of the course which isn't a reflection on the quality of the students, then yes.

The other question that has to be asked is the validity of the entry qualifications, every year, the tabloids say, decrease in standards of education, and then meanwhile more students get top grades at A level, now, either something very significant is happening between the ages of 13 and 18, so we are getting declining standards at the age of 13, but then suddenly from 13-18 we are getting improved standards or there is a slippage in the standards of A level points anyway.

Now that is a very big question and not one that I'd like to comment on. If that is the case that there is this step improvement in the education between 13-18, one question is why is the Government and tabloids so interested in declining standards at 13, if we are saying that those standards are being made up so rapidly between, or one says that the standards of A levels are declining and you can take one stance on this relative to another.

If you then take a BTEC student which is probably a better example, if you did separate out the BTEC students in terms of their entry qualifications, proportions of distinctions, to merits, to passes, then you could rank them and it would be quite easy. Someone who gets all distinctions they are better than someone who gets all passes for instance.

But then you get a very interesting point, in that, although it is a national recognised qualification and it is held under the EDEXCEL banner who ensure standards are consistent, we know for a fact, if you go to some small town FE college you will come out a greater proportion of

distinctions than if you came here. Now that means it is a bit pointless looking at how many distinctions because you need to cross tabulate that with the ranking of the institution.

L 12: Could I make an assumption that the standards of A level points and the standards of degree classification are the same? because I think that the standards of A level points and the standards of degree classification are supposed to be the same, or we try to make them comparable, although in reality, it is not true.

C 12: I think your safest using the A level, rather than BTEC, like I say you could rank BTEC entrants relative to each other. Similarly with GNVQ students and you are probably fully versed in the controversy there is in the differentials of standards applied in GNVQs across the country and across institutions and some institutions produce very good students and in a very true reflection of their ability, some produce very poor ones with excellent results and how do you know until you have taken in sufficient numbers whether it is the qualification or the institution so A levels is probably the safest standard in that although there may be drift that will take a lot longer to work through and although there may be differences between one A level board and another A level board like AEB used to be considered very lowly whereas Oxford and Cambridge A levels were considered very high, when I was a student. So there are still differentials but probably a better indicator of standard

L 13: Do you think, maybe, we can do something to ensure that the standards of entry qualifications and the standards of degree classification are maintained?

C 13: I am not sure what you could do, I mean , the Government keeps trying to say that we are maintaining standards.

The problem is, it is so complex and that is why performance measures tend to be so crude and so open to criticism that they become, can become almost meaningless, it is a bit like the ranking of institutions on the quality assessments, Oxfords Library apparently got hauled to bits by their students, if our students had access to what the Oxford students have, they would be very pleased. So again it is qualitative information.

L 14: Do you feel that the standards of A level points are different from before and the standards of degree classification are different, do you think it is right ?

C 14: Well, I think it is right, I think that the students, I have been working here for 9 years and unless I am just getting old and tainted, but

the students you see with 3 Cs, do not seem as bright as students who used to come with 3 Cs, now that could be like I say me getting older and more jaundiced or on the other hand it could be indicative of something that is actually happening meanwhile you get students that have got 3 Cs, you get students who have got 2 Ds, and they can seem equally bright which brings me probably onto another thing which my, what I have been told previously, and you have probably researched this and know much about it, but there has never been found to be a correlation between entry requirements and final degree classifications.

L 15: It varies with different subjects. I think that there is a correlation between the A level and the degree class in the business studies, but it probably isn't very strong,

C 15: But then how are you measuring value ?, if you saying if you have a good A level student, they will be good at degree level and therefore to move them on

L 16: Yes, but we are using A level to select students in university admission.

C 16: It is a rationing procedure, we have so many places, so many students want to do it, how do you choose between them? It is like I went to a talk by somebody from the law department at Sheffield and they were saying to do law you are going to need 3 As or 2 as and a B. It is nothing to do with the fact that you need those qualifications to do a degree, you could probably do the degree with Bs but more people want to do it than there are places. How do we decide ? We up the marks

When you have 5000 applications for your courses and you have somebody who has 3 hours a week, to make decisions on 5000 students, what other method can you use, it is the university accepted method, you would really be putting the trend to.... Now I personally don't believe that it is necessarily a good indicator as to whether the student is suitable for the degree, whether they are going to cope with the degree, because they could get 3 As in biology, chemistry and physics it wouldn't mean that they would be much use at business whereas somebody who gets a C in business studies and C in media studies will probably be much better suited.

L 17: Right.

C 17: Just a further point about the ranking that you have got here, with the ranking how much does this actually reflect value added ? Because what you are saying is, if they rank highly on their input grades and rank more lowly on their output grades, there is less

value added. But if they come in with quite low results and they get quite high degree classifications then there is more value added. But if you take, I know with the input grades, there is quite considerable distance but with this 43% Vs 45%.

L 18: That is quite close

C 18: That is very close.

The other thing is then, are these the A level points of the people who finish ? Or are they the A level points of the people who started

L 19: Started

C 19: Because it may be that weaker A level points here, we had a bigger withdrawal rate because they couldn't cope with taking a converse perspective.

L 20: The entry qualifications and the degree classifications were matched by student ID number

C 20: So it is only the output people that you have got the entrance qualifications of, it is matched by individual students.

L 21: Yes.

C 21: The thing is, A level entry and degree outcome on nice clean measures, assuming that there are no other variables to impact almost.

But it doesn't reflect the richness of what happens for instance, this year there were 82 students with A level entry and here there were 45, now I could not tell you for certain, but this cohort looks like it came from a large cohort whereas this cohort looks like it came from a small cohort because there is fewer people on this one and how well they did seems to reflect the size of the cohort probably as much as.... You might then say the quality or the size of the programme rather than the A level points affects the outcome for those students, that would be an alternative interpretation.

what you could read these statistics as, in 1990 there were a lot of students, they received less individual attention the class sizes were bigger. And that had an impact on their performance. Here where we have got a smaller group, their attainment was a lot higher, which you might say was value added, you might say with these students that they may not have attained anymore.

The idea of value added assumes that there is a correlation between A level points and the potential to attain high levels, but it is predicated on that, as evidence does not necessarily support that, then rather than the measure of value added from where they start to where they Finnish, a better method might be to say, right, if we rank the outcome and forget about the input; The qualitative factors which influence the output then is that as good a measure of value added

L 22: Yeah, that could be another way to measure it, but entry qualification is one of the factors which has an impact on degree classification. Although it is contestable whether it can indicate students' capability to do a degree, it is the only measurement available of their starting points.

C 22: You might use as a measure of output though, the proportion entering a job within 6 months and average salary.

L 23: Yeah, I could do that

C 23: Which might be a better indication for a vocational course, it adds to their value added.

L 24: Yes, but there are also many factors that could influence whether these students get a job within 6 months. It could also be the same as a degree classification.

C 24: I agree, what I am saying is, the whole issue is complex, you can simplify it, but by simplifying it you've missed out on many of the factors and this is why many performance measures, and value added measures are contestable, can be criticised and consequently can fall into disrepute.

L 25: If we do it in two stages, the first stage is to produce a result - value added. But we don't use it for allocating resources, and the next thing would be to investigate what factors might have some impact on this result. Would that be useful to look at the statistics in this way?

C 25 Like I say I suppose it depends on which standpoint you come from. If you think that A levels are an indicator of someone's ability, and if you think that degree classification, particularly from vocational courses is a true reflection of somebody's capability and what it is predicated on is the proportion / how many people can pass exams well, then yes it works. If you don't believe that then...

Yes, I would agree that there is not a true definition of what is the difference between degree classifications, changes in university

assessment rates can overnight impacts on what the degree classification is, it can affect whether somebody actually gets a degree or not. On top of that, you have got further complications when it comes to comparing courses, like I said coming back from different backgrounds, so comparing one course with another course, a course from with a course from science, just would not work and then meanwhile, you have got the local changes, like an external examiner stands up and says you have to do this and overnight you see the effect. We are expected to respond to the concerns and so on of the external examiner but consequently it looks like we have had far less value added ever since, by responding to a quality concern. We could have as many first class students as we want, if we just alter the way we asses them, how generous we are with our marking, if we alter the assessment regulations but that is not saying the quality of the students, nor their experience

L 26: No, one can manipulate the thing, the standards

C 26: Add to that the changes in course, for our students they find it terribly difficult to do quantitative things, if we had the course redesigned and input more quantitative elements the degree classification would come down over night.

L 27: So the degree classification does not mean a lot, but we still award students degree classifications.

C 27: I mean that is why, with a lot of employers, don't even know what a degree classification means unless they have got a degree themselves, undoubtedly with the increase in higher education more managers would have knowledge of that and it would therefore mean something, but as much as anything they will look at what university the student goes to, now if that is value added, somebody could come out with a third from Cambridge and probably still pick up a better job, than someone with a first from here.

L 28: So you mean the employers use the reputation of the institute as a criteria. What about if the two students who are all from this institution, how do they judge that ?

C 28: If they are comparing two students, ones got a 2-2 and ones got a 2-1 and they are both from here ?

L 29: Yes, Do they use the degree results ?

C 29: I couldn't have any evidence on that but it may be that the 2-2 doesn't even get an interview because they use that as a cut off point

L 30: What do you mean by 'cut off point'?

C 30: 300 applicant for a graduate training job, anybody with a 2-2 we don't look at and that narrows it down, and we'll now consider 50 applicants.

L 31: Yes, in reality it is very complicated.

Analysis of Interview with Carol

Stage 1: The points made by Carol were placed on the analysis sheet:

No. 1 Name: Carol School: Business

No. of paragraph	Points	Categories
C 3 - C 4 (1)	BTEC students are different	
C 4 (2), C 5, C 7, C 8,	Degree classification does not necessarily indicate how good the students are	
C 6, C 9, C 10	What degree class you get is dependent on what subject you study because different academic communities assess their subject differently	
...	...	

Stage 2: Categories were named. This was done by:

- marking the transcript with a coloured stripe (green) down the left hand margin;
- cutting up those paragraphs which deliver the points (see the above analysis sheet);
- putting paragraphs C 4 (2), C5, C6, C7, C8, C9 and C10 together as a group (group 1) because these paragraphs deliver the same points:
  Degree classification does not necessarily indicate how good the students are, naming this group as 'factors',
- putting paragraphs C3 and C4 (1) together as another group (group 2), and naming group 2 as 'differences between BTEC';
- after reading the transcripts of the individual groups and referring to the previous literature and the public debates, naming the categories firmly:

  Group 1: 'Factors that have an impact on degree classification'.
  (1) paragraphs C 4 (2), C 5, C 7, and C 8 : 'The way the examination boards behaved',
  (2) paragraphs C6, C9 , and C10: 'Academic community'.

  Group 2: 'Value-added and comparability'

- reviewing the categories and putting the names of the categories on the analysis sheet (see the analysis sheet on next page)

Stage 3: Points, groups, and the names of categories were all checked again.

No. 1 Name: Carol School: Business

No. of paragraph	Points	Categories
C 3 - C 4 (1)	BTEC students are different	Value-added and comparability Differences between BTEC students
C 4 (2), C 5, C 7, C 8,	Degree classification does not necessarily indicate how good the students are	Factors which may have an impact on degree classifications Factor suggested: The way the examination boards behaved
C 6, C 9 , C 10	What degree class you get is dependent on what subject you study because different academic communities assess their subject differently	Factors which have an impact on degree classifications Factor suggested: Academic community
...	...	

Transcription of Interview with Margaret and Ann
in School of Health and Community Studies

L = Li , M = Margaret, A = Ann (The names of the interviewees are not real)

L 1 Before you comment on my research, can you just talk about value added as a performance indicator very generally?

M 1 I've been involved with the Applied Social Studies degree programme for a number of years - in the 80s I did quite a lot of work on the admissions side, both admitting students to the 3 year route and also to the 4 year route, which was then the Applied Social Studies degree with a certificate of qualification in Social Work.

And of course that selection process was, it goes without saying, not only would someone, could they get a good degree, but would they make a good social worker? So we had to interview all those students and we were looking at life experience, relevant work experience - either in a paid or voluntary capacity - so the selection process was, you know, partly dealing with the academic side of it, the ability to do a degree, but also the vocational, professional side of it - could they, would they make a good social worker?

So, I think when we're looking at value added, when we're looking at students who are on the social work programme, and I think the same would go for some other degree programmes in this school - perhaps quite a lot of the professional courses at the university - there's value added in terms of what they are starting with as formal academic qualifications but there's also the sort of value you're adding about other things, which are not perhaps so easily measurable but maybe about work experience, life experience, qualities of personal maturity.

When you're recruiting students to these sort of programmes, for instance to social work, you have a responsibility to the student obviously as with all recruitment, would this student do well on this course. Because you don't want to recruit students to courses they're going to fail, that's not in their interest. You also have to think about the potential clients in the case of social work or patients in the case of say health services. Because if somebody is a bit mentally unstable or just wouldn't have the right professional attitude, they could do a lot of damage to the clients. Indeed, on the social work route for instance, students were, in the last two years of the programme, going on placement where they were doing social work, they were doing it under supervision, supervised by experienced practitioners, but certainly one was very conscious with particular social work colleagues when they were interviewing students, this represents responsibility not only to the students, but also responsibility to potential clients, without at being at least sure the person would do no harm and hopefully would do some good. Is that relevant at all?

L 2: Yes, I think it is a question about what is the value. Here I only measured the degree result, it's part of academic achievement, and what you have said I think is quite important, but the difficulty is ... it's very difficult to

measure it, especially when you try to do some statistical things. I think that's a big issue in value added as a performance indicator. It's very difficult to measure something like social value, or as you said professional quality.

M 2: (1) Yes, I think a lot of us who work in higher education institutions like this one are very committed in a sort of social and political sense to the concept of value added, because we know we're not always recruiting students who've got the highest A-level grades. We're recruiting, certainly on the social studies programme, which is perhaps one of the reasons why you've chosen this study, a lot of mature students, people who may have been in education for a number of years, and may have been out of the labour market bringing up children, may have worked in particular industries. Half the time people have been made redundant in the coal mining industries, the steel industries and so on, or engineering. So, they're not people who've just come straight through the educational system, been to, you know, a good sort of secondary school, grammar school, public school, got good A-levels and come straight here. You know a lot of them, as mature students, have families, they may have health problems, they may have experienced distressing life events like marital break-up or being made redundant before coming here.

(2) So, I mean, yes we very much believe, politically, in the idea of value added as a justification for what we're doing and in terms of a fair assessment, a fair performance indicator - because if you have no notion of value added, if you just compare the results of how people perform in final examinations without saying well, look some people have had every advantage to do with social class, supportive family background; other people have had to cope with all sorts of discrimination - perhaps to do with a disability they have, perhaps people coming from a working class background. If we don't take into account those differences, then we're not fair to the students or to the teachers. I mean I've known students and I've seen them go across the platform at the conferment ceremony and they've got a 2:2 degree - but I've known that given all the things that have been going on in their life; you know, while they've been here they may have suffered the death of several close relatives, or they've been ill, they've had poverty to contend with or whatever; it's still a heroic achievement to get a 2:2 degree.

(3) So I think the value added concept is important. I mean, what concerned me about your paper when I read it is how you measure it, because you can measure what they come in with in the way of academic qualifications, you can measure what they come out with - but of course there are so many factors that influence output, there's quality of teaching, there's motivation of the student, there's their financial situation.

(4) I think one of the things many people in higher education are worried about now is the increasing amount of time students are working in paid employment to support themselves through their studies, and we don't see that getting better as a situation, but it is detrimental for the students. When I was a student, people may have worked in the holidays, but they didn't work much in term times, whereas now because students are working two or three evenings a week, or they're leaving classes early or whatever, they

don't have the time for the broader reading, they don't have the time for discussing ideas with other students - all the deepening, enriching parts of higher education are being lost with it.

L 3 Yes, I agree. There are limitations of what I have measured. There are things like what you have talked about which I didn't measured, but may be, I could use this to identify problems.

M 3 (1) I think it's very useful if your work can identify just what are some of the problems about a value added methodology. You see, another aspect of it, which I think is going to be a topical, political debate I'm afraid soon, is what do the students get out of the degree? And you've got that in two ways. Obviously there's all the stuff about personal development in terms of doing a degree, which I still believe in.

(2) But there's also what do they get jobwise. Now my current responsibility is to be a final level tutor in the Applied Social Studies programmes, so I write quite a lot of references for students. And sometimes I write references for the students, and they're going back to the job they did before they came here. They were a care assistant in an old peoples' home before they came here, they've come here and got their degree, and now they're going back to that job. And if you look at the statistics, the way they're massaged by the Careers Office, they now define a graduate job as any job a graduate does. Right - so it's no longer the case that being a graduate in itself guarantees you any particular place in the labour market.

(3) So, I think there's going to be that issue for students of "what do I get out of it" in that sense, and I saw in the papers at the weekend that the Dearing Committee is talking about charging students £1,000 a year - I think for a place like this, and particularly a degree programme with a lot of mature students, people just won't come, they won't think it's worthwhile.

(4) So I think that's going to be another problem about value added, if one's saying, what does it add, either to the earnings of the individual or to the economy in terms of getting a degree? And that very much ties in with how far a student's motivation - academic, to do with interest, how much is it financial? It ties in with all these questions. But I think your model is useful in terms of identifying all these factors.

(5) I suppose the other thing too, if you're doing a longitudinal study over time, is the whole issue about how degrees are calculated, what the standard of a degree is, and that's obviously a very thorny problem, but is that changing at all over time? That may be an issue that comes in as well.

L 4: So, you think that it is necessary to develop a value-added performance indicator.

M 4 Well, we're being required by whatever government is in power to have certain performance indicators to justify public money in terms of how universities and colleges are run.

I think what's important is that performance indicators should not first of all consume so many resources to calculate the information, that they take

away resources from delivering the service to students; and the other issue is of course, if you measure anything, the danger is that that becomes the thing people focus on, and they ignore other things. Now, I mean, clearly one of the important things about higher education is helping students to get good degrees - that is our priority. The degree processes are also about personal development, about higher education and so on. If we are to look at standard of degrees, it is probably all right although of course we've got to make sure that the standard isn't inflated simply by grade inflation processes.

I think it's a question of how you measure what they come in with, because it's one of those difficult to predict things - I don't think there is any clear evidence people have got about how particular entry qualifications work out and define a class of degree. I mean if we did have that, the whole process of university admissions would be an much more exact science that it is. So I think that's one of the issues.

I think it's interesting to look at how one can measure value added, and obviously there are a number of ways. One could be by interviewing or questionnaire studies of students to find out what they themselves feel they've gained from the university education. Then obviously there's the academic attainment measures, which is what you're working on, which is another sort of area. There's is the whole thing about what it translates to - if it does in terms of jobs. That's another big area.

A 1: Can I ask about the methodology that you've used? Because this idea of value added came about, as I remember, about ten years ago. There was a PCAS booklet published on it, and my understanding at the time was that this was in the context of widening access to higher education and the idea that taking students with different types of qualifications - access or GNVQ or whatever, you had to look at what they gained when they were here in terms of the value that was added to that original starting point. So that they may come out with a Third Class degree, but the value added to them from the starting point to the degree was quite substantial.

Now I understood that that had run into some difficulties with measurement, and that people were finding it very difficult to measure the starting point and the finishing point. Looking through this, I wasn't clear how you identified the ... you've obviously identified different types of qualifications to enter, have you ranked these qualifications?

L 5 That's a big problem in measuring value added, because the types of entry qualification are different, so it's very difficult to give scores to these different types of entry qualifications. Here, what I did is, before ranking any qualifications, I looked at the populations and then I found out that A level students and Access and conversion course students are two major groups, so I didn't take into account students with other entry qualifications, because there are very few students with other entry qualifications. Then I made separate measurements for these two major groups, comparing A-level students with A-level students, and comparing access students with access students. And for the A-level students I used their A-level points as a measurement. Because access students didn't have any grades, I assumed they were the same.

A 2: Right, because within the access group there are huge differences in what students come in with. They could have been on a variety of courses - full-time, part-time. But in terms of credit levels for entry, the higher education certificate is given to an access student with 16 credits, 12 of those being at what they now call level 3, which is entry to higher education, equivalent to A-level. Now a lot of students on access courses do actually come out with far more than 16 credits. They may come out with as many as 35.

So there is a huge difference within that group in terms of how many credits that they've got, and you would have thought that would have some impact then on how well they do here. Students who've got experience of coping with a lot of work, all at the same time, come out with a lot of credits, because the credits relate to the volume of work, rather than the level. Then you would expect those students to possibly be better prepared. So, I'm slightly doubtful about you putting all the access students together in one big group, I think there are differences between them in the same way there are differences within the A-level cohort.

L 6: I agree, this is a limitation of this study. I didn't get their actual credits. What I got, the data is only that these students came in with access or conversion course. So that's a limitation.

A 3: So how do you, how did you work out the value added then? For different groups?

L 7: For the A-level students, because they've got A-level points I can calculate the mean of A-level points for a certain cohort, for example 88, and then I can rank the mean of A-level points. I'll show you here. Here, this is the mean and this is the rank. And then I calculated the proportion of good degree for different cohorts and rank and when I compared the ranks I got this result. These scores are relative, this shows that this cohort moved from there... their output is in the first, and input is third, so that means that they have had value added.

A 4: Right, plus 2. So does that mean that the 1990 cohort, their value added was a lot less, so that's the way it works?

L 8 Yes. It is relative, and it is only the result by comparing with these three cohorts. If you take one of the cohorts to compare with other cohorts that would be a different result. So it's relative

A 5: So it looks here then that the cohort, the 1990 cohort, that has highest mean level of entry points has the lowest value added.

L 9: Yes

M 5 Because they have started at a higher point.

L 10: Yes, that is the question. it looks like - it's difficult to add the value to students who has high A-level points. Is that true? if we just talk about academic attainments.

A 6: There are quite a few studies about the relationship between A-levels points and degree outcomes. I don't know if you've read any of that material - but my understanding was that the general view was that the higher A-level points, the better the degree that students come out with. Now, in one level that does not measure value added, does it? So what you're saying is that they might come in with high A-level points and go out with a high degree score, but the value added is a lot less than people coming in with low A-level points and maybe a mediocre degree level. Is that what you are saying?

L 11: Yes.

M 6 It's difficult reflecting on it from the experience of teaching, because I think we see some students, particularly at the top end of the range, where we are conscious of very definite intellectual progression here, that they really do make academic progress, that they are ready, you know, when they've graduated to go on to post-graduate study, and looking at their performance we can see a clear improvement over time. That's absolutely the top end of the range.

At the very lowest point of the ability range, there are students who struggle through each year with referrals, and who are going to come out with a very low 2:2, possibly a Third, where one feels that they're not really operating at the sort of level you would expect intellectually of a university graduate. I mean these are a very small percentage I'm talking about, but they stay in the memory because one is very worried about them, because they just never seem to get it intellectually, in terms of understanding the more sophisticated nature of questions, the more discursive elements or the finer points of theory. They still operate in too simple-minded a fashion really, and it's like, however hard you try with them, they really don't get it.

Then there are a number of students, there are a lot of students one gets in the middle of that range. There are some who always get 58%, they always get a good 2:2 but never a 2:1, and sometimes these are very conscientious and hard-working students, and you do the best you can to help them but they just haven't got the intellectual sparkle and flair. They just haven't got the capacity for the more penetrating intellectual thought. I know that there are some who work steadily and do come out with a 2:1, so a lot of our students obtain satisfactory degrees, but not brilliant degrees. These are the people who perhaps feel they need to get a degree to get a job, they're not necessarily sparked by the love of learning or anything like that! And so with a lot of those, there perhaps isn't a lot of value added but they plod through it at a fairly pedestrian way and come out with a high 2:2 or a low 2:1. As I say there are those who do a lot worse, who really struggle but just somehow manage to get some sort of a degree.

And then there are a small number one sees who, one feels, really do make a lot of intellectual progression. But I think even, you know, with all the students who've got 2:1s that there has been intellectual progression so, yes we can see value being added in terms of comparing, say. the first draft of an essay with the final one, or in our final year we require students to do a dissertation and that I think is a very valuable learning process, I think a

M 6: lot of the value gets added there. It's an assignment that causes students a

lot of anxiety at first because they have to define their own problem for investigation, so it's not like we give you a set of essay questions and you pick one and go away and do it. It's a very different learning process and they have to work on it all the year. But sometimes, not always, but sometimes when you get the final dissertation to read having commented on rough drafts, chapters throughout the year, sometimes you're pleasantly surprised, and you realise the student has got it together, and they have achieved some intellectual development. So we do try very hard as part of our teaching and learning philosophy to develop students as independent learners and to have some notion of intellectual progression, so that as they go through the years the essay questions become more challenging intellectually and require the students to be more analytical and more focused. But as I say some manage to do it and some don't really.

A 7: (1) One of the largest things; I mean obviously ability is a key factor in terms of how fast you progress; but I think possibly equally as important is motivation. And whatever students come in with, if they're highly motivated they generally come out with a decent performance at the end.

(2) But one of the things that might not be measured here are the students that fall by the wayside. Does this model take account of failure or withdrawal?

L 12: No. I got some students whose results were 'refer', that means some students who may fail in one unit, and they haven't got the final result, they may re-do it or may not, so I feel it's difficult to measure those results. so I didn't take them into account.

A 8: So do you take the people who are graduating and then trace them back, is that what you did? Or did you take the people that were enrolling and trace them forwards? Because there's a very big difference in terms of what you look at.

L 13: I traced them forward. If some students came in the second year they were not taken into account.

A 9: So the ones who withdrew or failed as they went through, what did you do with those?

L 14: I didn't take them into account because I think that another performance indicator may be used to measure this, and it is very difficult to measure how much value is added to these students if they withdrew or came in the second year. Yes, that is what I found. I only measure a limited number of students, a part of the students.

A 10: It might be important to think about how many of these different types of students are not managing to make it, because you might find that the value added to, say, A-level students who stay is great, but a fair number of them actually withdraw, and you know it may be different with other types of students. I mean for access students we know that a fair number of them don't actually make it, a lot of them withdraw or fail as they go through, but the value added to the ones that make it to the end of the course might be quite significant.

L 15: Yes, I think if you talk about value added, you have to know who you're

comparing with. If you're comparing students who finished the course with those who didn't finish the course, than certainly all the students who finished have added more value and those who didn't finish.

A 11: There are so many factors in here, it's very difficult.

M 7: Yes, I think the interesting thing about the sort of changing nature of higher education - and you've probably looked at some of the books published by the society for research into higher education, because they're very good at covering these debates - is so much the changing nature of the student population. Because traditionally in Britain it was a very elite system, it was just a limited percentage of the population going to university. Now we're aiming for much higher participation rates, more a sort of lifelong learning and people coming back into higher education, so students are so much more diverse as a group.

And it seems to me that perhaps implicitly any concept of value added methodology is still operating by measuring these people, who are very diverse as students, against a norm of the sort of, you know the full-time under-graduate who goes to university generally at age 18, as I did with A-levels, graduates at 21, and who doesn't in that time have children or have any serious illness or have anything that disrupts their studies, and goes straight through the system. So we still have that notion of the "typical" student, or the "typical" under-graduate. Whether they're even still statistically the norm now I'm not sure.

But we deal with so many students who've had different sort of patterns. I mean certainly on a degree programme like this, it was only the social work part of it, at that time the student was untypical. We've got a lot of part-time students as well in the university. When you are looking at that sort of system, which is much more like the American community college system, it's very difficult to say, well, what is the typical student? And if we are talking about value added, people coming from so many different starting points are going to so many different destinations. That seems to be an interesting question politically about your work, that in one sense there is still that model we've all got in our heads of what we think the typical student should be, and indeed we deliver the curriculum and we teach it to some extent to that model. But we know that's not what life is like for a lot of our students.

L 16: Yes, but if you look at the entry qualifications you may find certain groups of students is the majority, for example access students, for certain courses. And then you can look at how much value was added to them.

M 8: Yes, I think certainly one of the groups we're interested in, certainly what I very interested in. When I was an admission tutor is, how students from different access courses fared. Because I remember being an Admissions Tutor back in, must have been the very early 80s, when access courses were quite new, and there weren't many of them, whereas now they very common. And one of the things I was conscious I had to deal with when I was looking at the forms, was some access courses were full-time, some were part-time. Some lasted one year, some lasted two years. Also, some simply had a pass/fail distinction at the end. Some had

pass/merit/distinction.

M 8 So what they were saying to the students is, well, we won't fail anyone because access course tutors are loathe to fail students because they feel they've taken a really big fail once by the educational system - so their whole sort of professional ideology is about achievement, empowerment, giving students positive feedback. Well that can mean they're not very good at saying to students who can't really cope with higher education, well, perhaps it's not for you.

And the problem is, it's like students go from a warm bath on the access course, to a cold shower here. Because they come here, it's selective, it's competitive, feedback is about what they've done wrong as well as what they've done right, there is a real possibility of failure, they've got to compete with others who are also quite capable. It's not all sort of lovely and supportive and empowering and empathetic, it's more about critical thinking and developing analytical skills. And what it is, it has to be because we have to get the students to a certain level of academic achievement - we can't just tell them how good they are all the time.

But I think there is that problem, and it's much better I felt, if the access course had the three grades of pass - the pass, the merit and distinction - because then they could use the merit and distinction as qualifiers to higher education. But the problem is sometimes the word "access" seemed to be used in two meanings. There's access - to a return to learning for people who have been out of formal education, which has a sort of taster quality, people can use it to get a taste of education; do they like education, is it their scene, is it what they want to do with their lives? And that's a perfectly valid and legitimate educational role for such courses, to have an access course which is about return to learning, a sort of very broad, wide access. Then there is access, as in preparing you for higher education, which is obviously more selective, more rigorous and must be about trying to do an equivalent process to A-level. Now one does wonder with a lot of these courses whether they can get people in one year to the point which an A-level course gets people in two years. So I think there are all these questions about types of access courses and clearly the more information one can get about how well the access course is succeeding in preparing students for higher education is very important.

Now in one way that may be profoundly unhelpful in your work, because obviously access courses group as a category. Having a sort of feel for the area, one had the feeling that they're not all equal, some access courses were better preparations than others. For instance, on some access courses students didn't sit any formal examinations at all. In others they perhaps did them, at least as a learning experience. Now, I can see the difficulties for access tutors - that if a student never sits an exam - well maybe people are frightened to do exams, and they've got to sit down and write for three hours, you know, an unseen examination - that can be seen as very intimidating, and frighten many students - but if they've got to do it when they come here, it's much better for them to have some practice sessions on the access course, even if it's only as a learning experience rather than a formal tool of assessment. So, I think there are all those issues about access courses that a lot of us are concerned about; partly in fact the sort of culture, and of course also the role of tutors is difficult on

access course. We get reports from students on access courses of how lovely and accessible the tutors were, how they were always there to see them, spent loads of time with them. But these people were employed solely in a teaching function - and it is in our jobs we're expected to do teaching, we're expected to do research, we're expected to do course management - and maybe serve on university committees. We can't be accessible to see students all the time, because if we were accessible to see students all the time, we would be failing to do significant other parts of our jobs, other parts of our contract. So again, when they come into higher education they have to learn the lecturer has a slightly different role. The lecturer is here to teach, to give tutorials and so on, but lecturers also have other duties. So I think there are all these issues about how well access courses actually do prepare students to be independent learners.

L 17: When you recruited access students, how did you select them ? What was the criteria that you used?

M 9: We sometimes used to interview them. Obviously where the qualification was a sort of pass/credit/distinction we could require some sort of credit. The problem with a lot of access courses, it was simply pass or fail. So if you made an offer to the student, it would appear as a conditional offer on the application forms, but in practice a conditional offer to a local mature student was pretty close to a firm offer.

So in terms of the sort of number-crunching side of it, because if you're an Admissions Tutor, you're job is to try and get the right number of students on the programme enrolled in September; because if you get too many you've got a problem because we haven't got the staff to teach them, and if you get few you've got a problem because jobs are at risk because there are not enough students coming in. So it's a wonderful exercise in managing uncertainties - wonderful management training exercise in coping with uncertainty being an Admissions Tutor! And we give it to new and junior members of staff as a 'perk'.

But the thing about it is, whereas with A-level students we're always making a conditional offer - it would probably need something like about ten offers to recruit one student in September. It was quite, you know, quite a crazy process really. With a local mature student, if we say like two offers would generate one person actually enrolled. So with a local mature student on access course, say somewhere in the College A, the reality was you'd nearly made a firm offer. Formally speaking you were making a conditional offer, but in practice you near as damnit made a firm offer. Because you knew they were extremely unlikely to fail the course.

Another problem actually, of course, for access students was the whole timing of the admissions process. When I was Admissions Tutor it was the polytechnic's central admission system. But the whole thing was that applications had to be in by some date in December, so the tutors on the access courses had to write the references about students, in sort of October, November time. Now this was all right for sixth forms, because the sixth form tutor would have taught the pupil all through the lower sixth year, so they've already taught them for a year, then in the Autumn term they would write a reference for them, so they could often do it sort of

September time, perhaps with a predicted A-level grade and I'd get the reference.

But the reference I got from the access tutors was often useless, because they'd been teaching the person for six weeks and they had to write a reference for them. Well you can't make much of a professional assessment of a student's ability - often one got the sort of famous sentence ...

L 18: Why is it only six weeks?

M 9 Well, that type of ... I mean, they start teaching in perhaps September, the application form's got to be in by mid-December - well yes they could take a little bit longer.

But the thing is, you have to start making offers as soon as you get the forms, you don't wait till you've got all the forms in, so it's not in the student's interest to put the forms in at the last minute. Although the access ones did tend to come in later than the sixth form ones, and you'd get sentences in these references like, "it is expected that by the end of this year the student will be the right standard to enter higher education" - which means of course they're not now But they might be or they might not be at the standard by the end of the year, but they couldn't really tell I don't think. I mean you could interview people, you could sometimes look at essays they'd written, but in some cases tutors would help very heavily with writing those essays. So it was a bit hit and miss, but, while access courses did a lot a good work with lots of students in terms of study skills, essay writing skills - how far they could really be used as an effective selection method to judge ability and motivation for higher education is harder to say I think. We've had some wonderful students from access courses, and some have been struggling. I don't know if Ann agrees with my impressions ...

A 12: Oh yes. Definitely

L 19: So I understand now that there are big differences between access students. And we need to measure the differences.

A 13: You can probably measure how many credits they've come in with, but you would need to go through College A to do that. They keep a record of the number of credits that each student is awarded.

L 20: What are these credits?

A 14: (1) They're notionally based on time spent studying, so they need to be up to level 3 to get into university. But the number of credits is a separate thing - it's related to how many hours they've spent studying, which is important for preparation of students to deal with degree level work. The fact that they can deal with a lot of work at once, which is reflected in a high number of credits, is quite important. So you could probably measure that.

(2) But what you couldn't measure is equally important, and that's things like the domestic situations of the students; whether they've got broken marriages; children; no children; financial difficulties or whatever; and also the motivation of the students - and that tends to be variable.

(3) Generally access students are motivated, probably far more than some A-level students, who you feel, or you tend to have a gut feeling, that these are people who are expected by their parents to come to university and they're doing it, they're not really bothered about it - they just see it as a normal kind of thing for people their age to do. Whereas access students have generally made a lot of sacrifices to come. So they are generally motivated.

(4) But what tends to happen when they get here is that there's an interaction between the whole domestic situation and the academic life here, which often leads to many students dropping out. And that's something that is very difficult to measure.

L 21: You have taught this course?

A 15: Yes

L 22: So you think that for the access students, mature students, their domestic problems are a big factor to affect their studies?

A 16: Domestic problems, financial difficulties, yes

M 10: I think it's important to say that not all mature students come through access courses - some come through things like the Open University foundation courses, some take A-levels, so some of them do come in with an A-level which they've studied at a part-time evening course.

We have known of cases of students where the marriage is breaking up because the partner can't cope with the new person they've become. I mean in some ways if that's happening it's better if the break-up comes before they start the degree programme. But in some cases, particularly I think for mature women students, there can be even active hostility from the family. There are some studies of women in higher education which have quoted these sorts of things as well, so there's a lot of experienced studies which show this sort of thing - where it's all right for the woman to study but she must have a nice meal ready on time, she mustn't leave her college books lying around the home. She mustn't reveal her knowledge when she's talking to people in her family or in the pub. She's got to pretend to be the person she was. I think women get this sort of pressure from the family worse than men. But I think men may feel bad about not playing the breadwinner role, and not supporting their families.

And sometimes, too, people have very real problems with teenage children, who are perhaps getting into trouble with the law or being tearaways - and that's very, very difficult. Or indeed even sometimes possibly grandchildren. I think one of the problems for a lot of working class students is if they are the first person in their family to go into higher education, the family may not appreciate the demands of higher education, they may think it's a bit of a skive being a student, not see it as a full-time job. So, women mature students in their 40s, for example, may be asked by their daughters to look after their grandchildren or something like this, when they should be doing their college work. So sometimes, if the woman is playing the traditional role, of the woman in the family, holding the family together, she's taking on too much to do as well as her studies. So, there could be that sort of

problem. And then there are cases of very nasty divorces.

Sometimes men if they resent their wives education, they get very mean and they just don't give them any housekeeping money. I knew of one student who was refused money to repair the washing machine when it had broken down, or refused money to repair the cooker, so she was expected to cook and wash without a functioning cooker or washing machine. Now she still got a very good degree and so on, but had she got a divorce?.. So sometimes there is almost that sort of persecution.

L 23: So although mature students have all these sorts of domestic problems, but some of them still obtain a good degree. Why? what is the factor to make them obtain a good degree?

A 17: Motivation, I would think.

M 11: I think motivation, and I think good organisation.

L 24 What do you mean by 'good organisation'?

M 12: I mean they don't leave work until the last minute. This is a mistake that a lot of the younger undergraduates make.

L 25 They know how to organise their lives?

M 13: And they get on with work early. For instance for the dissertation they make sure that they work at it steadily throughout the year. But for some of them, yes it is a heroic achievement. But I think it's about them making sure that they work effectively when they are working and they do manage to keep some time for doing their work and for getting on with it. But it is difficult. But I think that thing about, you know, where maturity can be a positive asset is about being able to organise your time, whereas younger students may, not necessarily waste more time, but they use it in different ways - they don't get up so early and they don't get on with things so quickly.

A 18: But against that, in the past we used to make attempts to accommodate mature students far more than we are able to do now. We used to be able to timetable them between 10 and 3 in the afternoon so that they could do their education around the children being at school, we used to take account of half terms, we used to give them reading weeks when the children were at home for half term holidays.

What's happened increasingly over the last few years is that the university, the centre as it's called, has imposed a more and more rigorous schedule on terms, semesters, assessment and also the teaching day. We're now technically teaching 9 in the morning till 9 at night. It's become increasingly difficult for us to accommodate mature students and take account of their special needs. So although I think individual teaching staff do spend more time with mature students when they need it, there's now a limit to what we can do to make their academic life smooth for them.

M 14: Yes, I mean it's difficult to generalise too much about students - but I think mature students are more conscientious about using tutorial time, coming in

and getting help. I think it's partly they know they can't afford to leave it to the last minute, whereas younger under-graduates can sometimes fail to use tutorials for weeks on end and then suddenly expect you to give them hours of time at the last minute, when you can't necessarily. I think that is a bit of a difference. And I always take the view that if people recognise they've got learning deficiencies or ask for help, that is in some way a sign of intelligence. The students one is most worried about are the ones who are too stupid to realise and don't think they need help, think they know it all and discover at the end that they don't. So, I think there is that aspect. For some mature students too I think it's worth saying, if say they went to a grammar school to the age of 16 they may have had a good basic education, in terms of good basic standards of literacy and numeracy, whereas for some of the younger ones those qualities may be weaker. And that can help them, even if they've been out of education for a time.

L 26: Just some questions I've listed, some of which you've explained, but can you give me any explanations or suggestions why 1990 cohort A-level students had much less value added?

A 19: Have you looked at the numbers of admissions that year? Wasn't the 1990 cohort particularly big, wasn't that the one where all ..?

L 27 I've got the full information here.

M 15: There were some variations. I mean one of the difficulties with part of the programme, certainly the Applied Social Studies part, ...

L 28 That's the total number, the whole groups, whole cohort.

M 16: ... you may with male students in some groups have a very small sample and it may be difficult to make, I don't know whether there are just too few men sometimes in the sample to make a reliable statistical analysis.

A 20: It was the 1991 cohort that was huge. The 1990 cohort was OK - so it's not the size of the intake is it?

M 17: Of course if you're comparing 1991 and 1988 you can see whether there are any variations in standards of A level marking or perhaps grading system.

A 21: You've got fewer access students on the 1990 cohort, so it's not that there was a sudden big intake of access students who had an effect. The A-level grades in 1990 were better weren't they?

L 29 1990 was the highest. And it's interesting, the 1990 access students, they had most value added compared with other cohorts.

A 22: I think it's probably something intangible. You sometimes find that you get years of students that just don't seem to gel, the groups don't fit together very well, the students don't seem to be integrating very well - and that seems to happen some years doesn't it? I can't think of any other reason.

M 18: No, I mean you do get variations in the standard from one year to another, and I think it's difficult when you're looking at quite a small sort of sample, to make reliable statistical predictions. The other problem is for instance if one looks at indices about the high proportion who are failing, at certain times

the university has changed regulations about whether marginal fails in the 35 to 40 range be compensated or not, if they've got to pass a unit in aggregate, or whether they got to pass both components of a unit - we've got some of that sort of stuff going on, which might sometimes be what's behind, what appears to be a change in the statistical outcome, but really it's just that the ways of measuring students performance have changed - there can be some of that factor.

L 30 Do you have any memories that in 1990, the standards increased or something?

M 19: I don't - I mean one would have to check back through examples, regulations. The Academic Registrar Department might be able to help with some of that. I just think that sometimes there may be things like that happening - so if for instance, this year we're going to have more 2:2 degrees and fewer 2:1s than last year - well that's actually because the university's changed the method of calculating the degrees. It's not that the students are any different necessarily.

L 31 There is one more question - Do you think if the value-added performance indicator can be used with formal incentives and penalties?

M 20: I would definitely say I strongly disagree with that.

L 32 Why?

M 21: Why? Well partly because we have a proverb in English, you can take a horse to water but you can't make it drink - and with some students you do your best and they still don't learn anything. You can't necessarily measure the quality of the education simply by what students achieve, because there are things to do with the students ability, there are things to do with motivation. Now we can have some impact on motivation, we can get some students interested - or we can bore them to death! - but there are some students, take those for instance who have got serious drug problems. We're not going to be able to motivate them - that's a very small number - but sometimes, however hard we try to motivate students, they've got, I don't know, mental health problems, or they've got things like that going on which makes it really difficult. Similarly with academic attainment - I'm sure some of my students could achieve better academically if they had more money, and didn't work three or four evenings a week in a supermarket. But, I can't personally do anything about that. And I think these are some of the issues that ... and they are adults, they do have a choice how hard to work. Our responsibility is to make it clear to them what they need to do to get a good degree, to convey information about our subjects clearly - which is partly about presentational skills but it's also fundamentally about subject knowledge. So I think there are ways you can measure the quality of teaching, but there isn't a one to one relationship between the quality of teaching and what students achieve, because there are other factors that come in to it. Family influence sometimes, student motivation - and of course just what they brought from their primary/secondary education. Because if they've got gaps in numeracy, which is probably some of the science areas, or literacy, which really should have been addressed at primary or secondary level, we can do what we can, but we're not actually qualified remedial teachers.

A 23 In addition, if you introduce performance indicators of that nature, then what will happen is that people will make their courses easier, make their assessment easier, and give higher marks.

M 22 That's the other side of teaching by results.

Analysis of Interview with Margaret and Ann

Stage 1: The points made by Margaret and Ann were placed on the analysis sheet:

No. 7 School: Health and Community Studies Name: Margaret and Ann

No. of paragraph	Points	Categories
M1	there is also the sort of value you are adding about other things which are not so easily measurable.	
M 2 (3)	The difficulty with measuring value-added is that there are so many factors that influence output.	
M2 (4)	The increasing amount of time students are working in paid employment has an impact on student achievements	
M 3 (1) - (4)	Value of a degree in terms of employment	
M 3 (5)	The ways in which degrees are calculated and how standard of a degree changes over time	
A 2	Within the access group there are huge differences in what students come in with.	
A 7 (1)	Motivation is a key factor. Whatever students come in with, if they are highly motivated they generally come out with a decent performance at the end.	
A 14	Domestic problems, and financial difficulties have a great impact on mature students' academic achievements.	
A18	A rigorous schedule on terms, semesters, assessment has made mature students' academic life difficult.	
...	...	

Stage 2: Categories were named.

- The transcript was marked with a coloured stripe (orange) down the left hand margin.
- The paragraphs which deliver the points (see the above analysis sheet) were cut off.
- The paragraphs which deliver similar points were put together as a group:

  Group 1: paragraphs M1, M3 (1) -(3)
  Group 2: paragraphs M 2 (3), M 2 (4), M 3 (5), A 7 (1), A 14 and A 18
  Group 3: paragraph: A 2

  Each group was given a provisional category:
  Group 1: other value than degree results
  Group 2: other factors than quality of teaching
  Group 3: differences between access students

- After I read the transcripts of individual groups and referred to the previous literature and the public debates, the categories were named firmly:

  Group 1: The concept of value-added
  Group 2 : (1) factors that may have an impact on student achievements
  (2) factors that may have an impact on degree classification

Group 3: Value-added and comparability

- The categories were reviewed and the names of the categories were put onto the analysis sheet:

No. 7 School: Health and Community Studies Name: Margaret and Ann

No. of paragraph	Points	Categories
M1	there is also the sort of value you are adding about other things which are not so easily measurable.	The concept of value-added
M 2 (3)	The difficulties with measuring value-added is that there are so many factors that influence output.	Factors that may have an impact on students' achievements
M2 (4)	The increasing amount of time students are working in paid employment has an impact on student achievements	Factor: The increasing amount of time students are working in paid employment
M 3 (1) - (4)	Value of a degree in terms of employment	The concept of value-added
M 3 (5)	The ways in which degrees are calculated and the standard of a degree changes over time	Factors that may have an impact on degree classifications
A 2	Within the access group there are huge differences in what students come in with.	Value-added and comparability
A 7 (1)	Motivation is a key factor. Whatever students come in with, if they are highly motivated they generally come out with a decent performance at the end.	Factors that may have an impact on students' achievements Factor suggested: Motivation
A 14	Domestic problems, and financial difficulties have a great impact on mature students' academic achievements.	Factors that may have an impact on students' achievements Factor suggested: domestic problems and financial difficulties
A18	A rigorous schedule on terms, semesters, assessment has made mature students' academic life difficult.	Factors that may have an impact on students' achievements Factor: management of teaching
...	...	

Stage 3: Points, groups, the names of categories were all checked again.