Issues relating to the statutory assessment of design and technology at key stage three - 1989 to 1993

Richard Tufnell

submitted for the award of

Doctor of Philosophy

University of Warwick
Institute of Education

September 1995

Contents

	page no
Abstract	v
Chapter 1	
The Context and Aims	1
An outline of the National Curriculum	2
The Task Group on Assessment and Testing (TGAT)	3
Validity and Reliability	9
Issues relating to criterion-referencing	12
The impact of assessment activities on learning experiences	14
The relationship between different modes of assessment	17
Chapter 2	
Towards a National Curriculum in design and technology	23
The making of a subject: Responding to social change	24
The rise of the subject	27
Design and technology comes of age	29
The promotion of design and technology	32
The vocational movement of the eighties	34
The influence of the Assessment of Performance Project	37
Defining a National Curriculum subject	40
The assessment of design and technology: From marking to assessment	45
The recognition of project work and course work	48
A common examination for all	52
National Curriculum assessment - criterion referencing	56
Conclusion	58
Chapter 3	
The context surrounding the research and development of statutory tasks	63
The context	64
Developing standard attainment tasks - the development agencies	64
The steering procedure	67
Main steering September: 1989 to August 1990	68
Main steering September: 1990 to August 1991	70
Subject steering December: 1989 to August 1991	72
The Summer of 1991	76
Research and development: September 1991 to August 1993	79
National curriculum design and technology: Blue Peter or Micky Mouse?	81
The first statutory assessment	85
Confrontation with the unions	89
Conclusion	91
Chapter 4	
Developing the Research Instruments	96
The development of Assessment Instruments for Statutory Assessment	97
The suitability of the statements of attainment for assessment purposes	100
Research and development - the assessment instruments for practical tasks	102
Research and development - the assessment instruments for tests	112
The development of the pupil material - practical tasks	119
The development of the pupil material - tests	126
Conclusion	139
Chapter 5	
Pupil Performance in Major trials and PilotsThe method of analysis	141
Section A - Practical tasks	142
The 1990 Trials	144
The 1991 Pilot	151
The 1992 Pilot	157
The 1992/3 Pilot	163
Section B - Pencil and Paper tests	167
The 1992 Pilot	167
The 1992/3 Pilot	175
Conclusion	185

	page no.
Chapter 6	
Teachers' evaluation of the assessment process	188
The approach to teacher evaluation	189
The 1991 National Pilot	190
Teacher evaluation - the context and the respondents	190
The administration of assessment tasks in school	190
Manageability	191
Assessment and moderation	196
Other issues	200
The 1992 National Pilot	**208**
Teacher evaluation - the long task	208
The manageability of the long task	208
Teacher evaluation - the tests	211
Conclusion	215
Chapter 7	
Evaluation and conclusion	217
Reaching conclusions	218
The climate in which the research was undertaken	218
Issues of validity and reliability	219
Issues relating to criterion-referencing	227
The impact of assessment activities on learning experiences	235
The relationship between different modes of assessment	240
Final conclusions	**244**
The role of assessment in changing the curriculum	244
The style of assessment device	245
The nature of assessment criteria	247
The interpretation of assessment criteria	248
The nature of practical tasks	249
The standardisation of pupil material	230
The nature of tests	250
The assessment process	251
Appendices	
2.1 The design and technology attainment targets 1988 - 1995	256
2.2 The statements of attainment - 1989 Statutory Order	257
2.3 Twelve general design abilities - CSE (NWEB)	261
2.4 Educational aims - GCSE C,D and T and Assessment Objectives	262
3.1 Issues relating to aggregation	263
3.2 Extracts from the 1991 MUTEC tender document	268
4.1 Stranded assessment matrix - 1991 Pilot	273
4.2 Stranded assessment matrix - 1990 Pilot	277
4.3 AT assessment matrix - 1991 Pilot	283
4.4 The level guide - 1991 Pilot	289
4.5 Remark Process - 1990 Pilot	291
4.6 Evaluation of the effectiveness of the assessment devices - 1991 Pilot	297
4.7 Task assessment matrix - initial proposals - 1992/3 prestatutory trial	306
4.8 Instructions and example assessment matrixs - 1992 tests	312
4.9 Task specifications - initial proposals - 1992/3 presatatutory trial	315
4.10 Example of 1992 Test paper - Te4	318
4.11 Example of 1992/3 prestatutory test paper -Te1	321
6.1 INSET evaluation - 1991 Pilot	329
6.2 Teacher evaluation - 1991 Pilot	333
A Research ownership	342
Bibliography	343
Primary sources	343
Published sources	346

Acknowledgements

I should like to acknowledge the following for their assistance and support in undertaking this research:

My supervisor, Professor John Eggleston, for his encouragement throughout my period of registration and his confidence in my eventual completion, this was most reassuring. I have valued in particular his perceptive insight and invaluable comments on this manuscript.

Professor John Cave, Tim Denning, Kate Johns, Margaret Goodchild, Richard Roberts and Peter Stensel, who as members of the project team worked on many other aspects of statutory assessment, complementary to this research, such as information technology, project management, Welsh issues, desktop publishing, special educational needs, Inset administration and culture fairness (see appendix A). The development of printed resources was jointly undertaken by this team.

The many teachers who administered these assessment devices, without whose cooperation and professionalism this thesis would have proved impossible. I appreciate that for many this represented an additional burden and one with which they may have had little sympathy.

The countless pupils who took the tests and tasks which are the subject of this thesis. I hope that these activities contributed in some way to their design and technology experience.

The School of Education at Middlesex University for support throughout my period of registration at Warwick University.

Valerie Tufnell for her encouragement throughout the four year period and patience in proofreading the text.

Declaration

The work presented in this thesis was part of a much larger project funded by the Schools Examination and Assessment Authority (see pages 19 and 20). This thesis draws only upon the aspects of the research for which the author was responsible. The outcome of this research was presented to SEAC at regular intervals in the form of confidential, unpublished reports (see Bibliography - primary sources). In addition, a paper was presented to the International Conference on Design and Technology Educational Research and Curriculum Development - IDATER 94. This paper sought to expose the research in this thesis to a wider audience for debate and discussion.

Abstract

The National Curriculum, introduced in 1988, brought the requirement that pupils' achievements should be measured and reported at regular intervals during their period of formal education. Design and technology was at the forefront of National Curriculum implementation, which was planned to take 13 years. This research was aimed at investigating and developing statutory assessments in design and technology to be administered to pupils at the end of key stage three, normally after nine years of schooling.* Design and technology, as defined by the National Curriculum, represented a significant change for the majority of schools in relation to philosophy, content, organisation and, particularly in the context of this thesis, assessment. Consequently, expertise and resources needed to be focused on the development of assessment procedures especially given that the National Curriculum is based on criterion referencing. This approach to assessment only recognises and records pupils' positive achievements.

In order to produce reliable and valid assessments, successive trialling and piloting took place over a four year period. The research and development of the testing instruments is fully reported, along with the detailed results of the application of these instrument in the major trials and pilots. This research resulted in the design of a wide range of innovative approaches to assessment, ones sufficiently robust for statutory assessment. It addresses both the nature of the assessment evidence and the devices by which the evidence is judged. Thus the repertoire of assessment in design and technology has been significantly extended, especially in the context of criterion referencing. Evidence is also provided which shows how assessments, which are comparable and fair, can be devised and applied across the breadth of design and technology activity.

The research was required to serve both political and educational objectives, consequently devising assessment procedures to meet their respective demands required compromise,. The key conclusions attempt to identify strategies which might in the future, if understood by all at the outset, produce satisfactory statutory assessments in this subject.

* These pupils, would by law have been taught the design and technology programme of study for this key stage from September 1990. They would not however, have been taught the programme of study for key stages one and two, the first six years of formal education.

Abbreviations

APU	Assessment of Performance Unit	
AT	attainment target	
CSE	Certificate of Secondary Education	
DES	Department of Education and Science	(to April 1992)
DfE	Department for Education	(from April 1992)
DT	design and technology	
ERA	Education Reform Act 1988	
GCE	General Certificate of Education	
GCSE	General Certificate of Secondary Education	
HMI	Her Majesty's Inspector	
IT	information technology	
KS1	key stage one	
KS2	key stage two	
KS3	key stage three	
KS4	key stage four	
LEA	Local Education Authority	
MEGNAP	Midland Examination Group National Assessment Project	
MUTEC	Middlesex University Technology Education Centre	
NC	National Curriculum	
NCC	National Curriculum Council	
NFER	National Foundation for Educational Research	
OFSTED	Office for Standards in Education	
PC	profile component	
PoS	programme of study	
SAT	standard assessment task/test	
SCAA	Schools Curriculum and Assessment Authority	
SEAC	Schools Examination and Assessment Council	
SoA	statement of attainment	
Te	technology	
TA	teacher assessment	
TGAT	Task Group on Assessment and Testing	
UCLES	University of Cambridge Local Examinations Syndicate	
Y	school year, i.e. Y1 to Y13	

Glossary of Terms

Attainment Target: defined in the Education and Reform Act as the 'knowledge, skills and understanding which pupils of different abilities and maturities are expected to have by the end of each key stage.' They refer to the subsets of attainments into which most National Curriculum subjects are divided. In design and technology the attainment targets are grouped at each of ten levels, they are not separately defined for each key stage, see appendix 2.1

Core subjects: the subjects which have been given priority in relation to curriculum planning and assessment. These consist of English, mathematics and science and, for Welsh speaking schools in Wales, Welsh.

Foundation subjects: The Education Reform Act defines these as, English, mathematics and science, for Welsh speaking schools in Wales, Welsh, technology, history, geography, a modern foreign language (key stages 3 and 4 only), art, music and physical education.

Key stages: There are four key stages which divide each pupil's schooling into prescribed periods; these are, KS1 - from the beginning of compulsory education to 7 years, KS2 - from 7 to 11, KS3 - from 11 to 14 and KS4 from 14 to 16.

Level descriptor: broad statements describing achievement at a level. Introduced in this research in 1990 and to be introduced for all subjects from 1995 as a result of the Dearing Final Report and the subsequent revision of the National Curriculum.

Levels of attainment: ten levels into which each attainment target is divided, 1 being the first level and ten the highest (art, music and physical education being the exceptions).

Long/practical task: an extended task undertaken during normal lessons in a prescribed time window yet still contributing to a statutory end of key stage assessment.

Profile component: a term used to group attainment targets for the purpose of reporting a profile of attainment. The National Curriculum subject technology had two profile components - information technology and design and technology.

Programme of study: defined by the Education Reform Act as ' the matters, skills and processes which are required to be taught to pupils of different abilities and maturities during each key stage. Technology had two programme of studies one relating to each of the profile components.

Standard assessment task: introduced by the Task Group on Assessment and Testing to refer to the many possible types of standard, externally devised, test materials to be used in making National Curriculum Assessments, particularly at the end of a key stage.

Statements of attainment: the attainment descriptors which constitute each level in each of the attainment targets, for each National Curriculum subject (art, music and physical education being the exceptions).

Statutory Orders: legal documents which put into effect the decisions of the Secretary of State for Education and the Secretary of State for Wales under the powers given them by the 1988 Education Reform Act, for example, the statutory Order for technology defined in detail the programme of study and attainment targets of this National Curriculum subject.

Teacher assessment : those undertaken by teachers as part of their normal teaching against national curriculum criterion and reported to parents at the end of each key stage.

Test: externally devised, pencil and paper, timed exam, invigilated by teachers and initially intended to be marked by teachers.

Tier (band): this refers to the subdivision of the levels of attainment into entry bands for statutory tests, for example, tier 2 - levels 3 - 6.

Chapter 1

The Context and Aims

Synopsis

This thesis is principally concerned with the statutory assessment of design and technology at key stage 3 of the National Curriculum. Design and technology is a foundation subject of the National Curriculum for England and Wales introduced under the auspices of The Education Reform Act of 1988. To establish the context in which this research was undertaken, this chapter describes in outline the National Curriculum. In particular the role of The Task Group on Assessment and Testing is considered and discussed as its key recommendations had a direct influence on statutory testing. In particular issues relating to end of key stage 3 assessment in design and technology, the focus of this thesis, are detailed.

The aims of this thesis are described in relation to the key questions which it seeks to answer. In brief these are concerned with the validity and reliability of national statutory tasks; the impact of adopting criterion referencing in design and technology; the effect of assessment activities on learning experiences and the relationship between the different modes of assessment. Each of these issues is reviewed in relation to the way in which they might effect statutory testing. In addition the aspects which will need to be considered in undertaking the research and development required are identified.

This chapter is divided into the following sections:

An outline of the National Curriculum
The Task Group on Assessment and Testing (TGAT)
Validity and Reliability
Issues relating to criterion-referencing
The impact of assessment activities on learning experiences
The relationship between different modes of assessment

An Outline of the National Curriculum

The National Curriculum was introduced under the auspices of the 1988 Education Reform Act[1]. This Act defined the content of the curriculum and instituted an administrative structure within which reform would take place. The principal objective of the Reform Act was to formalise and standardise mainstream practice in state schools as a means of raising educational standards and making education more relevant. The public perception was that the education system was failing children, expectations were too low and teachers were neglecting the basics, grammar and tables, in favour of more 'trendy' topics. Concern also extended to the teaching methods being employed. Formal strategies, often termed 'chalk and talk', had been replaced by child-centred approaches involving exploration and project work. These were seen as bringing disorder to the classroom and were believed to go hand in hand with the widely held view that there was a lack of discipline in schools. As long term unemployment trends continued to rise there was a growing opinion that education should be more geared to the utilitarian purpose of gaining employment and less self-indulgent.
'The public perception was of a system going rapidly downhill and that is the key in all the reforms that were to follow in the 1988 Education Reform Act.'[2]

The fundamental changes envisaged by the Act required new terminology with which the reader of this thesis must be familiar, see glossary p. vii. For example, the period of compulsory schooling, from 5 to 16 years of age in England and Wales, was divided into four key stages. Key stage 1 for pupils from 5 to 7 years of age, key stage 2 from 7 to 11, key stage 3 from 11 to 14 and key stage 4 from 14 to 16. Key stages 1 and 2 approximated to infant and primary; whilst key stages 3 and 4 divided secondary education into pre-examination and examination courses. In addition, a new nomenclature was introduced for the years of schooling. The year in schooling would now be described as a continuum from Y1, the first year in school for 5 year olds, to Y13, the last year in school traditionally referred to as the Upper Sixth. These changes were introduced to:
'Achieve consistency between types of school and between areas.'[3]
The adoption of this approach reinforced the continuous and progressive nature of a child's education, loosening the traditional institutional focus. The introduction of a common terminology and structure was an essential requirement for the national reporting and monitoring procedures introduced by the Act. This thesis is focused on key stage 3, the phase covering Y7, Y8 and Y9, normally pupils between the ages of 11 and 14 years.

The Act categorised subjects under two headings:
'The core subjects are Mathematics, English and Science. The other foundation

subjects are History, Geography, Technology, Music, Art and Physical Education at all stages and, for pupils in key stages 3 and 4, a modern Foreign Language.'[4]
Exceptionally the act continued by clarifying the content in relation to Technology:
'In practice Technology is understood as embracing many aspects of Design.'[5]
The Act prescribed that each subject should be defined by its own Statutory Order, which would consist of:
'i. Attainment targets (AT), defined as the knowledge, skills and understanding pupils are expected to have by the end of each key stage. They will provide objectives for what is to be learned in each subject during that stage;
ii. Programmes of study (PoS), defined as the matters, skills and processes which must be taught to pupils during each key stage. They will set out the essential ground to be covered in order to meet the objectives set out in the attainment targets; and
iii. Assessment arrangements, defined as the arrangements for assessing pupils at or near the end of each key stage, for the purpose of ascertaining what they have achieved in relation to the attainment targets for that stage.'[6]
It is important to note that Act provides a definition of the term 'assess'. Section 25(1) of the Act states that it includes, 'examine and test', allowing a variety of methods to be used as part of the assessment arrangements.

In the event the assessment arrangements were not included within the Statutory Subject Orders. Although the subject Orders could determine the minimum requirement for each subject, they could not prescribe the amount of time to be spent on any part of the programme of study, particular ways of providing a subject in the school timetable, teaching methods or approaches or text books and other teaching materials.

The Task Group on Assessment and Testing (TGAT)

The development of the assessment arrangements to accompany the National Curriculum was the subject of an independent task group. The first report of the Task Group on Assessment and Testing (TGAT)[7] was published in January 1988. This report recommended the principles and practices to be adopted. However, the then Secretary of State, the Rt. Hon. Kenneth Baker, asked the group for further advice relating to the administrative and other arrangements, such as INSET, needed to support their recommendations.

The Task Group established four general criteria which national assessment should meet:

'- the assessment results should give direct information about pupils' achievement in relation to objectives: *they should be criterion-referenced*;

- the results should provide a basis for decisions about pupils' further learning needs: *they should be formative;*
- the scales or grades should be capable of comparison across classes and schools, if teachers, pupils and parents are to share a common language and common standards: *so the assessments should be calibrated or moderated;*
- the ways in which the criteria and scales are set up and used should relate to expected routes of educational development, giving some continuity to a pupil's assessment at different ages: *the assessments should relate to progression.*[8]

These four key tenets resulted in the requirement that the assessment of the National Curriculum should be criterion-referenced, formative, ensure progression and moderated so that comparisons relating to performance could be made. These are commendable criteria and jointly they would ensure that national assessment would assist learning and support the professional development of teachers. Improved assessment methods and the exchange of information about expected outcomes would give teachers greater insight into the way in which their pupils learnt and the success of their teaching. These were ambitious objectives, and TGAT acknowledged that,

'..... no national assessment system has as yet been constructed which meets all these criteria.'[9]

The additional TGAT reports were published in June 1988. Following their publication the Secretary of State for Education and the Secretary of State for Wales then announced jointly, on 7 June, the adoption of TGAT's main principles as the basis for a national system of assessment and testing related to the National Curriculum attainment targets. The key features of the structure announced by the Secretary of State are given below:

- attainment targets will enable the progress of each child to be measured against national standards at the ages of 7, 11, 14 and 16;
- attainment targets should be grouped to make assessment and reporting, at these ages, manageable;
- different levels and overall progress demonstrated by tests and assessment should be registered on a 10 point scale covering all the years of compulsory schooling;
- assessment should be by a combination of national external tests and assessment by teachers;
- the results of tests and other assessments should be used both formatively to help better teaching and to inform decisions about the next steps for a pupil, and summatively at ages 7, 11, 14 and 16 to inform parents about their child's progress;
- detailed results of assessments of individual pupils should be given in full to

parents. Individuals' results should not be published, but aggregated results at the ages of 11, 14 and 16 should be so that the wider public can make informed judgements about attainment in a school or LEA;

- in order to safeguard standards, assessments made by teachers should be compared with the results of national tests and with the judgements of other teachers.[10]

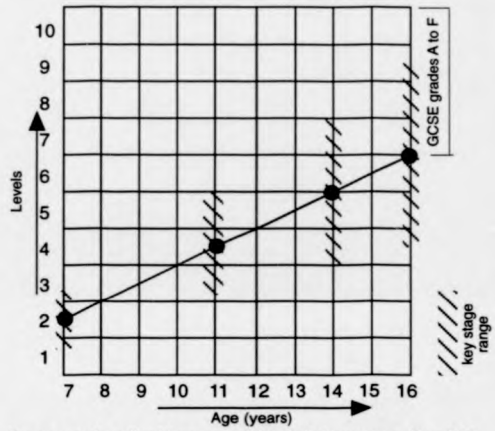
This announcement also detailed the actions which would be taken to implement this structure, In the context of this thesis the most important being:
'We shall also set in hand work on the development and planning of national tests.'[11]
The announcement was highly contentious as it confirmed a number of proposals which were bound to be greeted with hostility amongst the teaching profession. In particular the publication of aggregated results heightened speculation in relation to school performance league tables. School performance tables would only be possible with the introduction of national external tests. In its report TGAT had proposed that assessment would in future be characterised in two types, firstly teacher assessment (TA), principally formative in purpose:
'As a natural part of teaching, therefore, teachers are constantly assessing pupils to determine their progress and to plan the next stage of their learning. In the widest sense, such assessment involves a continuous comprehensive examination of all aspects of the pupil's learning, drawing on a wide variety of evidence from many sources to arrive at a general picture.'[12]
The second type was summative in nature and its aim was to:
'provide standardised, ie nationally comparable, assessment results.'[13]
This type of assessment was quickly to become synonymous with the term SAT, standard assessment task, introduced by TGAT.
'...our use of the word test will have a broader meaning. Ideally it might be better expressed by the phrase "standard assessment task" ....'[14]
Externally provided tasks and procedures designed to produce performance data on a national scale. It is with the evolution of statutory assessment tasks that this thesis is concerned.

TGAT's recommendation that pupil achievement should be directly related to the development of pupils' competencies as described by attainment targets rather than age-specific scaling was a significant break with tradition. It reinforced the notion that education in future would be a seamless continuum between the ages of 5 and 16. A pupil might move from one school to another but their achievements would go with them. The report insisted that:
'Only one set of criteria is required. The levels defined by the National Curriculum attainment targets will provide differentiated challenges at each age according to

the needs of the individual pupil. All pupils should then have the expectation of making progress in every profile component: progress which will be indicated by the achievement of new targets.'[15]

Progress in an attainment target would be marked by the achievement of successive levels over time. TGAT offered guidance on the sequence of pupil achievement between the ages of seven and sixteen, see diagram below. This sequence is based on the assumption that:

'a national curriculum level should be roughly equivalent to two years of educational progress.'[16]



Sequence of pupil achievement of levels between ages 7 and 16
(Figure 1 - Task Group on Assessment and Testing (1989) - A Report)

In relation to key stage 3 TGAT advised that, the average expectation for a pupil embarking on the key stage should be level 4 and for a fourteen year old at the end of the key stage an average pupil should be on the 5/6 level boundary. Should national assessments be calibrated to coincide with these performance recommendations? Although no view is offered by TGAT it is implicit in the report's arguments for adopting criterion referencing, that assessment procedures must be able to measure change, that standardisation was not part of the system. However, these recommendations do lay a norm-referenced veil over a criterion-referenced system!

Each level in an attainment target defines the knowledge, skills and understanding appropriate to achievement at that level. In the writing of the subject Orders it became necessary to dissect these as independent assessment objectives. Consequently, the term statement of attainment was introduced. This term was not part of the structure described by the TGAT report but had been universally

adopted by the time of the Policy into Practice document published by the Department of Education and Science in 1989.
'...the Orders will also contain statements of attainment which are much more precise and describe each of up to ten levels of attainment. The statements of attainment will provide the basis for the assessment arrangements.' [17]
The statement of attainment was intended to be the smallest assessment unit, sufficiently precise that a pupil could be deemed either to have evidenced it or not. In addition to implying an achievement, the term attainment also constitutes a goal - something to work towards. The same term is used by General National Vocational Qualifications (GNVQs). They use the term in contrast to a statement of competence, a term which is the preserve of National Vocational Qualifications (NVQs). In this context competence implies occupational or professional competence. It was defined by Jessop (1991) as :
'the ability to perform to recognised standards'[18]
In the context of National Curriculum design and technology, and other subjects involving specific skills, statements of competence might have been more appropriate than statements of attainment. Possibly the notion of competence is too closely linked to minimum achievement to provide a satisfactory curriculum goal.

The acceptance of TGAT's proposals allowed the draft subject Orders in mathematics and science to be published for consultation. As the three core subjects were given the highest priority the next group to be established was English. Each group had to define both the programme of studies and the attainment targets within the TGAT structure. The next working group to be convened was for design and technology. Under the Education Reform Act design and technology, uniquely among the foundation subjects, was designated as a mandatory element of each child's education throughout each of the key stages. Prior to the National Curriculum, pupils' experience of this subject was extremely diverse. The subject had been taught to very few pupils in key stage 1; it was in the process of being introduced more widely in key stage 2; the majority had some experience in key stage 3 but only a small proportion chose to do the subject post 15. (In 1993, the final year of pre-National Curriculum GCSE, 5.31% of all entries in all subjects were in craft,design and technology and home economics.)[19] A National Curriculum in design and technology clearly involved far more than formalising and standardising what already existed. In many schools it involved initiating and creating; clearly the introduction of a National Curriculum in isolation would be insufficient to turn the legislation into a reality.

The unique position in which design and technology found itself led to its being referred to as part of the extended core. Consequently, following the establishment

of the Subject Groups relating to the core subjects, one focusing on design and technology was the next to be established, see chapter 2. Following the established procedure a National Curriculum in technology was formulated and passed into legislation. During this procedure design and technology and information technology were combined under the title technology; each being defined as a separate profile component (PC) within the subject. Both the design and technology and information technology curricula, in common with the core subjects, followed the framework established by the Task Group on Assessment and Testing (TGAT). The first design and technology profile component was defined as design and technology capability and was composed of four attainment targets. The information technology profile component was defined as information technology capability. It was composed of a single attainment target. The use of the term capability in conjunction with each subject title reinforced their practical and purposeful nature. This study relates to the first of these profile components - design and technology.

Within the context of the National Curriculum at key stage 3 and the profile component design and technology this thesis aims to explore the following issues and propose answers to the questions which they pose:

- which aspects of the 1990 Statutory Subject Order in Technology, relating to design and technology, could be assessed in a valid and reliable way by national statutory assessment procedures?
- is it possible to devise and implement assessment procedures, based on criterion-referencing, which will differentiate pupils' achievement across the range of design and technology in a fair yet consistent fashion?
- can an assessment procedure be devised which will meet the political aims of National Curriculum assessment whilst providing pupils with valid learning experiences and pupils, teacher and parents with accurate and useful measurements of performance?
- to investigate the relationship between end of key stage assessment, terminal and summative in nature, and ongoing teacher assessment, continuous and formative in nature.

At the outset there was a range of issues underlying each of these questions which would need to be considered and explored in formulating answers. What these were, in relation to each question, will now be examined in turn.

Validity and Reliability

- *which aspects of the 1989 Statutory Subject Order in Technology, relating to design and technology, could be assessed in a valid and reliable way by national statutory assessment procedures?*

The trade off between validity and reliability has long been a central focus of analysis for test constructors. In the context of National Curriculum assessment both aspects are of critical importance. The scale of the operation, approximately 600,000 pupils and 40,000 teachers, posed significant problems. In addition the teacher would be required to operate in the unfamiliar role of an examiner and undertake a task for which they would receive no formal training. Classical measurement theory has identified different types of validity and reliability and has developed related methods of statistical analysis. However, many of these measures are not applicable to criterion referenced level scores. Traditional reliability coefficients rely on the proportion of variance in the obtained scores which is predictable from the true scores but criterion referenced measures do not tend to produce sufficient variance for these coefficients to be possible.

Validity concerns the soundness and legitimacy of the test/tasks. Four main types of validity can be identified as approaches which could be used to review assessment tests/tasks.

1. Content validity - This concerns the items and activities which make up the task. They should match the objectives of the curriculum and the actual content of the teaching and learning experience, i.e. does the task/test actually examine what the children are supposed to have learnt? In the case of the National Curriculum this would be the programme of study for the appropriate key stage.
2. Criterion-related validity - There are two types both of which are concerned with the relationship between the test and external criteria:
a) Predictive - This is concerned with whether the task/test can predict performance on another educational activity?
b) Concurrent - This is concerned with whether the task/test will produce results which are generalisable to another assessment made at the same time?
3. Construct validity - This is concerned with whether the task/test measures what it claims to measure in terms of the underlying concepts pertaining to the teaching, learning and assessing activity?
4. Face validity - 'On the face of it', does the task/test have validity, is it accepted by the profession as assessing fairly what they perceive the curriculum to be?
Content and face validity are best addressed by expert scrutiny. Criterion-related validity was not appropriate, given the novelty of National Curriculum assessment and the lack of criterion or standards against which to measure the validity of the

assessments. Similarly there was no existing data relating to relevant assessments to assess concurrent validity. The appropriateness of construct validity would depend on the structure of the assessment procedures.

Reliability concerns the consistency of assessment. It should be noted, however, that in any assessment process a degree of inconsistency is inevitable, given the numerous sources of variance in pupil performance, both internal and external to the child, and the interpretation of the procedure and assessment criteria by the teacher or marker. Establishing reliability therefore entails minimising these effects. In criterion-referenced assessment this requires the consistent placing of pupils in mastery or non-mastery categories. Three types of reliability can be identified as approaches which could be used to review assessment tests/tasks.

1. Internal consistency - has the marking been consistent? This requires the correlation of components, such as halves, of a test or task but it is problematic if the task/test is open-ended and does not have an item based structure.
2. Equivalence - do different tasks/tests of the same attribute produce comparable scores from the same population?
3. Stability - to what extent are total scores reproduced if pupils are tested on two separate occasions using the same task/test (test/re-test). Learning effects will obviously contaminate this approach to reliability. However it is also possible to examine the stability of a task/test from the perspective of the marker(s) rather than the student. Of these approaches re-mark activities offer a sound practical approach in reviewing both internal consistency and stability. Equivalence was a key issue if schools and possible students were to be given a degree of choice.

Validity and reliability in relation to design and technology

The 1990 National Curriculum in design and technology consisted of four attainment targets. The attainment targets are provided in full in appendix 2.2 For purposes of reference and so that every attainment target and statement of attainment is uniquely identified, design and technology attainment targets are referred to as Te1, Te2, Te3 and Te4. To identify a specific statement the attainment target is followed by the statement label. For example, Te1.3b is attainment target 1, level 3, statement b. The four attainment targets might be interpreted as the four component parts of any design and technology activity. Their titles were as follows:

Te 1	Identifying needs and opportunities;
Te 2	Generating a design;
Te 3	Planning and making;
Te 4	Evaluating.

As design and technology capability constitutes all four attainment targets important issues affecting validity are immediately apparent. Would it, for example, be valid to assess each attainment target independently and aggregate the resulting scores; are the attainment targets of equal value; do all four need to be assessed; how could the assessment of statements of attainment, or the individual items which they contain, be accurately undertaken; which was more important - the assessment of the whole capability; or the assessment of the part - the item within a statement?

Design and technology capability is frequently referred to as being holistic in character - the whole being greater than the sum of the parts. A commonly used analogy is that it is more difficult to juggle with four balls simultaneously than with each ball in turn. A survey undertaken by the National Centre for School Technology (NCST) in 1983 [20], claimed that:
'Technological capability is the capacity to take action to master the physical world and increase the quality of life by employing the problem-solving skills, certain knowledge about energy, materials and methods of control, and the ability to make value judgements.'
This survey was influential in the establishment of the Assessment of Performance Unit (APU) in design and technology. The Unit asserted that from the antecedents, such as the NCST survey, design and technology was an holistic capability since, 'design and technological activity requires an appropriate interaction of skills, knowledge and values and cannot appropriately be assessed by any process of merely aggregating discrete levels of performance.'[21]
The final report of the APU unit, published in June 1991 - two years after the publication of the technology statutory Order, believed it had established a prima facie case for assessment to be holistic. This approach contrasted sharply with the framework, itemised and specific, established by TGAT. The subject perspective to assessment was, therefore, contradictory to the approach adopted by the National Curriculum. To be valid tasks had to assess the statements of the National Curriculum yet the APU evidence pointed to assessments being significantly more reliable when based on an holistic overview than on individual elements.

The statutory Order in technology did not articulate a clear body of knowledge; what a pupil would need to know in order to demonstrate capability at a specific level. The acquisition of knowledge is the most straightforward attribute to assess especially if knowledge is defined simply as facts, which in itself is contentious. Assessing factual acquisition merely requires a recall test. Cognitive skills were implicit in many of the statements of attainment. At level 4 for example, the average level of attainment for pupils entering key stage 3, pupils have to be able to:

identify needs and opportunities, devise ways of, recognise points of view, explain, justify, estimate resources, adapt procedures, choose, adopt alternatives, review decisions, comment on, understand the implications.
All these require more than just knowing, they require understanding. One can argue that understanding is a higher order state of development than knowing, so it follows that understanding subsumes knowledge. If this is accepted, assessment of these characteristics should focus on whether the pupil could demonstrate a conceptual grasp of a level as defined by the statements at that level.
Consequently assessment should focus on the ability to apply the things which a pupil has been taught as required by the programme of study rather than just recall them.

Issues relating to criterion-referencing

- *Is it possible to devise and implement assessment procedures, based on criterion-referencing, which will differentiate pupils' achievement across the range of design and technology in a fair yet consistent fashion?*

This question is also clearly involved with matters of validity and reliability but it is more concerned with the range of subjects and materials which were prescribed by the National Curriculum as coming within the orbit of design and technology. The Order[22] brought together the following departments in order to deliver design and technology: art and design; business studies; craft, design and technology; home economics; information technology and textiles and required that at each key stage pupils should be given the opportunity,
'to work with a range of materials, including textiles, graphic media (such as paint, paper, photographs), construction materials (such as clay, wood, plastic, metal) and food.'[23]
Consequently, the assessment criteria, both the attainment targets and the statements of attainment, had to be equally applicable and relevant in the variety of environments and contexts which would result from these requirements.

To devise a national assessment system which was formative at every level and scale, TGAT decided to adopt a criterion-referenced approach. Monitoring changes in national standards requires knowing what the average child has actually achieved. A norm-referenced approach would conceal changes in national standards. Whatever the average child accomplishes is the norm and if the average child's performance changes the reported norm remains the same. Only criterion-referencing can monitor and communicate changes in standards. There were and are persuasive educational reasons for adopting a criterion referenced approach. Norm-referencing is comparative and discriminative in nature and

therefore provides the clearest information for selecting and sorting. Norm-referenced tests are designed to spread performance as widely as possible and are traditionally associated with fixed quota selection.

Criterion-referencing confirms achievement as it evaluates performance against an established criterion of achievement. It indicates a level of mastery and accurately identifies each pupil's strengths and weaknesses. It is poor as a selection tool because the aim of the teaching is to achieve mastery. It is sensible to assume that an individual should not be assessed against criteria which they had not been taught to evidence. The beneficial aspect of criterion-referencing from the learner's perspective are also important. Assessment is focused on achievement and clear, progressive objectives provide each individual with achievable goals and success relevant to their own level of ability. Whether criterion-referencing could be applied to a national assessment system would be a crucial factor in the successful implementation of the curriculum. More particularly could the statements in design and technology be applied reliably across the subject contexts and material specific environments, this would test the concept of the TGAT model to the full.

Differentiation is in many respects a novel issue in design and technology. Whether teachers understood it as such, the conventional approach employed by the majority was differentiation by outcome. This is an approach in which no attempt is made to match the difficulty of a task or activity to individual differences but where the pupils being assessed are presented with the same task and different degrees of success occur and can be marked. Degrees of success can be easily marked and translated into rank orders for purposes of comparison or sorting, but whether this approach would be suitable to a criterion-referenced regime would need to be explored. Differentiation by outcome might prove to be haphazard, it would be largely left to chance for a pupil to produce the appropriate evidence to record success against a specific task. Specific tasks, each focused on a particular statement of attainment, clearly provide the most appropriate evidence. However, would dissecting the activity into a series of discrete tasks or items and then aggregating the results produce an assessment of design and technology capability?

The argument for assessment by outcome in design and technology is persuasive, given the process nature of the subject. Attainment target 1, for example, requires pupils to identify needs and opportunities, decide what is worth doing, how they will set about the task and what will be the implications for the decisions which they take. Consequently, the teacher, whose main priority might be assessment, might not even be able to determine the pupil's task, if it is these aspects which he or she is seeking to assess. The pupil's outcome might as a result offer little appropriate

evidence against the other attainment targets, even though they have been involved in a purposeful and rewarding task. This issue was most acute at key stage 3, as only at this key stage were standard attainment tasks required to cover all ten levels. So could a single task provide differentiation across such a wide range of achievement and would such a task offer these pupils the opportunity to evidence their best level of achievement? Which approach would be the most reliable and valid, differentiation by task, by outcome or by some form of combination, and would the most appropriate mode be politically acceptable?

The Impact of assessment activities on learning experiences

- *Can an assessment procedure be devised which will meet the political aims of National Curriculum assessment whilst providing pupils with valid learning experiences and pupils, teachers and parents with accurate and useful measurements of performance?*

Post-war governments have made various attempts to tackle the issue of standards in education. These attempts were based on the universal recognition that the growing demands of a technological age required a more skilful and educated population. Successive Secretaries of State for Education had, however, left the curriculum in the hands of the professional educators, regarding it as an aspect which was taboo to politicians. The decision to change this unwritten policy was taken by the Prime Minister, James Callaghan in 1976 in a speech at Ruskin College Oxford. In the speech, subsequently published as a Yellow Paper, he referred to low standards and education's failure to meet the needs of the industrial and commercial community. By signalling an end to what he termed 'keeping off the grass', the Prime Minister had no doubts that he was responding to growing public concern:

'I take it that no one claims exclusive rights in this field. Public interest is strong and legitimate..... Parents, teachers, learned and professional bodies, representatives of higher education and both sides of industry, together with the government, all have an important part to play in formulating and expressing the purpose of education and the standards that we need.'[24]

In addition to advocating a core curriculum the Yellow Paper also articulated the need for:

'adequate means of obtaining information about pupils' performance in schools.'[25]

Concern over falling standards in education increased and reflected a much wider one of both industrial and national decline. The Conservative administration headed by Margaret Thatcher, elected in 1979, gradually adopted a greater degree of intervention in relation to the curriculum, culminating in the Education Reform Act. This administration was, unlike previous Conservative Governments:

'.....radical, technical and pragmatic in its orientation, sufficiently divorced from the older values.....The Prime Minister and her closest associates did not come from the old school. Their origins made them better able to implement reformist policies aimed at restructuring British Industry, to a degree at least through education and training.'[26]

Assessment, and the reporting of the results, was key to the Government's drive to raise standards. The need to quantify performance at every level, pupil, teaching group, school, LEA and national, was essential both as a quality assurance measure and as a means of justifying expenditure. During a period when public finances were increasingly placed under severe scrutiny, value for money exercises could only be undertaken given reliable statistics as the basis for comparison.

'We all want to see money put to good use. Above all we want to see the concrete results of improved standards. The best way to justify these great increases in investment by the taxpayer is for teachers to carry out tests so we can see improved results.'[27]

Government clearly required assessment which would be nomethic in character, that is the assessment of individuals with a view to making comparisons among them or to generalise to larger populations. This was fundamentally opposed to the notion of idiographic assessment, assessment designed to understand the individual in their own right without comparison with others. This is an approach to assessment which in the view of the majority of teachers is at the heart of the teaching and learning process. Nomethic assessment seeks to categorise individuals, and unfortunately all too frequently this results in individuals being neglected by being classified as 'slow learners', 'underachievers' or 'behaviour problems'.

This type of classification can also be favourable, 'high-fliers', or 'A streamers'. This type of labelling focuses on similarities, what individuals have in common. The assessment procedure devised by TGAT was conducive to this mode of generalisation. For example, in the future a key stage 3 teacher may have substituted, 'level twoers and threers' for underachievers and 'level eighters and niners' for A streamers! Teachers have always been opposed to such labelling as they would claim that it is all too frequently self-fulfilling, making the learner other or lesser than he or she might have been. Nomethic assessment is closely linked to summative and norm referenced assessment. However, it is only by assessing and classifying that statistics can be produced which would allow a government to, for example:

- detect objectively a rise in standards/performance;
- correlate a rise in standards/performance with the enrichment of resources;
- determine the effect on standards/performance of different teaching strategies.

Nomethic assessment can be an extremely powerful tool in detecting the cause for an effect. Nomethic predictions are similar to the actuarial approach of insurance companies. If, for instance, the consistent national average for design and technology capability, at the end of key stage 3, was level 5.5 with the exception of one year, the cause of that effect could be explored. If in that year a government initiative had provided INSET for two teachers from every school and the average level rose by 0.2 of a level. It would be reasonable to conclude that the enhancement of teacher quality was the cause. The cost of raising standards could be determined and planning decisions taken as a consequence. But as Rowntree (1977) observes:

'We must not jump to the assumption that what is 'actuarialy' true for the group will also be 'clinically' true for the individual.' [28]

In the year in which the average rose there would undoubtedly have been many pupils whose relative standard fell. That is why, fundamentally, educationalists mistrust assessment focused on producing information which will be used to generalise, as their concern is with the individual and detecting differences rather than similarities.

Idiographic, formative, assessment is the only approach which is truly conducive to the educational development and progress of an individual. In principle criterion referenced assessment, the foundation of national curriculum assessment, is idiographic as it seeks to determine what the individual learner can do and can not do. This provides sound information for informing future learning. The key question for every teacher would revolve around the educational relevance of an assessment procedure. If it appeared that the principal purpose was to gather information for making general, global observations then these assessments would be seen as inappropriate and imposed. For example, if the generalisation of performance was the prime motive all pupils would need to be assessed at the same time. However, from the perspective of the individual this might be quite inappropriate. In the case of design and technology, relevance would also relate to the mode of assessment. It is only relevant to assess a practical purposeful subject by a practical purposeful task. Traditionally, assessment was only respectable if it involved examinations. Course work and practical tasks were viewed with suspicion, especially by those for whom they had not been part of their own educational experience. As development got under way opinions polarised. For politicians, examinations, pencil and paper tests, were seen as straightforward, cheap and providing the right kind of information. For teachers the relevance of information gained from such a test to design and technology capability would be more crucial.

The relationship between different modes of assessment

- *To investigate the relationship between end of key stage assessment, terminal and summative in nature, in relation to ongoing teacher assessment, continuous and formative in nature.*

As previously described, the TGAT report clearly articulated two distinctive types of assessment, end of key stage assessment and teacher assessment. In traditional assessment terminology the first is terminal and the second continuous. Continuous assessment is an essential requirement of any programme of learning. It enables the teacher to reflect on the success of the learning process and adjust teaching tactics according to the development of the pupil(s). In making these judgements a teacher can draw on all the evidence at his or her disposal. Assessment is frequently categorised as either permanent or ephemeral. In the context of design and technology permanent evidence usually relates either to the process - a design folder or the product - the final solution in what ever form it might take. Such evidence can be stored, seen by other assessors and agreement can be reached on what the evidence is worth. Ephemeral evidence exhibits none of these characteristics. It might concern conversations which a teacher has had with a pupil or observations on how a pupil tackles a task. Such evidence can be extremely revealing and provide genuine insight into a pupil's capability. However, such evidence often only exists in the memory of the teacher, its collection and retention may have been partial and spasmodic, it cannot be moderated and standardised against similar evidence from other pupils. Permanent evidence can also be fallible. It might be tokenistic, the recording of events and decisions after they have taken place, not because they are of value in completing the task successfully but because they have assessment value. The recording of an event is not always the best evidence although it might be more acceptable as summative evidence.

A good example relates to safety, an issue covered by several statements in Te 3 in the technology Order. Operational safety is of course the key issue. The evidence in a pupil's folder might be that:
'The best way to cut out the shape is to drill a lot of holes using the drilling machine. When using the drill I must follow the safety rules.'
This is evidence of an awareness of the issues not confirmation that the pupil carried out the task in a safe manner. It is only the teacher who observes the pupil carrying out the task who can confirm that the pupil can:
'....demonstrate by their choice and use of a variety of equipment that they understand the principles upon which these work and the requirements of safety and accuracy.' Te3.5c

There are other instances where the teacher experiences the pupil's performance but cannot preserve it, the assessment must take place in real time and the judgements made must be taken on trust.

Evidence, both permanent and ephemeral, provides the teacher with a rich picture of a pupil's capability accumulated over a period of time. In the case of key stage 3 - three years. Is it possible to provide a comparable terminal end of key stage assessment task? Will it not, because of the constraints, appear arid in comparison? It could be argued that the view of the teacher has become too subjective and that only an objective external assessment can reveal actual achievement. So regardless of the nature of the task, it provides a calibrating and standardising tool not subject to local distortions. The crux of this issue is which of the two modes of assessment has priority, which is the one that really counts? Does the summative level ultimately accorded to the pupil depend purely on performance over a short period of time at the end of a course or are previous on-course performances taken into the reckoning also? Sudden death play-off or cumulative record?

Assessment is all too frequently viewed only from the perspective of the information required by the person carrying out the assessment. The views of those being assessed must also be considered. It is generally claimed that continual assessment is less stressful than a final examination. This must be balanced by the view that continuous assessment also creates stress as the pupil might believe that they are under constant surveillance. Every action whether positive or negative will be used as assessment evidence. This effect was captured by a student writing to the Times Educational Supplement who compared continuous assessment with examinations as:
'Comparing months of nagging toothache with the short sharp pain of having a tooth removed.'[29]
Even if students get used to the regular stress of continual assessment, the question must be asked, does it effect their attitude to learning? The answer is almost certainly, yes. There is good evidence that only those tasks with an assessment weighting will be taken seriously. For example, Platt (1972)[30] accounts how undergraduate students in their final year concentrated only on the assessed assignments and neglected non-assessed coursework.

Perhaps the fairest assessment procedure from the pupil's perspective would be only to record the assessment for the best piece of work from a course or period of study, regardless of whether it occurs during the period of study or at the end. In a system using criterion referencing this should be the only acceptable approach. Once a statement has been achieved it cannot be 'unachieved', even though a

pupil may fail to repeat this level of capability in a subsequent task. This failure may be due to other issues such as degree of motivation or external factors. A system which has a policy of 'best work only to count' should create the least possible stress for the pupil. However it would create a situation in which end of key stage assessment would only count if it improved on continuous assessment results. Given the importance placed on summative assessment it is difficult to see this approach being adopted; if for no other than reasons than cost-effectiveness. An assessment system which aggregates or combines criterion referenced performance, or only recognises levels gained in specified contexts must have motives other than observing the best achievement of each individual. Because such procedures miss the principal aim of a criterion referenced approach - to record what a pupil can do.

The TGAT report discussed the issue of which assessment mode should take priority and recommended:
'....that the national assessment system is based on a combination of moderated teacher's ratings and standardised assessment tasks.'[31]
This recommendation, by requiring teacher assessments to be moderated, recognised that individual judgements would need to be brought in line with general standards. As schools' catchment areas are not representative of the national population, normal or average spread of achievement in one school will not be the same as in another. Consequently, it argued, there will be a need to reconcile teacher judgements to reach an agreed standard. The pupils might reasonably complain of assessment overload, being subjected to both continual and terminal assessment. Indeed teachers might also complain that their task had been distorted given the high priority accorded to the assessment of teachers.

The question of manageability although recognised at the outset became more important as the first statutory assessment approached. Statutory assessment would cast teachers in the role of assessor, impartial, detached and objective. Assessment systems and procedures would need to be implemented successfully by a wide variety of teachers in a range of environments and the activity, by its nature and origin, would be highly contentious and subject to intense scrutiny and criticism. Any approach which was slightly unrealistic, over demanding, appeared repetitious or superfluous, or in any minor way educationally questionable would not succeed.

Support and advice

In seeking answers to these questions the author of this thesis received support and funding from the Schools Examination and Assessment Council. Following a

national competitive tendering procedure and in liaison with the Midland Examining Group the author was joint director of the agency commissioned to carry out the necessary research and development to produce statutory tasks in technology at key stage 3. This contract, worth £1.9 million was initially for a three year period - September 1989 to August 1992. However, it was terminated by SEAC at the end of two years. Following another round of competitive national tendering the author was awarded a further contract for two years as the Director of an independent agency based at Middlesex University. The second contract had a value of approximately £700,000. Without this level of financial support and consequential leading role in the implementation of the National Curriculum this level of research and cooperation would not have proved possible. The work in this thesis draws only upon the aspects of the research for which the author was responsible, many other issues were dealt with by the agency but they are not the subject of this thesis. However the author readily acknowledges the cooperative nature of the exercise and the support, both practical and intellectual, which others members of the team provided.

References

1. Department of Education and Science. *The Education Reform Act 1988: The School Curriculum*, Section III.1. London: HMSO
2. Graham, D. and Tytler, D (1993). *A lesson for us all.* London: Routledge, page 2
3. Department of Education and Science (1989). *The Education Reform Act 1988: The School Curriculum and Assessment.* Circular No. 5/89. London: HMSO
4. Department of Education and Science. *The Education Reform Act 1988: The School Curriculum*, Section III.1. London: HMSO
5. Department of Education and Science. *The Education Reform Act 1988: The School Curriculum*, Section III.1. London: HMSO
6. Department of Education and Science. *The Education Reform Act 1988: The School Curriculum*, Section III.1. London: HMSO
7. Department of Education and Science. (1988.) *Task Group on Assessment and Testing - Report.* London: Department of Education and Science
8. Department of Education and Science. (1988). *Task Group on Assessment and Testing - Report.* London: Department of Education and Science, para, 5
9. Department of Education and Science. (1988). *Task Group on Assessment and Testing - Report.* London: Department of Education and Science, para, 12
10. Statement by The Secretary of State for Education (7th June 1988). London: *Hansard*
11. Statement by The Secretary of State for Education (7th June 1988). London: *Hansard*
12. Department of Education and Science. (1988). *Task Group on Assessment and Testing - Report.* London: Department of Education and Science, para, 41
13. Department of Education and Science, (1988). *Task Group on Assessment and Testing - Report.* London: Department of Education and Science, para. 45
14. Department of Education and Science, (1988). *Task Group on Assessment and Testing- Report.* London: Department of Education and Science, para, 45
15. Department of Education and Science. (1988). *Task Group on Assessment and Testing - Report.* London: Department of Education and Science, para, 109
16. Department of Education and Science. (1988). *Task Group on Assessment and Testing - Report.* London: Department of Education and Science, para, 108
17. Department of Education and Science. (1989). *Policy into Practice.* London: HMSO, Para. 3 -11
18. Jessop, G. (1991). *Outcomes: NVQs and the emerging model of education and training.* London: Falmer, pages 123 to 127
19. Schools Curriculum and Assessment Authority (1994). *GCSE Examination Results 1993.* London: School Curriculum and Assessment Authority
20. National Centre for Schools Technology. (1983). *Understanding school technology.* Loughborough: NCST, part 1, page 2
21. Schools Examination and Assessment Unit. (1991). *The Assessment of performance in design and technology.* London: Department of Education and Science, page 13

22. Department of Education and Science/Welsh Office. (1990). *Technology in the National Curriculum*. London: HMSO

23. Department of Education and Science/Welsh Office. (1990). *Technology in the National Curriculum*. London: HMSO, page 19

24. Callaghan, J. (1976). *Towards a National Debate*. Speech at Ruskin College, Oxford

25 Callaghan, J. (1976). *Towards a National Debate*. Speech at Ruskin College. Oxford

26. Bernbaum, G & Mathieson, M. (1988). 'The British Disease: A British Tradition?' *British Journal of Educational Studies* Vol. XXVI No. 2

27. Portillo, M. (1993). Speech to The Hertford Conservative Association.

28. Rowntree, D. (1987). *Assessing Students: How shall we know them?*. London: Kogan Page, page 160

29. Cowell, D. (1972). Letter to *The Times Educational Supplement*: London

30. Platt, J. (1972). *Report of Research into the Operation of the new pattern of BA Final Assessments*. Lewes: University of Sussex

31. Department of Education and Science. (1988). *Task Group on Assessment and Testing - Report*. London: Department of Education and Science, para, 63

Chapter 2

Towards a National Curriculum in design and technology

Synopsis

This chapter reports the evolution of design and technology by tracing the origins of the principal influences on the 1990 Statutory Order. The first part describes the way in which design and technology has become more responsive to the changes taking place in society. It charts how these changes, along with a belated re-evaluation by society of the importance of 'practical learning', have helped to enhance the status of the subject.

The influences from the mid-sixties onwards are examined in detail to assess the impact which they had on the first Statutory Order for the subject. The process by which this Order was developed is charted critically in an attempt to shed light on its final content and structure.

The second part of the chapter details the developments which have taken place in the manner and purpose of the assessment of design and technology. The gradual shift of the assessment focus from the product to the process is detailed along with the parallel shift from norm- referencing to criterion-referencing. It describes how the adoption of both these changes culminated in the 1990 Design and Technology Order.

The chapter is divided into the following sections and sub-sections

The making of a subject

- Responding to social change
- The rise of the subject
- Design and technology comes of age
- The promotion of design and technology
- The vocational movement of the eighties
- The influence of the Assessment of Performance Project
- Defining a National Curriculum subject

The assessment of design and technology

- From marking to assessment
- The recognition of project work and course work
- A common examination for all
- National Curriculum assessment - criterion referencing

The Making of a subject: Responding to social change

Design and technology is one subject in the school curriculum which must respond to the changes which occur in the nature and values of society. Indeed the subject's rise is a reflection of changing social values throughout the past one hundred years. Subject practitioners have constantly sought to raise the profile of practical education but little would have been achieved had not society reassessed its view of the value and importance of learning through doing. The content of the curriculum is a reflection of the value which society places on certain types of knowledge and learning. As Young (1971) has suggested:

'Curricula in this country involve assumptions that some kinds and areas of knowledge are much more 'worth-while' than others..... Further, that as we assume some patterns of social relations associated with any curriculum, these changes will be resisted insofar as they are perceived to undermine the values, relative power and privileges of the dominant groups involved.' [1]

Design and technology, with its emphasis on learning through practical experience, has had to overcome the traditional resistance to non-academic subjects which is prevalent amongst those who have the power to make changes.

Young continues by describing how status and hierarchy is established within the curriculum:

'The contemporary British Educational System is dominated by academic curricula with a rigid stratification of knowledge. It follows that if teachers and children are socialised within an institutionalised structure which legitimates such assumptions, then for teachers high status (and rewards) will be associated with areas of the curriculum that are (1) formally assessed (2) taught to the ablest children (3) taught in homogeneous ability groups of children who show themselves most successful within such curricula.'[2]

Design and technology has, as yet, not achieved this level of recognition of academic legitimacy and only if the National Curriculum had been fully implemented would it have done so.

The National Curriculum represents the most extreme centralisation of control over the curriculum yet attempted in this country. One frequently stated reason for the taking of such control was to provide a curriculum which would be more in tune with the nation's economic needs. If this is the case then technological knowledge should now be accorded high status because as Apple (1978) comments:

' the corporate economy requires the production of high levels of technical knowledge to keep the economic apparatus running effectively and to become more sophisticated in the maximisation of economic expansion.'[3]

In the United Kingdom high status has always been bestowed on academic rather than technical education. It is doubtful even now that technical education has

achieved comparable status but the recognition of its importance has forced the subject to surface rapidly from the backwaters of the curriculum.
'Thus we have seen the advancement of technological subjects from their low status in the hierarchy of school subjects to fourth place in the National Curriculum league table.' (Blenkin et. al. 1992)[4]
Even if true, however, the promotion might have been too rapid and too radical.

The traditional concept of how a subject becomes part of the school curriculum has been explained by Goodson (1988) in the following terms:
'Once a discipline has established an academic base it is persuasively self-fulfilling to argue that here is a field of knowledge from which an academic school subject can receive inputs and a general direction.'[5]
It is apparent that design and technology does not correlate closely with this concept. There was certainly no academic base on which the subject was established and it is doubtful if there was, or is, a field of knowledge which is the specific domain of design and technologists. If there is an academic base to the subject it is primarily concerned with teacher education, undertaken by those who have been deemed to be 'good practitioners' of the subject at school level. The requirement for teachers of the subject is, consequently, the sole reason for the subject's establishment in higher education. If those in higher education focus their energies only on supplying the school demand for teachers there will be a natural tendency to introversion and self-perpetuation. This will result in a failure to promote an objective and critical appraisal of the subject's parameters and potential. But design and technology may not be alone in this. Even though most subjects in the school curriculum are held by society to be disciplines in their own right, in the view of Jenkins and Shipman (1976) many have characteristics in common with design and technology:
'Many school subjects are barely disciplines let alone forms of thought. Many are unclear about their most fruitful concepts, forms of explanations and characteristic methodology.'[6]
It is not untypical then for a school subject to be divorced from its originating discipline, if indeed it has one, to be autonomous and to establish a subject community in its own right. Design and technology, due to the low esteem in which it was traditionally held by society, has not encouraged self-appraisal. This has resulted in a subject community which, with notable exceptions, did not seek to establish a conceptual understanding of its fundamental principles and practice.

There are several models which attempt to explain how a subject becomes established in the school curriculum. The most pertinent to this discussion relates to science, a subject which also has a practical dimension, and was conceived by Layton (1972) a member of the National Curriculum Working Group for technology.

He defined three stages of development:
First stage - The callow intruder stakes a place in the timetable, justifying its presence on grounds such as pertinence and utility. During this stage learners are attracted to the subject because of its bearing on matters of concern to them. The teachers are rarely trained specialists, but bring the missionary enthusiasm of pioneers to the task. The dominant criterion is relevance to the needs and interests of the learners.
Second stage - A tradition of scholarly work in the subject is emerging along with a corps of trained specialists from which teachers may be recruited. Students are still attracted to the study, but as many by its reputation and growing academic status as by its relevance to their own problems and concerns. The internal logic and discipline of the subject is becoming increasingly influential on the selection and organisation of the subject matter.
Third stage - The teachers now constitute a professional body with established rules and values. The selection of subject matter is determined in large measure by the judgements and practices of the specialist scholars who lead inquiries in the field, students are initiated into a tradition, their attitudes approaching passivity and resignation, a prelude to disenchantment.[7]

Layton's model is linear and has much in common with traditional developmental models, such as those relating to landscape formation, which identified development in terms of youthful, mature and old age. The emergence of design and technology does not appear to correlate exactly with this model, because, despite a hundred years of evolution in these terms it has not reached the seasoned confidence of the third stage. The correlation between the youthful or first stage is apparent and in many aspects the subject still exhibits these features. Possibly, this is because the subject is not rooted in a discipline which draws on a defined body of knowledge. Consequently, major developments which have attempted to change substantially the subject's identity and nature are always accompanied by a 'missionary enthusiasm' and 'pioneering spirit'. Experience also indicates that, 'learners are still attracted to the subject because of its bearing on matters of concern to them.' In contrast to Layton's model, design and technology is best represented by a cyclical pattern, one which involves a process of re-evaluation and rejuvenation as the subject periodically appraises both its content and practice. A dynamic process in which the subject attempts to adapt to varying circumstances and to respond to changing pressures. Design and technology in this sense has some of the characteristics of a product rather than a discipline. Each manifestation of the subject only has a certain life cycle and to stay relevant and appropriate the subject needs to be continually interpreting the technological world in forms which will be appropriate to school pupils.

Some aspects of Layton's second and third stage are still to be achieved. It is doubtful if students have ever been attracted to the subject because of its status and only now is a professional body emerging which might have the equivalent influence to, for example, the Association for Science Education (ASE). Committed practitioners have consistently tried, via the reality of relevant curriculum practice, to keep the subject in touch with the rapid changes in society but the subject has been perceived to be slow to change. Design and technology has reached its current status not just because of the aspirations of the subject's professional community but also because perhaps, uniquely, society has changed its view of the subject's potential, even though this potential has yet to be fully realised.

The rise of the subject

Evidence of this heightened value was the inclusion of design and technology as a foundation subject in the National Curriculum. The subject's unique capability to contribute to every child's education has finally been recognised. For the first time it is a subject which all pupils must follow for the eleven years of compulsory education. This will permanently change the traditional perception of the subject, those who take it and hopefully those who teach it. As Penfold (1988) comments: 'It is now part of educational folklore that those amongst the shavings and the swarf were the academically less able boys - the 'thickies'. Having failed at everything else, they were deemed to be 'good with their hands.'[8]
Every educational report relating to the curriculum, such as Hadow, Crowther, Newsom et al. [9], has extolled the benefits of practical subjects for all pupils but the view persisted that it was a subject for low achieving boys. It was not until the Education Reform Act of 1988 [10] that it became a compulsory element of every child's education from 5 to 16. Legislation has brought about what many have advocated, but it will not, alone, convince the sceptics of the subject's value.

The Statutory Order in technology was the culmination of a century of consistent evolution. With the introduction of this Order the subject should have been poised to embark on a period of consolidation and refinement. The reality has proved to be very different. The subject, in accepting its place in the mainstream of education, has had to recognise the rules that pertained to this elevation. As Kimbell (1992) commented:
' But the fruits of this success became clear when it (design and technology) appeared in the extended core of the National Curriculum policy documents. Someone was going to have to decide what it was that we were all doing so successfully. It had to be tamed and institutionalised.'[11]
The subject had, for the first time, to be clearly defined and needed to establish a relevant and appropriate curriculum for both boys and girls for the eleven years of

compulsory education. In addition the subject had to accept the regulations of the National Curriculum and specifically all that entailed in relation to assessment and reporting.

Those charged with the task of proposing subject Orders, in addition to their own specific background and philosophy of the subject, were possibly influenced by the direction and developments in the subject, current at the time. During the formative period of this work developments were dominated by the Assessment of Performance (APU) Project based at Goldsmiths College.[12] This project evolved a specific model of design and technology and had direct roots in one of the three traditional influences which have shaped the subject. Each of these three perspectives has at varying times been in the ascendency and provided the catalyst for change. The first of these influences and the prime reason for the introduction of manual training in the 1890s, is focused on the vocational potential of the subject. The second seeks to enhance the subject's links with scientific principles and their application in a practical context and the third recognises the benefits of the subject as an enriching educational medium, offering schoolchildren the opportunity to grapple with problems and take decisions for which they will be accountable. This was the philosophy underlying the work of the APU.

The subject design and technology, as defined by the 1990 Order, has emerged in the past thirty years from the traditional handicraft element of the curriculum. Handicraft was dominated by an emphasis on skill acquisition, an approach which frustrated an increasing number of teachers. They could see that this bias to the mastery of skills was no longer relevant to the needs and did not satisfy the aspirations of many pupils and students. Design and technology evolved due to the effort and enthusiasm of a minority of teachers. These teachers tried out ideas and approaches, often in a rather arbitrary fashion, but gradually independently, and jointly, they distilled the essence of the subject.
' The subject grew from practice rather than theory; from teachers in classrooms trying out innovative and often idiosyncratic activities and programmes rather than an academic analysis of a field of knowledge.' Kimbell (1992)[13]
The new approach was characterised by offering pupils the opportunity to design the objects which they were to make. This gave pupils the opportunity to translate their ideas into real things and consequently they started making decisions about what they would do and how it would be achieved.

Design and technology, like all school subjects, must be responsive to, and reflect, the fundamental effects which technology is having and will continue to have on society. Therefore, it is difficult to envisage a time when design and technology practitioners and others will not be evaluating the subject's role and purpose.

Historically this has been the case and it is likely to suggest that this will continue to be so in the future. Even by the end of the nineteenth century debate was taking place concerning the most purposeful approach to the subject. Initially, the approach to acquiring practical competence had been via prescribed exercises. Circular 44 issued by The Science and Art Department in August, 1890 stated:
'... the course be adapted to teach the use of tools employed in handicrafts, not so much as an initiation in a special handicraft, but as a disciplinary educational course to train the hand and eye to accuracy by a progressive series of exercises connected with free-hand drawings and drawing to scale.' [14]
However, this approach was questioned, by amongst others Saloman (1894) one of the pioneers of the 'Sloyd' approach in Scandinavia. He believed that personal motivation was essential if effective learning was to take place:
'by making practical useful articles and not abstract preparatory exercises, the pupil gains his skill and experience.'[15]
Saloman believed that children should be confronted with situations which required them to exercise judgement and make decisions; technical achievements, in his view, were secondary to those of an educational nature. This basic debate revolving around the relative importance of the educative process in relation to the acquisition of knowledge and the quality of the product has been a continuous theme throughout the subject's evolution and is key to the debate concerning the nature of design and technology in the National Curriculum.

This debate has not hindered the subject's development, possibly it is reasonable to argue that this dichotomy, as to whether the subject is primarily knowledge or process based, has provided one of the main stimuli to the subject's progression to acceptability. Indeed, without a genuine appetite for reform the subject could not, in little over a hundred years, have evolved from manual training in the playground shed of a Paddington school to a compulsory element of every child's education between the ages of 5 and 16. The rise of the subject, from its inception, has been well documented by amongst others Dodd (1978)[16] and Penfold (1988)[17] They and others have acknowledged that it has not always been a smooth route to acceptability. For example, Eggleston (1985) commented:
'After its (craft, design & technology) years in the outer reaches of the curriculum - and often the school buildings - it has painfully fought its way out of the cold. CDT teachers, advisers, inspectors and teacher trainers have devoted decades of committed work to the struggle.'[18]

Design and technology comes of age

The turning point in the subject's fortunes can be traced to the two national projects initiated in the 1960s. Both were of paramount importance. These projects were

established with funding from the Schools Council, one, more concerned with content and the acquisition of specialist skills and the other committed to wider educational issues. Each, however, can be seen to have initiated crucial aspects of design and technology as defined in the Statutory Order of 1989.

Project Technology was established in September 1966 and ran for five years under the directorship of Professor G. H. Harrison at Loughborough University. It was initiated by a Schools Council field report which posed eight fundamental questions (Schools Council Field Report 3 Technology in the Schools. 1966)[19], some of which, even now, have not been fully resolved. In concert with this report a pilot study was established in 1966 which set out to answer these questions by investigating, 'the place of technology in the school curriculum.' The answers were provided in The Schools Council Working Paper 18 (1968) which established the aims and objectives of the project. The principal aim of Project Technology was described in this paper as:
'to help all children to get to grips with technology as a major influence in society and as a result, to help them lead effective and satisfying lives, to encourage technological activities in school and thereby develop a range of abilities and provide motives which are often overlooked by more traditional approaches.'[20]

The Design and Craft project began with a feasibility study in April 1967. This resulted in Schools Council Working Paper 26 Education through the use of Materials. This paper articulated the possibilities which change might offer handicraft teachers:
'handicraft teachers, through their technical skills and knowledge and the resources at their disposal, are in a strong position to assist their pupils to develop not only motor skills and craftsmanship, but also many intellectual qualities that hitherto have been associated mainly with other subject areas.'[21]
These qualities, or curriculum objectives, were detailed in the appendix of the report and were categorized into six areas:
motivation to achieve objectives; personality traits and attitudes; creative development; skill; logical processes/strategies and knowledge. These objectives were far wider than those to which the subject had previously dared to aspire.

The paper resulting from the study convinced the Schools Council to establish a Research and Development Project in Design and Craft Education under the directorship of Professor S. J. Eggleston at the University of Keele. The Project's work was described in one of its publications, 'Materials and Design - a fresh approach,' in the following terms:
'At the heart of the Project's work has been the development of problem-solving approaches suitable for use in secondary schools. Students are encouraged to

identify and investigate design problems - in the home, for example, or in the community, in leisure pursuits, or in work situations - to produce and realize solutions and finally to evaluate end products.' [22]
This statement predicts much of the National Curriculum even though it predates it by twenty years.

Both projects fundamentally changed the nature of handicraft/technical studies and in turn craft design and technology and design and technology. These two projects established the foundations on which the subject built in the 1970s and 1980s. However, the outcomes from them did not always naturally integrate in the school context. Eggleston (1992) observed:
'The effect of the two projects was to create two separate reforming movements in the school curriculum. Project Technology sought to introduce a new subject into the curriculum that was additional to what had already been there. The Design and Craft Education Project sought to transform existing subjects to create the new design orientated approach.'[23]

Frequently, either departments or individual teachers exhibited a bias towards one of the two approaches. Project Technology with its leaning towards engineering and the control technologies which underpin design in this sphere, provided obvious links with the Science department, whilst Design and Craft Education placed much greater emphasis on problem-solving in the context of making value judgements, offering opportunities for collaboration throughout the curriculum. Project Technology was more interested in individuals acquiring a technical expertise which would be of value to industry whilst Design and Craft Education was concerned with using the medium to develop problem solving skills, which it advocated were transferable.

Both projects sought to change teachers' attitudes. Project Technology attempted to persuade reluctant science teachers, mainly theoretical in their outlook, to become involved in practical situations whilst encouraging craft teachers, lacking in mathematical skills, to overcome their reluctance to expose these limitations. In contrast the Design and Craft Education project sought to change radically the attitude of the teacher to the pupil:
'The concept of the developing pupil - experiencing, initiating, creating, and taking decisions - must be matched by the concept of the teacher who is effectively sharing these activities with him. The objective of this project is to assist more teachers and thereby more pupils to reach the take-off point for this partnership in education.'[24]

The philosophical difference between the two projects is perhaps best characterised by their approach to the dissemination of outcomes, although the way in which the work of both projects was disseminated may appear remarkably similar. The Design and Craft Education project did not advocate ready made 'suggestions for teachers' whilst in contrast Project Technology produced materials capable of immediate adoption. Both projects spread new approaches and curriculum initiatives via journals, textbooks and pamphlets. These were, in the main written by teachers for teachers. However, it is now apparent that such resources cannot, independently, effect permanent structural changes to the curriculum. As Dodd (1978) observed, in relation to Project Technology:
'Perhaps because of the dearth of literature available to teachers of Technical Studies at the time, the enthusiasm with which they collected the material was commendable. The use they made of it varied considerably, and they were very reluctant to report on their success or failure.'[25]

The promotion of design and technology

The best of the work produced under the auspices of these two projects demonstrated the potential of this element of the school curriculum to educate pupils for a changing world. Consequently, during the 1970s Government and its agencies, such as The Design Council, alongside Industry started to nurture and promote design based work in schools. In 1973 The Department of Education and Science funded a study at the Royal College of Art into, 'Design in General Education' directed by Professor Bruce Archer. This team strove to establish design as the third component of education, alongside physical sciences and the humanities. They defined this third area in the following terms:
'...the collected body of knowledge based upon sensibility, invention, validation and implementation.'[26]

This project produced only limited resource materials and consequently, in practical terms, its effect was rather limited and it was openly criticised by those with a bias towards the technological spectrum of the subject. Also in the mid-1970s the Department of Trade and Industry established The Industry Education Unit to foster links between industry and education[27]. The Unit summed up its objectives as:
'to act as a link between the different organisations active in this field and to spread the ideas being developed in various parts of the country as widely as possible in the hope that they may stimulate others to take action in their own area.'[28]
This unit clearly had a bias towards Project Technology and the promotion of careers in engineering. This was exemplified by The Young Engineer for Britain prize scheme which, with a number of different sponsors, has been held for nearly twenty years.[29]

This was balanced by The Design Council which promoted The Schools Design Prize successfully for a decade. This scheme, sponsored by major industrial companies such as GEC, Rolls Royce and Thorn-EMI, had a bias towards school work which exemplified creativity and ingenuity, often in a social context. Its objective was:

'to encourage the development of creative talents which are educationally important to the children themselves and to the future of British manufacturing industry'[30]

Other companies, notably BP with schemes such as the 'Build a car/hovercraft/ bike', capitalised on design based work in school to achieve a wide-range of objectives. In all these schemes the administrative and promotional costs were far higher than the prize money. (For example, The Schools Design Prize was estimated to have cost in 1980 approximately £45,000 excluding staff costs, when the total prize money was £10,000.) The value of such prize schemes should not, however, be underestimated. They were especially important in bringing to the notice of politicians and other decision makers the changes which were taking place in this element of the curriculum. For example, when in 1982 the then Prime Minister, Margaret Thatcher, presented the winners of the Schools Design Prize with their prizes she observed:

'The value of design education is in encouraging firstly, a professional approach to problem solving and, secondly, a greater awareness of the appearance of products and a more informed consumer demand..... Design education has a vital part to play in the economic future of the nation'[31]

This, amongst other events, helped convince her of the importance of design in economic terms and subsequently she held a seminar at Downing Street in April 1982 specifically on design education. The seminar set out to explore the promotion of design education. The seminar was reported in The Designer:

'We have devoted the whole of the Designer to a discussion of the important initiative taken by Mrs Thatcher in identifying design as one of the crucial areas of development in the regeneration of our industrial life, and instigating an enquiry into the steps that need to be taken.'[32]

In 1978 an HMI discussion document, Curriculum 11-16, argued the case for a common core for the school curriculum. This document expressed the important role technology would have to play if such a curriculum were introduced:

'Any school curriculum seeking to educate 11-16 year-olds must prepare them to operate within this technology-based industrial society....First we need, and will continue to need, people who can operate competently within such a society....Secondly school leavers should be equipped to look at technology critically and to be part of a society that seeks to master technology, not to be enslaved by it.'[33]

The report also commented on some of the issues which would become prominent in the National Curriculum debate:
'Technology is not a given body of facts, nor a unique set of ideas - it is a process with definite applications and origins in time ..... many disciplines can contribute ..... Responsibility for bringing technology into the curriculum ought to be shared by any department which can make a useful contribution in this matter.'[34]
This is a clear statement by HMI of the questioning, process-driven and cross-curricular nature of technology and all of these aspects were to become essential ingredients of the 1990 Order.

These views were supported in July of 1980 by an influential report published by The Design Council. This was the result of a working party, under the chairmanship of Professor David Keith Lucas, on Secondary Education set up by the Education Advisory Committee of The Design Council. This report fully endorsed the notion that design education was an essential element of every child's education:
'there can be few more important educational experiences for children than to grapple with the sort of problems they will meet as adults - problems of the environment, of man-made things and how they can be improved, of the quality of living - or, in other words, 'design' in all its forms. As such, design education is the concern of all boys and girls, not just those who might eventually go on to design-related occupations.'[35]

The vocational movement of the eighties

Clearly by the beginning of the 1980s a consensual view had arisen that design and technology education, as it is now known, had a central role to play in every child's education. However, in the early 1980s, following the election of the Conservative administration of Margaret Thatcher, a radical shift was taking place in economic policy. Monetarism and its effects were to impinge on every facet of Government strategy. Inflation had reached stubbornly high figures and unemployment had risen to 12.5 %. The Government had to be seen to act on a wide front to reverse these trends. Shaw (1991) commented:
'Ideological financial policies were matched by an ideological return to utilitarianism in education.'[36]
Pre-vocationalisim was the educational equivalent of monetarism. Through a range of initiatives, most notably TVEI[37], the Government sought to influence all aspects of the school curriculum and persuade schools to take more responsibility for preparing pupils for the realities of the employment market.

Vocationalism was not a new concept to technology education and this area sought to capitalise on the opportunities which it offered. As noted earlier,

vocationalism was one of the arguments central to the introduction and growth of technical education. Ripper (1886) noted:
'We have to regret the total absence in elementary schools of instruction specifically bearing upon industry, an omission which a manufacturing community would do well to remedy.'[38]
Even though such sentiments were one of the principal reasons for the Technical Instruction Bill of 1889 reaching the statute book the subject, following its introduction, did not foster close links with Industry. Penfold (1988) comments;
'No sooner had the ink dried on the Technical Instruction Act than pressure to align manual instruction with industrial regeneration began to evaporate. Despite the standing of those who promoted the introduction of manual instruction to this end, it was based on a false premise.'[39]

There were consistent attempts to align this aspect of the school curriculum with the nation's needs but few brought effective and permanent change. The Butler Education Act (1944) led to the reorganisation of the secondary sector into a tripartite system; introducing technical schools specifically to develop the skills needed by industry. Commenting on secondary technical schools, Ministry of Education Pamphlet No. 1 (1945) said:
'Industry on the production and manufacturing side has been deprived of its reasonable share of the national talent - a loss which a country so highly industrialised as this cannot afford to carry.'[40]
The technical school concept, except with a few notable exceptions, did not have the desired impact. These schools did not have the prestige of the grammar schools, frequently they were housed in poor premises and were short of resources, and indeed there was no understanding of how to recognise technical ability. This was acknowledged by the Ministry in its next secondary pamphlet when it referred to the:
'notorious difficulty of recognising technical aptitude at 11.....it may be necessary to transfer those pupils who turn out to have been wrongly allocated.'[41]
Technical schools sought to raise their profile and when the ban on their pupils taking ordinary public examinations, such as the GCE, was lifted they started to compete openly with grammar schools. Government's attempt to meet industry's needs through state education, as in so many instances, was once again thwarted.

Pressure on schools to be more responsive to the needs of industry continued to grow. Porter (1967) in the Introduction to, 'A Schools Approach to Technology,' the second curriculum bulletin issued by the Schools Council observed:
'During the past five years or so there have been many appeals to the schools that they should adapt their curricula and their methods to meet the needs of a society which is in the throes of a second industrial revolution. Such pronouncements have

often come from individuals or from institutions concerned for the economic health of the nation, and especially for the supply of suitably qualified and skilled manpower to meet its needs.'[42]

Despite the curriculum initiatives of the 60s and 70s, described earlier, criticism of schools continued to mount. The criticism culminated in the speech made by the then Prime Minister, James Callaghan at Ruskin College, Oxford (1976). Callaghan chided the educational system for its irrelevance to working life. Under the title, 'Preparing future generations for life,' Callaghan said:
'...the goals of our education are clear enough - they are to equip children to the best of our ability for a lively, constructive place in society and also to fit them to do a job of work.'[43]

The theme of this landmark speech is still the tenet which is used to advance the need for educational reform. Pring (1986) summed this up in the following terms:
'People, particularly employers, were concerned about standards, about attitudes to industry, about the values being (or not being) promoted, about lack of basic skills.'[44]
Callaghan's speech not only launched the great debate on education it also established the tone for the debate in which the curriculum, so it was claimed, should become more relevant and indeed more vocational. This was summed up by a high ranking Treasury Official who was able to say:
' We took a strong view that education could play a much better role in improving industrial performance. The service is inefficient, rather than unproductive, and does not concentrate scarce resources in the areas that matter most. The economic climate and imperatives are clear; the task is to adjust education to them.'[45]

The Technical and Vocational Education Initiative (TVEI) announced by the Manpower Services Commission in 1982 was a major initiative aimed at rectifying the situation. This initiative enabled consortia of schools and colleges to develop curricula from 14 to 18 that would meet certain general criteria, for example, greater technical and vocational emphasis; links between schools and colleges and the world of work and, of specific interest in the context of this thesis, the introduction of regular assessment based on previously established criteria. TVEI was indicative of Government adopting new strategies as it represented the first major development in schools which was not funded by or wholly responsible to the normal education authorities.

These funds were significant and enabled schools, especially those involved in the piloting phase, to update and introduce equipment previously beyond the reach of

schools in the state sector. One of the principal beneficiaries was the technology element of the curriculum. This area has an appetite for resources which can never be fully satisfied. Following TVEI the Government has continued the policy of targeting funds through a range of selective initiatives, such as City Technology Colleges (CTC) and Technology Schools Initiative (TSI)[46]. Consequently, the availability of resources and facilities is no longer consistent in State-maintained secondary schools, an issue with serious consequences following the imposition of the National Curriculum. The lasting effect of TVEI is debatable. Some, such as Shaw (1991), conclude it was rather superficial and any benefits which might have acrued were subsumed by the National Curriculum reforms; he observes:
'In schools TVEI flared up briefly, over a hundred different projects, but it was rapidly taken over by teachers despite the early intentions, and assimilated to the mainstream of education away from the narrower vocationalisim and instrumentalism of the founders..... Yet it may well be that TVEI will come to be seen as marginal in curriculum terms.'[47]
This might be the case in relation to the school curriculum but the philosophy of TVEI has, as Smithers (1993) observed, been influential in the development of vocational qualifications:
'The cornerstone of the Council's philosophy, which owes much to TVEI, is loosely derived from behavioural psychology and argues it is what people can do that counts, their competence; what they know or understand can be inferred from what they do.'[48]

Perhaps Government advisers were aware that, yet again, an initiative with a specific focus was all too easily hijacked by teachers with other objectives.
'How many teachers 'took the money and ran'. Certainly many of them went to some lengths to justify their 'selling out' to the MSC in such terms. And, indeed, it is quite clear that in many cases the money was taken and used for purposes few would be inclined to quarrel with. Further, it must be recognised that this was invariably done with the full support of the MSC itself, or at least of those who were employed to handle its affairs and distribute its funds.' [49]
The conclusion was being reached that if the curriculum was to be permanently changed it would only be achieved by the compulsion of legislation.

The Influence of the Assessment of Performance Project

Throughout the development of design and technology, via its many identities, such as manual training, heavy crafts, handicraft, technical studies and craft design and technology, the focus of research had always been on improving the quality and relevance of the pupil's experience. Understanding and assessing the essence of the subject and defining its generic characteristics has in contrast been of

secondary importance. It was providential, therefore, that the last major project, before work commenced on the National Curriculum, set out to remedy this situation. In 1975 the Department of Education and Science had established the Assessment of Performance Unit. The APU was established to record what children were actually achieving, following the curriculum diversification which had taken place in the previous two decades. The pressure for such a unit had been created by the mounting criticism of school standards, exemplified by the Black Papers (1969 and 1970)[50]. The role of the unit was described by Jean Dawson, at one time its administrative head, in the following terms:
'The APU's main purpose is monitoring children's performance, to provide objective information about national standards of children's performance, so that those concerned - teachers, local authorities and central government - may have a reliable and dispassionate measure of the performance of the education system and can then better decide on improvements.'[51]

The Unit was not established to change the curriculum, but it was apparent that it might provide evidence which would promote change. The establishment of the Unit was symptomatic of the gradual move to centralisation and of the increasing importance of assessment in monitoring standards. An indirect, but key consequence of the Unit's work was identified by Eggleston (1991) who concluded that:
'By demonstrating the reality of achievement in the new aspects of the curriculum it was able to reinforce the standing and esteem of the subject areas in which it worked. Conversely subjects not in receipt of APU assessment were unable to enjoy this objective, public recognition of their capabilities.[52]

Only four subjects have been the focus of major projects, three being the core subjects of the National Curriculum and the exception being design and technology. The ground for a project in design and technology had been prepared by a feasibility study undertaken in 1981 at Trent Polytechnic which published its results in 1985. In 1982 an APU report, 'Understanding design and technology' attempted to isolate what made the subject unique:
'The dominant feature.....is the bringing of skills, experience, knowledge, understanding, imagination and judgement, whatever their limitations, in the execution of a specific task. In practice, it involves the integration of a complex of activities which are specific - because they relate to a particular need; inventive - because they call for a creative response; effective - because the end result should reflect a better fit or match between need and provision than existed formerly; and evaluative because the designer is called upon, throughout the process to exercise value judgements of many kinds when arriving at the proposed solution. Evaluating the efficacy of the final solution against the original need is perhaps the most

demanding task of all.'[53]
This represented a complex challenge for the Unit and consequently much deliberation took place before a project was established at Goldsmiths College. The project covered a five year period from 1985 to 1990 with the final report being published in 1991. This project's influence on the work of the National Curriculum Working Group (NCWG) in design and technology is acknowledged in the introduction to the report:
'Our work.....has spanned the period of operation of the NCWG in design and technology, the period of consultation following their deliberations and the production of the Statutory Order for technology.....There has inevitably been a good deal of intellectual exchange between us and the developments that have been taking place in relation to the National Curriculum.'[54]

Despite, or maybe because of this influence, the problems which have beset the subject since the introduction of the Order were identified in the team's first publication:
'Because of the cross-curricular nature of design and technology activity, and its relatively recent appearance in school's curricula there is less clarity and consensus about traditions and practices.....From the basis of existing practice, however, it is difficult to achieve an acceptable and all embracing definition first of design - and second of technology - and hence of design and technology'[55]

The team developed its model of design and technology on the basis that:
'design and technology is an active study involving the purposeful pursuit of a task to some form of resolution that results in improvement (for someone) in the made world. It is a study that is essentially procedural (ie deploying processes/activities in pursuit of a task) and which uses knowledge and skills as a resource for action rather than regarding them as an end in themselves. The underlying drive behind the activity is one of improving some aspects of the made world, which starts when we see an opportunity to intervene and create something new or something better,'[56]
This approach supported the view which had emerged through earlier HMI documents that the knowledge and skills required to operate effectively were subservient to the process and should be acquired on a 'needs driven' basis.
'Knowledge and skills have to be seen as resources for action - not as ends in themselves. The design and technology curriculum cannot just be taught as a body of knowledge, or just 'delivered' to pupils.'[57]
This clear statement concerning the concept of design and technology set a precedent for an APU project as Blenkins (1992) observed:
'this statement had some influence on the way in which the subject was viewed and conceived in schools as well as on the methods by which it was taught.'[58]

The Unit's research was influential from the outset. Results were being disseminated and were affecting the school curriculum whilst the project was still in its early stages. It is apparent that the design and technology project team were seeking to change the curriculum and not just monitor standards. This team's philosophy of prioritising the process was fully accepted by NCWG which from the outset never questioned the notion that assessment of the process would be the heart of the subject.

The APU approach fostered a model of design and technology which was project based, there was no place for knowledge and skills to be taught via learning situations which have their acquisition as a specific objective. The APU team also postulated that the context in which a project was set was of paramount importance if pupils were to take 'ownership' of the activity. In their first publication the team described their motive for design and technology:
'For pupils to see the activity as genuine, they must be able to identify and understand the need and be able to consider the value issues in the task that develops from the need.'[59]

During the later years of the eighties this approach was widely embraced by the cognoscenti of the subject. The notion that all tasks (term used by APU and subsequently widely adopted) should be contextualised and that pupils should, from their first experiences in design and technology, identify potential needs and opportunities became fashionable and clearly influenced the NCWG. Although the APU had advocated this approach they recognised that it was not the only way of instigating design and technology tasks. By the time of their final report (1991) they commented:
'The assumption is that the Order requires pupils to start with a context, and then identify a frame of reference and subsequently to tie down a specific task.....we do not believe that this is an approach that is slavishly to be followed on all projects. Neither do we believe it is universally required by the Order.[60]
However, others interpreted the approach in a rather simplistic fashion with predictable consequences.

Defining a National Curriculum subject

It was with the subject's fluctuating development and wide spectrum of opinion that the Working Group would have to grapple. Clearly, efficient implementation would depend on issues beyond the control of the Group but a good starting point would be to produce a document which would meet with the approval of the majority of practitioners, whilst recognising the powerful influences outside the world of education. The establishment of the Group did not take place without

some apprehension. Firstly, there was a considerable delay in setting up the technology group, following the setting up of groups in the foundation subjects of English, Mathematics and Science. Then, secondly, as Eggleston (1991) describes, more fundamental concerns arose:

'The concern was heightened when it was announced that a few additional members with technological expertise were to be added to the science Working Group so that the Group could also report on primary school technology. Many people feared that the powerful science lobby had achieved what amounted to a takeover of technology - that it would become applied science and that science would effectively control 20 per cent of the curriculum rather than the 10 per cent specifically labelled 'science'.[61]

The working group established by the then Secretary of State for Education and Science, Kenneth Baker, had ten members including the chairperson Lady Parkes. The team came from a wide range of backgrounds. It was apparent that the team might be lacking in subject expertise but it was quickly appreciated that this would not hinder them from producing both an independent and a visionary solution. The team included two teachers (one of whom was a re-trained Drama teacher, the other's expertise being Technical Drawing), an information technology expert from higher education, an LEA adviser for Business Studies, a Chief Education Officer, three representatives from professional/Government agencies and an independent consultant. The HMI observer, the Staff Inspector for craft, design and technology, was undoubtedly an influential voice in the group's deliberations. The Group felt it necessary to co-opt two people with subject expertise, a Professor of Science Education and an LEA Inspector for CD&T. By the time of the final report these two had become full members of the group. During the first months of its work the group renegotiated its terms of reference so that by July of 1988 the group's remit was extended to all key stages and all aspects of the subject, confirming the integrity of design and technology. In presenting her group's Interim Report to the Secretaries of State for Education and Science and for Wales, Lady Parkes commented:

'We are conscious that with this interim report we are breaking new ground. Our aim has been to develop an approach to design and technology which will enable pupils to achieve competence by engaging in a broad range of activities which are currently undertaken in a number of different school subjects. This approach will require a far greater degree of planning and cooperative working on the part of schools than currently occurs. We believe that teachers will welcome the opportunity to develop coherent programmes of activities for design and technology to which they can contribute their own subject expertise.'[62]

The integration of these subjects into design and technology was ambitious. Even though there were definite overlaps in their philosophies this realignment would

prove for many to be a 'shot-gun' alliance.

The report proposed a model for design and technology based on a single profile component: design and technology capability, comprising five attainment targets, see appendix 2.1, page 256. In addition the report provided examples of both the statements of attainment and of the programme of study. The proposed linkage between the programme of study and the statements of attainment was, in this report, much clearer than in final Order. The attainment targets appeared directly descended from the published material of The Design and Craft Education Project, cited earlier (see 12 above). The interim report was far more circumspect in relation to the programme of study. Basic principles were outlined for comment, but the daunting task of prescribing their content and structure was not tackled in detail. The examples were rather insignificant and general in nature. This sent a powerful message to the subject practitioners concerning the pre-eminence of the attainment targets, these were to be the foundation of the subject. The attainment targets would be the reference point from which all other matters would be determined. Another clear signal from this report was the importance of context in relation to making tasks relevant to pupils. The group thought it of such significance that they defined the term:
'We use it in the sense of the situation or set of circumstances in which arises the need or opportunity that prompts design and technological activity....a range of contexts is important not only to encourage motivation for learning, but also to ensure for them a balanced experience of the use of different resources of knowledge and skills and of the appraisal of constraints.'[63]
It would have been helpful if other terms had similarly been defined and not left to individual interpretation.

The Secretary of State welcomed the report, noting and approving:
'the view of design and technology as an essentially practical activity.... the relationship between knowledge and skills and the practical activity of design and technology being complex and interactive.... that a knowledge led approach to design and technology would not reflect the essentially practical nature of design and technology.'[64]
He agreed that the report would;
'..serve very well as a basis for informal consultation,'[65]
and asked the group in its future work to focus on a number of points. In relation to the content of the subject he posed two questions: How can technological problem-solving help to develop economic and careers awareness and business understanding? Can design and technology cover all aspects of design or will other aspects need to be covered in other areas of the curriculum? In relation to more general issues he asked the group to coordinate recommendations with other statutory proposals and to consider the vocational possibilities for key stage 4 by

drawing on the experience of the TVEI unit. These points indicated his intention to raise the profile of business awareness and the vocational potential of the subject and steer the group towards the demands being placed on education to be more relevant to the needs of industry.

The Working Group's final proposals were published in June 1989. In her letter to the Secretaries of State Lady Parkes articulated clearly the group's aim:
'...our proposal for design and technology is to prepare pupils to meet the needs of the 21st Century; to stimulate originality, enterprise, practical capability in designing and making and the adaptability needed to cope with a rapidly changing society.'[66]
The report's proposals would also test the education system's capacity to adapt rapidly if its proposals were to be implemented efficiently. The five attainment targets were compressed to four. The correlation of these generic components to those produced at Keele University in 1972 was now even sharper. The benefits of Project Technology could also be explicitly detected in the programme of study. These were now detailed under 16 headings. There was a significant technological content and several elements, such as energy, systems, structures and mechanisms, owed their definition and inclusion to the development work undertaken by the Project. The Secretaries of State's request had been heeded as a strand relating to business and economics was also included. The programme of study was described by level, providing the possibility for direct links with the statements of attainment, a possibility not fully realised.

The Working Group now handed the task to the National Curriculum Council to conduct a statutory consultation and then report the findings to the Secretary of State. This was completed by November 1989 and in submitting this report to John Macgregor, who became Secretary of State for Education and Science in the July reshuffle, Duncan Graham, Chief Executive of NCC, highlighted the key points:
'Consultation showed general agreements with the attainment targets for the design and technology profile component.....the programmes of study were considered by some to be too ambitious..... Council recognises the need for differentiation in the programmes of study, and believes that it has been able to improve them, without reducing their rigour, by producing a general programme for each key stage with qualifying statements for pupils working towards each level.....technology is an activity which goes across the whole curriculum, drawing on knowledge and skills from across all foundation subjects....Council has been advised by its task group on technology.....that its recommendations are in every sense practical and realistic.'[67]
These recommendations represented a significant shift in the programmes of study. This was justified via the consultation in the following terms:
'the programme of study covered technology adequately but there were significant

reservations that there was a craft, design and technology bias, that the range of the programmes of study was too wide and daunting and that the new technologies needed more emphasis. Many local education authorities and organisations advised a key stage approach. The sixteen sub-headings were not popular.....many felt the list was unnecessarily complicated and would make planning difficulties for teachers.'[68]

Council Officers had produced a programme of study which attempted to meet all the principle criticisms. With the benefit of hindsight, it can be seen that although the result mollified the criticism, the solution was unwieldy and unmanageable and lacked focus and definition. It appears that there was a lack of clarity and an inability to define precisely what the content of the subject should be. By reducing the 16 headings to 4: developing and using systems; working with materials; developing and communicating ideas and satisfying human needs; a softer perception of the subject was created, one with less rigour and compulsion to get to grips with the technological end of the spectrum. The folly in not linking these four strands to the four attainment targets would quickly become evident. Criticism of the programme of study was inevitably going to be more informed as it dealt with what teachers taught, it detailed the content of their everyday teaching and they naturally felt more confident of their views and more willing to fight for their own area of expertise. In contrast the attainment targets, being content free, met with much greater and less critical, approval. They also represented an approach to assessment with which teachers were unfamiliar, and they therefore could not predict the problems which might occur when adopting this approach. The statements of attainment were, however, significantly different from those which the working group had proposed.

These proposals, with minor modifications, were accepted and after the Statutory Instrument was placed before Parliament in March 1990 they became the Education (National Curriculum) (Attainment Targets and Programme of Study in Technology) Order 1990, see appendix 2.2, page 257. The Order came into force in accordance with articles 2 to 5 of the Statutory Instrument. These articles detailed the enforcement of the Order in relation to the four key stages. If these articles had stayed in force the first pupils to have had a full statutory period of education based on the Order would have reached the end of compulsory education in 2001.

The 1990 Order was significant for many reasons. Apart from the apparent change of status and recognition which compulsion created it also sought to establish some basic tenets on which the subject could be founded. It enshrined the pre-eminence of process over product in relation to the assessment of the subject. This

resulted in a fundamental shift in the way in which teachers delivered and organised their teaching. This revolved around the demands of Te 1 (Technology Attainment target 1) - Identifying needs and opportunities and to a lesser extent Te 4 (Technology Attainment target 4) - Evaluating. These two attainment targets, interpreted logically as 50% of the subject, were extremely time consuming and to many teachers and pupils they would prove challenging if not confusing. The requirements of these attainment targets changed fundamentally the way in which the subject was taught and the way in which learning occurred. This shift produced a subject more reflective and academic and therefore less practical in nature.

The Order also encapsulated the ambitions of those projects which had sought to broaden the educational objectives of the subject. The subject could have been used solely as an instrument to advance technological change, utilitarian in its approach to the economic needs of society. The Order established that the subject involved far more than just making young people technologically skilled and aware. It recognised that technological change is inevitably accompanied by social, moral and political change and that these value issues are an essential ingredient of the subject. These issues coupled to the adoption of a criterion-referenced approach to assessment were to bring about perhaps the most turbulent period in the subject's development. This was in no small part due to the opposition to the Order from groups with specific subject interests which, despite initial wide spread acceptance, coalesced into a faction which, from the outset, was bent on undermining the integrity of the Order. This was a novel position for design and technology as it was no longer on the periphery it was now in the eye of the storm.

The assessment of design and technology

From marking to assessment

From the introduction of manual training and for the greatest period of its relatively short life, subject practitioners have paid little attention to the way in which they should assess the success of either those taking the subject or their own teaching. Assessment has traditionally been based on a pupil's skill in reproducing the teacher's model or the accuracy with which the pupil translated a drawing into an object. Marks were awarded solely on the quality of the outcome. Being wholly practical in nature, the subject did not fit the conventional examination mode. Consequently, in examination terms the subject was seen to have little value or importance:

'The secondary teacher often regards the handwork portion of the curriculum as so much 'waste' time, or, at least ineffective time so far as examination work is concerned.'[69]

The established approach to assessment, based only on the quality of the product, as Kimbell (1982) observed was extremely narrow:

'Assessing the quality of work in a traditional craft context would simply require the application of criteria of craftsmanship and success will be measured by the accuracy with which the child has conformed to the teacher's plan. Nothing else is required of the child so nothing else can be assessed and the teacher can decide the details of the assessment scheme right at the start, All children can be scored according to their performance in reproducing the craft 'job'.'[70]

Indeed, if assessment reflects, as it invariably does, the objectives of the teaching and learning then they, by association, were similarly narrow.

Teachers in marking the quality of craft work could, and did, apply very precise criteria. Judgements concerning the accuracy of a piece of work could be based, for example, on measurements made with precision instruments and templates. Frequently the exercise being assessed had been devised so that it involved accurate fitting or interlocking components, such as joints, so that the quality of the outcome was all too evident. It followed that the judgements made and the marks awarded were generally sound, they were both valid and reliable and the evidence, on which the judgement was based, was permanent. Such standardisation was achieved because the rigour involved in the training of handicraft teachers focused on common and agreed norms in relation to the quality of craftsmanship. These standards, which were transmitted from tutor to teacher to pupil verbally and by exemplification, were based on the trained craftsman's eye in making comparisons. This approach to assessment was associated with phrases such as, 'the quality of finish', 'the cleanness of the lines', 'honest use of materials', 'a bodged job'

(incorrectly used to infer a clumsy or spoilt piece of work); terms which had their roots in the Arts and Crafts movement. This is not surprising as it was the philosophy of this movement which provided the bedrock of manual training. However, in education these values survived unchallenged until the 1950s whilst in industry they did not outlast the social transformation which occurred after the first world war. For over half-a-century craft teachers judged work by these established norms, ones to which they had been trained to aspire to and which many had achieved. When work in schools revolved around sea-grass stools, coffee tables, table lamps and magazine racks then it was sufficient that:
'attainment was confined in the main to motor skills and low order cognitive abilities, it was perhaps sufficient to use end products as the evidence on which to base assessment of pupils.'[71]

It was only with the introduction of the General Certificate of Education that pupils taking these subjects could achieve qualifications in this subject area comparable to those in academic subjects. The subjects on offer to pupils up until the sixties, centred on metalwork, woodwork and technical/engineering drawing, only the later having some degree of respectability and prestige. Teaching and assessment based solely on hand craftsmanship was beginning to be called into question and in the early sixties Porter (1964) observed that:
'Since, however, most schools would claim to pass on certain standards in hand craftsmanship, it may well be asked whether this is still possible or relevant in the world of automation. We have already reached a stage when people who are otherwise well educated have little appreciation of the finer points of craftsmanship.'[72]
Craftsmanship as a sole objective was losing credibility due to its increasing irrelevance. So here was the start of the transition from marking based principally on skill to a realisation that assessment should value both the potential of the production process, as an educative medium, and the technological knowledge on which the production was based. Once the pupil is given some say in what he or she will design and the form and structure of that design then, as the learning objectives become more ambitious so, accordingly, will the potential for assessment increase. This shift from product to process as an attribute of the activity worthy of assessment was clearly established by the end of the sixties. The Schools Council Working Paper 26 stated:
'Orthodox examination techniques may provide a reasonable measure of such objectives as the acquisition of skills and knowledge, but a problem arises in assessing teaching that seeks to attain such objectives as 'the development of desirable personality traits and attitudes'. In this the subjective assessment of a teacher may be of value, and if it is conducted within a clearly defined structure it may give a useful appraisal of the attainment of pupils in terms of specified educational objectives.'[73]

The objectives articulated by this paper were extremely persuasive and started to effect examinations in design and technology in the early seventies. Changes to the assessment techniques used by schools have invariably been introduced by external assessment procedures introduced by examining bodies. Examination syllabuses have always been a powerful influence on forming and shaping the curriculum as Grady (1976) noted:

'Examination syllabuses, over-riding and inescapable curriculum influences, suggest methods of work in good faith as teacher guidance. Too often these harden into rigid codes of practice leaving little room for innovation..... Tightly specified examiner expectations, exemplified by syllabus instructions, reinforced by result experience, easily translate teacher guidance and pupil activity into an exercise in satisfying requirements.'[74]

Novel approaches, required for external examination, demonstrated to teachers the benefits and improvements of these techniques. These assessment techniques rarely penetrated down to pre-examination classes. Some teachers and schools adopted these procedures but, possibly because of the prescriptive nature indicated by Grady, all too frequently teachers modified only the content and subject matter, to reflect the changes in the syllabus. For the majority of teachers the most common assessment strategy has always been based on a norm-referenced, 10 mark scale.

The recognition of project work and course work

Examination bodies, at CSE, GCE and 'A' level progressively introduced a design component into the syllabus and started allocating marks for students' performance in these areas. The move to assess the design process as well as the outcome coincided with the introduction of the CSE (Certificate of Secondary Education). This was the first external examination to accept and encourage project work as part of the formal assessment process. Responsibility for assessing project work had to be undertaken by the teachers supervising the pupils engaged in these projects. This required the examining boards to have confidence in these teachers as examiners; a trust which some sectors of society have never fully accepted. Project work, taking place in normal teaching time, raised many practical issues:

'Course work and project work are suited to teacher assessment, since the teacher is the only person able to watch the development and progress of the work, and to disentangle the contributions of individual pupils to a communal piece of work'[75]

In addition to these practical concerns project work raised more fundamental issues. For example, what were the pupil abilities which teachers could assess via project work and could these behaviourial qualities be assessed with consistency and objectivity? If examining bodies required quality assurance then the mechanism to provide this should have been central to the syllabus and the

assessment procedure. Initially, this was not the case.
'An analysis of 73 CSE mode 1 syllabuses, each containing the facility to assess project work, did not reveal one list of objectives or precise specification of attainment. The majority of syllabuses contain aims expressed in a vague and generalised form. Where assessment procedures required teachers to mark subjective work the criteria were equally broad. Qualities such as 'creative ability', inventive design', standards of craftsmanship', 'pride in achievement', occur frequently in syllabuses both as aims and in marking schedules. Hence doubt exists of the validity of teacher assessment in all these syllabuses.'[76]
The first examinations to specify behaviourial objectives, so that they could be used as assessment criteria, achieved this by dissecting the design process and allocating marks for the various stages. In this sense assessment started to dictate the fashion in which tasks or problems were solved. (It should be noted that this was a significant departure from the way in which commercial designers, in all sectors of design, were and are judged. Commercially it is primarily the success of the end product by which achievement is recognised.) Consequently, an educational model of designing evolved which valued all components of the process as highly as the final outcome. There are also many examples where the fluidity of the process was disrupted to ensure sufficient assessment evidence was produced. Many syllabuses, for example, asked children to produce a defined number of solutions to a problem before choosing which one to develop. The number of design stages varied in detail and complexity, often according to the level of the examination. For example, a North Western CSE Examination (NWEB) in design studies, devised by the Schools Council Project in Design and Craft Education, had seven stages whilst 'Design' an Oxford Local Delegacy Examination 'A' level syllabus, required teachers to make judgements using a complex assessment matrix which could be converted by visiting examiners into scores.
'The staged process was converted into an assessment scheme within which all the stages became - in turn - the focus of explicit assessment.'[77]

The North West Examining Board course was called, 'A course of studies in Design'.[78] This syllabus identified a set of 12 generalised abilities. These were derived by analysing a record of pupil 'behaviour' when engaged on design projects, see appendix 2.3, page 253. This was recognition that if pupil performances were to be compared, core abilities needed to be identified for three key reasons: the wide range of activities with which pupils would engage; the impossibility at the beginning of a design activity to predict what outcomes might emerge and to ensure the assessment had, 'face validity' - an acceptable workload for teachers. The challenge of this type of assessment for teachers was significant as it required them to observe pupil activity, relate this activity to an objective

(stage in the process) and then make a judgement about the quality of performance. The exam group recognised that although some evidence would be in the form of an end product or project report much would depend on teacher observation and consistency in reaching a judgement about the level of achievement. The advocates of this syllabus could quite rightly claim:
'The first step appears to have been taken towards assessing children rather than 'jobs' - we do seem to be getting away from the hang-up on end products.'[79]

In 1971 the Oxford Delegacy of Local Examinations introduced an 'A' level entitled 'Design'. This syllabus also broke new ground in the assessment of design-based project work. A key issue at this level was to gain recognition in relation to University entrance. This was one reason why this examination consisted of two three hour theory papers in addition to an assessment of the candidate's course work by a visiting examiner. However, the course work attracted 60% of the final mark with the major product accounting for the majority of these marks. This project was assessed, firstly by the teacher, against ten headings, then these assessments were moderated by a visiting examiner who conducted a viva voce with each candidate. The headings represented both stages and outcomes from the design process and unlike many other syllabuses asked teachers to make judgements about qualities such as self-reliance and initiative.

Although both the CSE and the 'A' level syllabuses broke the process down into stages, the way in which the teacher judgement was made was markedly different. The CSE syllabus required a teacher to allocate a score from 0 (no real attempt) to 5 (comprehensive grasp of the capability); these scores were aggregated to produce an overall score. As Kimbell (1994) concludes:
'It was in fact a fine example of norm-referenced assessment, for the practice of teachers was typically to rank-order their pupils and then distribute the order across the mark range on the assumption that the best pupils should get top marks and the worst should get the bottom marks.'[80]
The 'A' level syllabus also had five response categories but in place of marks the teacher was provided with descriptors, different ones against each heading. For example, against the heading 'Quality of Evaluation' the rating at the highest level of performance stated:
'Able to give good and dispassionate criticism; finished product thoroughly tested';
and at the lowest level of performance:
'Unable to criticise own actions or design'. It was in effect a five point scale and clustering about the norm or mid-point was probably as prevalent as in any other assessment system of the time. This tendency was encouraged by the language used which described positive achievement on one side of the mid-point and negative achievement on the other. In essence this was norm-referencing using

descriptive criteria. A genuine criterion-referenced approach would have presented a picture of positive achievement even at the lowest level. However, such an approach would have created serious problems of interpretation as these two lower categories generally correlated to failure in relation to the exam being taken. Although this examination put a high value on project work it did not trust the supervising teacher with the assessment. The visiting examiner had the authority to make any changes they wished to the judgements in converting them to marks and indeed it was not common knowledge that the headings did, in the final analysis, have different weightings. It was also the visiting examiner's task to convert the judgements into marks.

Examinations were also introduced which focused on technological knowledge and application. Invariably these were modular in structure, each module dealing with a specific component, such as: structures, mechanism or energy. Such dissection allowed knowledge to be assessed using traditional techniques, alongside a project which required a pupil to integrate knowledge and skills, acquired independently, in solving a practical problem. Projects were generally functional in character with frequently little attention being given to any potential users. One of the first examinations of this nature was Control Technology which emerged from The Schools Council, Project Technology. Introduced in 1974 as either a mode 1 or 3 CSE course it was well supported by the National Centre for Schools Technology and was eventually offered in 1975 as a mode 3 'O' level and in 1982 as a normal GCE, by the Associated Examination Board. Other GCE examinations in this area, such as the Cambridge 'O' level Technology and the similarly titled Southern Universities Joint Board exam (both examined for the first time in the Summer of 1979), followed this developmental route. The Cambridge course, for example, was derived from a CSE offered by the East Anglian Examination Board. In the Cambridge examination assessment was equally split between examinations and project work. There were two examinations. The first worth 30% tested knowledge, the second, a design task of an engineering nature, accounted for 20% of the final assessment. A typical design task was one which required pupils to design a coolant feed for a shaping machine. In this paper, as in the project, marks were allocated for various stages in the design and development process. Pupils, as indicated earlier, were asked to suggest a number of solutions before selecting one worth developing. This became a tokenist approach which claimed to assess a pupil's creativity and capacity for lateral thinking. Many of these tasks are more correctly defined as problem-solving as there were no requirements to identify issues or assess implications and the pupil was generally rearranging existing elements. These developments led to examinations of an increasingly specialised character as teachers developed their expertise in, and enthusiasm for, a particular technology, for example, Cambridge 'A' level; - Electronic systems.

During the nineteen seventies and early eighties the assessment of course work was adopted, for commercial as well as educational reasons, by most examining boards at most levels of external examination. As Kimbell (1982) noted:
'Assessment can never be the aim of a design exercise, it can only be a by product and the information gleaned from a design examination will differ only marginally from that acquired by the assessment of normal project work.'[81]
In the acknowledgement of this examination boards began to rely progressively more heavily on the continuous assessment of course project work. As this resulted in a move to teacher assessment the boards required a device for ensuring quality control. This resulted in the increasing demand for itemised assessment judgements in the belief that this would lead to more reliable assessments. If this was the case for external assessment there is, however, no substantive evidence that the itemisation of assessment judgements permeated beyond course work and project work undertaken for this purpose. The time required to develop complex assessment strategies confined such systems to external examination. Before pupils chose an examination course in design and technology, at the age of 14, it was deemed entirely normal and satisfactory that any judgements reported to parents would be a subjective commentary based on teacher observations of how pupils had coped with course work. Some schools did continue to use conventional, knowledge focused, examinations but such approaches tended to disappear as they were deemed to be inappropriate.

A common examination for all

The objective of General Certificate of Education (GCE) and Certificate of Secondary Education (CSE) had been to assess pupils in the top 60% of the 16-year-old population. GCE, introduced in 1951, was designed for approximately the top 20%. At that time the remaining 80% were not deemed capable of coping with external examinations. It was felt that they would benefit from an education free from the pressure of examinations. As a result of the Beloe Committee,[82] the CSE was introduced in 1965, this examination targeted the next 40%. But once introduced, the divisive nature of fitting pupils into one of three categories, non-exam, CSE or GCE, quickly convinced teachers of the need for a common examination at 16+. By 1970 the Schools Council was also convinced and they embarked on a wide ranging development and feasibility study. The result was a report in 1971 [83] in favour of a common examination. Consecutive Secretaries of State procrastinated for almost a decade, balking at introducing a common examination and terminating the GCE and all it stood for in relation to an elitist approach to education. Shirley Williams set up the Wadell Committee; it reported in 1978 but its recommendations, following a change in Government, were not accepted until 1980. It was a further four years before Sir Keith Joseph announced

in June 1984 that GCE 'O' level and CSE would be replaced by a common 16+ examination, the General Certificate of Secondary Education (GCSE), from the Summer of 1988.
With the introduction of a new examination system new assessment procedures were also to be introduced. The Department of Education and Science recognised that here was an opportunity to increase levels of comparability between the regional examination groups. The work to achieve this had been put in place in 1980 with the development of national criteria in the twenty most important subjects. This work was carried out by the Joint Council which was established by the GCE and CSE Boards. In the case of design and technology there were initially four groups, one for each of Metalwork, Woodwork, Combined materials and Technical Drawing. These independent groups were soon amalgamated into a single group charged with monitoring Craft, Design and Technology. The group was chaired by Lady Parkes, who was to chair the National Working Party for Design and Technology. It included representatives from each exam group and a number of independent subject experts. The group evolved, via a consultation process, a single set of grade descriptors which all exam groups were required, by the Schools Examination Council, to employ.

The grade descriptors were incorporated into the National Criteria for Craft, Design and Technology. These requirements, for the first time, controlled centrally all aspects of the curriculum: the aim of GCSE courses, the content, the relationship between assessment objectives and content and the techniques of assessment, see appendix 2.4, page 263. GCSE examinations placed more emphasis on assessment than any previous exam. Consequently, teachers started to focus more on these issues than on the content:
'Whereas in the past the choice of content has perhaps tended to dominate the process of syllabus development, the emphasis has now changed with the result that considerable effort has been spent articulating the aims and assessment objectives of courses and working out the implications of these aims and objectives for techniques of assessment'[84]
This trend was to continue with the introduction of the National Curriculum.

The GCSE had many novel features perhaps, most importantly for Craft Design and Technology, the General Criteria (para. 19(e)) prescribed:
'the principle of fitness for purpose must be observed: all examination components and assessment procedures should reflect and be appropriate to the nature of the subject, its educational aims and its assessment objectives'. [85]
The practical consequence of this was that examinations could include a significant element of course work which would need to be assessed by the teacher and moderated by the examination group. Many syllabuses grasped this opportunity

and developed courses in which the final award was derived totally from course work undertaken during the two years of the course. For example, the Welsh Joint Education Committee offered a GCSE, 'design and technology', based on three projects undertaken during the two years. Two minor projects, which accounted for 50% of the marks, had to be selected from three provided by the Board and the third project, the major project, was self-determined. Consequently, design and technology shifted away from the constraints of terminal examinations in either the practical or knowledge based format. This in turn resulted in less emphasis being placed on the knowledge which underpinned the subject and aspects such as materials technology only received attention when they impinged on some aspect of the practical activity.

Assessment objectives did not specify any particular knowledge or skill, instead they served to emphasise the central importance of the design procedure:
'The aims of CDT provide a general direction and purpose for all activity, the objectives provide almost a checklist of specific performance goals against which to measure the candidates.'[86]
The central responsibility for assessment now had to lie with the supervising teacher:
'They (teachers) are the only ones who are really in a position to make such assessments. Not only can they see the evidence that students might present to an examiner, but they are also aware of all the unrecorded events that the students suffered from or gloried in, and which might be relevant to forming an accurate picture of the candidate's real ability.'[87]
The notion that the teacher was the only person capable of making an accurate assessment is undeniable, however, this did not guarantee the standardisation of these assessments. It is likely that teachers ranked their pupils accurately but the consistency of the interpretation of the national grade criteria was dependent on the strategies employed by each Examining Board.

One of the purposes of any examination is to discriminate between candidates. GCSE required examiners to employ differentiation to achieve better discrimination. Paragraph 16 of the General Criteria stated:
'All examinations must be designed in such a way as to ensure proper discrimination so that candidates across the ability range are given opportunities to demonstrate their knowledge, abilities and achievements: that is, to show what they know, understand and can do. Differentiated papers or differentiated questions within papers will be required accordingly in all subjects.'[88]
This was a radical step, traditionally candidates had been assessed as much by their failure to answer questions as their ability to answer them correctly. Now strategies had to be employed which would focus on each pupil's positive

achievements. Subjects devised appropriate procedures to meet this challenge. Those which essentially tested knowledge used stepped questions and stepped papers, differentiated papers or papers with differentiated sections. This is commonly referred to as differentiation by task. Subjects which focused on the assessment of procedural and behaviourial aptitudes based assessment on identifying different levels of achievement in students' responses to common tasks, ones neutral in respect of difficulty. This is commonly referred to as differentiation by outcome. Differentiation by outcome was widely employed by examiners in Craft, Design and Technology.

GCSE presented a major shift in examination practice as it represented a genuine attempt to assess each pupil's capability rather than to judge pupils against their contemporaries. On reflection it is evident, that for secondary school teachers, it was part of a consistent evolutionary progression from the traditional examining procedures to those that were to be articulated by the Task Group on Testing and Assessment in relation to the National Curriculum. However, those recommendations were never fully implemented. For teachers of Craft, Design and Technology the emphasis of GCSE examinations to focus on assessment by outcome, of course work by teachers and a shift away from the formal assessment of knowledge and assessment by task would create conflict when statutory assessment at key stage 3 was introduced.

Even though GCSE was, by design, criterion referenced the majority of syllabuses in design and technology did not use performance criteria as, in reality, they described response on a continuum from poor to excellent, detailing negative as well as positive achievement. Each criterion was converted into a mark and individual marks aggregated to provide an overall grade. For example, the NEA syllabus in Design and Realisation assessed course work under two main headings: design and manufacture. Design was sub-divided into 7 sections and manufacturing into 5. The sub-sections in design were described in four levels of response. A typical set of responses, in relation to recognition of problem, is given below:

	mark range
No suitable area of study identified	0
Considerable guidance needed	1 - 2
Some help or guidance needed	3 - 4
Problem identified unaided	5 - 6

This approach moved teachers towards criterion referencing but in reality many would still have related assessment to the norms pertaining in their teaching group

or school. In the NEA scheme teachers were required to judge pupils in twelve categories. Others, notably the Graded Assessment scheme for Craft, Design and Technology offered by LEAG, converted all teacher judgements into yes or no decisions (can do statements) in the belief that this would produce a most accurate and reliable measure of a pupil's achievement. This approach to assessment has the effect of greatly increasing the number of judgements which have to be made as Kimbell (1994) has observed:

'It follows that, as the units of assessment get smaller and (supposedly) more precise, they have also to get more numerous in order to cover the same ground.'[89]

But the objective of any assessment at sixteen is predominately summative in nature and however the judgements are made and regardless of the scale of the scrutiny process, the judgements have to be converted into a single mark and grade. The greater the number of judgements, the greater will be the reliance on the application of a formula in reaching a final grade. The itemisation of assessment can clearly result in lengthy and unmanageable assessment procedures. Perhaps of more concern, itemised assessment could result in a formulaic approach being adopted by pupils as they strive to meet each individual assessment criteria.

National Curriculum assessment - criterion referencing

The management of assessment proved to be the dominant issue in the implementation of National Curriculum assessment. The introduction of any new system would, initially, require more time. The nature of the assessment criteria (the statements of attainment) in design and technology exacerbated the situation not just because of structural and linguistic issues, but also because of their novelty. Despite the many achievements of GCSE it still allowed teachers to make judgements in a traditional manner as assessment criteria articulated both positive and negative performance, whereas National Curriculum judgements were based solely on positive achievement. Assessing only positive achievement requires a confidence and intimacy in the assessment criteria being applied, something which the majority of teachers would not acquire within the relatively short life of the 1990 Statutory Order. The National Curriculum Order represented a final, though possibly premature, break with marking. Teachers where now charged with equating performance with a statement of attainment. Pupils no longer received a mark of 7 out of 10 for a sheet of design work, instead they were told they had evidenced Te 2.5b and Te 2.5c. The resistance of the Schools Examination Council to provide a formula which would allow teachers to aggregate pupil performance into an achievement level was interpreted by many as indecision. SEAC though appreciated the dangers of this approach if used in relation to all aspects of assessments. Rules introduced for statutory assessment were inevitably

seized on by teachers and used out of context as many had an instinctive desire to convert assessments into a comparative structure. A desire similar to that of the Government's, albeit for entirely different reasons.

Given the government's prime reason for introducing a National Curriculum, to raise standards via accountability, it is surprising that it did not perceive that the Task Group on Testing and Assessment's (TGAT) report would prove a major obstacle to this being achieved. TGAT introduced a new vocabulary in relation to assessment, one which confused teachers and beguiled politicians. National Curriculum assessments, they decreed, were to be made in two categorises, teacher assessment or TA as it soon became known and end of key stage assessment, synonymous with SATs or standard assessment tasks. Different types of assessment with different purposes, novel names but there was nothing novel about the concepts. Rowntree (1971) defined the various modes of assessment which might be used by teachers and noted that:
'The teacher must certainly plan and evaluate his (sic) assessment methods in relation to the purposes he is pursuing. As soon as he does so he will find himself caught up in balancing the claims of various modes of assessment, e.g.:

informal	vs.	formal
formative	vs.	summative
continuous	vs.	terminal
course work	vs.	examination
process	vs.	product
internal	vs.	external
divergent	vs	convergent
idiographic	vs.	nomothetic.'[90]

Some of these terms were discussed by TGAT in relation to their proposals. TGAT envisaged a system in which:
'assessment tasks exploit a wide range (far wider than normally envisaged for tests) of modes of presentation, operation and response, and their numerous combinations, in order to widen the range of pupils' abilities that they reflect and so enhance educational validity.'[91]
However, as the system evolved and was implemented the two modes of assessment which they prescribed could, in general terms, be used as column headings for Rowntree's modes of assessment. Teacher assessment was intended to be idiographic by nature, focused on the individual and reflecting all the evidence of achievement produced, whilst end of key stage assessment was broadly nomothetic as it was intended to monitor systems at a range of scales from an individual class to national performance. However, the dominance of the

formal, summative end of key stage assessment and the relegation of the informal, formative teacher assessment convinced teachers that National Curriculum assessment procedures would not improve the quality of teaching and learning but would so infect the curriculum that it would undermine their professional integrity leaving them merely as operatives capable of being judged by performance figures.

Conclusion

The status of design and technology had been constantly rising during the thirty year period leading up to the National Curriculum. Since that high point the subject has failed to live up to the expectations of those responsible for its elevation. The demands which were placed on the subject, in addition to the legal demands, were probably achievable but only with time and careful nurturing. The story of the subject's rise was probably not as progressive as this account implies. It has focused on the many positive steps forward without detailing the steps sideways or backwards. The events of the last few years are still too close to reflect on with confidence, as to how they will affect the subject in the long term. It is still though a remarkable achievement that design and technology, as a result of the National Curriculum, became and still is, an essential and required element of every child's education from 5 to 16.

The way in which the subject is assessed is still evolving. As the key objectives are better understood then the assessment strategies devised to reflect performance accurately will become ever more sophisticated. It is to be hoped that sophisticated will not in turn mean complex. At the heart of this research has been a desire to produce elegant and efficient assessment devices, ones which would be of value to the teacher and the taught. A lasting legacy of the National Curriculum will be the increased awareness of the teaching profession of the value of assessment in improving the quality of teaching and learning. In design and technology there is still much to do. This chapter has hopefully established the logical context for the developments which were to follow and how they contributed to the subject's understanding of assessment issues.

References

1. Young, M.F.D. (1971). 'An approach to the study of curricula as socially organised knowledge' in Young, M.F.D. (Ed) *Knowledge and Control: New Directions for the Sociology of Education.* London: Collier-MacMillan, page 31
2. Young, M.F.D. (1971). 'An approach to the study of curricula as socially organised knowledge' in Young, M.F.D. (Ed) *Knowledge and Control: New Directions for the Sociology of Education.* London: Collier-MacMillan, page 34
3. Apple, M.W. (1978). 'Ideology, reproduction and educational reform', *Comparative Education Review*, No. 22, page 380
4. Blenkin,G.M., Edwards,G. & Kelly,A.V. (1992). *Change and the Curriculum.* London: Paul Chapman, page 15
5. Goodson, I.F. (1988). *The making of Curriculum.* London: Falmer, page 164
6. Jenkins,D. & Shipman,M.P. (1976).*Curriculum: an Introduction.* London: Open Books, page 107
7. Layton,D. (1972). 'Science as General Education' in Goodson, I.F. (1988). *The making of Curriculum.* London: Falmer, page 165
8. Penfold, J. (1988). *Craft, Design and Technology: Past, Present and Future.* Stoke-on-Trent: Trentham, page 20
9. Board of Education (1926). *The Education of the Adolescent* . (The Hadow Report)
   Ministry of Education 15-18, (1959). *The Crowther Report.* London: HMSO
   Ministry of Education (1963). *Half our Future,* (The Newsom Report). London: HMSO
10. Department of Education and Science (1988). *Education Reform Act.* London: HMSO
11. Kimbell, R. 'A problem of success', *The Times Educational Supplement* (16th. October 1992)
12. The Assessment of Performance Unit Project in design and technology. (1985 -1990). Goldsmiths College: University of London
13. Kimbell, R. 'A problem of success'. *The Times Educational Supplement* (16th. October 1992)
14. Department of Science and Art: Form 813. *Suggestions for Manual Instruction - Circular 44* (5th June 1890). London: Department of Science and Art
15. Saloman, O. (1892). '*The Theory of Educational Sloyd'*
16. Dodd, T. (1978). *Design and Technology in the School Curriculum.* Sevenoaks: Hodder & Stoughton
17. Penfold, J. (1988). C*raft, Design and Technology: Past, Present and Future.* Stoke-on-Trent: Trentham
18. Eggleston, J. (1985). 'Craft, Design and Technology - the uncertain future' in *Technology in Schools,* edt. by Cross, A. & McCormick, B. (1986). Milton Keynes: Open University, page 162
19. Schools Council Field Report 3. (1966). *Technology in Schools.* London: Schools Council
20. Schools Council Working Paper 18. (1968). *Technology and the Schools.* London: Schools Council
21. Schools Council Working Paper 26. (1969). *Education through the use of Materials.* London: Schools Council, page 7
22. Schools Council. (1974). *Materials and Design: a fresh approach.* Eggleston, J. *et. al.*, London: Arnold, page 1
23. Eggleston, J. (1992). *Teaching Design and Technology.* Buckingham: Open University, page 7
24. Schools Council. (1975). *Education through Design and Craft.* Eggleston, J., London: Arnold, page 29

25. Dodd, T. (1978). *Design and Technology in the School Curriculum.* Sevenoaks: Hodder & Stoughton, page 41
26. Royal College of Art (1979). *The Three R 's' Design in General Education.* London: Royal College of Art
27. Department of Industry. Industry/Education Unit. London
28. Editorial. *View.* (1981). London, Industry Education Unit, Department of Industry
29. Young Engineer for Britain sponsored by British Industry and organised initially by The Department of Industry Education Unit and subsequently by The Engineering Council via SATROs (Science and Technology Regional Organisations)
30. The Design Council (1982). *Schools Design Prize.* The Competition Brochure, London: The Design Council
31. Thatcher, M. (October 1982)'*Presentation Ceremony - Schools Design Prize.* Design Council
32. Editorial. *Designer* (April 1982). Society of Industrial Artists and Designers
33. HMI Discussion Document (1978). *The Curriculum 11-16.* London: Department of Education and Science
34. HMI Discussion Document (1978). *The Curriculum 11-16.* London: Department of Education and Science
35. The Design Council. (1980). *Design Education at Secondary Level.* London: Design Council, Chairman's foreward - page 2
36. Shaw, K.E. (1991). 'Teaching Design and Technology ', *Perspectives 43,* Exeter: Exeter University, School of Education
37. The Technical and Vocational Initiative. (1982). The Manpower Services Commission. London: Department of Employment
38. Ripper, W. (1886) 'Technical Training as a main feature in National Education', Report of the fifty sixth meeting of the *British Association for the Advancement of Science*
39. Penfold, J. (1988). *Craft, Design and Technology: Past, Present and Future.* Stoke-on-Trent: Trentham, page 19
40. Ministry of Education (1945). *The Nation's Schools, Pamphlet No 1.* London: HMSO
41. Ministry of Education (1945). *The New Secondary Education, Pamphlet No 9.* London: HMSO
42. Porter, D.J. (1967). *A School Approach to Technology,* Curriculum Bulletin No. 2. London: HMSO, Introduction - page 1
43. Callaghan, J. (1976). *Preparing future generations for life.* Ruskin College. Oxford
44. Pring, R.A. (1986). 'The Curriculum and the new vocationalism', *The Stanley Lecture 1986.* London: Royal Society of Arts
45. Pring, R.A. (1986). 'The Curriculum and the new vocationalism', *The Stanley Lecture 1986.* London: Royal Society of Arts
46. TSI - Technology Schools Initiative (1991-3). - A competitive funding initiative aimed at enhancing technological resources in schools. Awarded by the Department for Education on the basis of school tenders which had to demonstrate positive collaboration with local industry.
47. Shaw, K.E. (1991). 'Teaching Design and Technology ', *Perspectives 43,* Exeter: Exeter University, School of Education
48. Despatches (1993). *All our Futures.* London: Channel 4 Television, page 16 - para 4.6
49. Blenkin,G.M., Edwards,G. & Kelly,A.V. (1992). *Change and the Curriculum.* London: Paul Chapman, page 88

50. Cox,C.B. & Dyson,A.E. (eds) (1969/70). *Fight for Education: A Black Paper.* London: Critical Quarterly Society
51. Blenkin,G.M., Edwards,G. & Kelly,A.V. (1992). *Change and the Curriculum.* London: Paul Chapman, page 90
52. Eggleston, J. (1992). *Teaching Design and Technology.* Buckingham: Open University, page 8
53. Assessment of Performance Unit (1982). *Understanding Design and Technology.* London: Department of Education and Science, page 2
54. Schools Examination and Assessment Unit. (1991). *The Assessment of performance in design and technology.* (The Goldsmiths' Report). London: HMSO, page 7
55. Department of Education and Science: Assessment of Performance unit. (1987). *Design and Technological Activity: A Framework for Assessment.* London: HMSO, page 7
56. Schools Examination and Assessment Unit. (1991). *The Assessment of performance in design and technology.* (The Goldsmiths' Report). London: HMSO, page 17
57. Department of Education and Science, (1977). *Curriculum 11-16,' Working Papers by H.M. Inspectorate: a contribution to the current debate.* London: HMSO
58. Blenkin,G.M., Edwards,G. & Kelly,A.V. (1992). *Change and the Curriculum.* London: Paul Chapman, page 95
59. Department of Education and Science: Assessment of Performance Unit (1987). D*esign and technology activity: A Framework for Assessment.* London: HMSO, page 8
60. Schools Examination and Assessment Unit (1991). *The Assessment of Performance in design and technology.* (The Goldsmiths' Report). London: HMSO, page 233
61. Eggleston, J. (1992) *Teaching Design and Technology.* Buckingham: Open University, page 16
62. Department of Education and Science and the Welsh Office. (1988). *The Interim Report of the National Curriculum design and technology Working Group.* London: HMSO
63. Department of Education and Science and the Welsh Office. (1988). *The Interim Report of the National Curriculum design and technology Working Group.* London: HMSO, 1.18, page 18
64. Department of Education and Science and the Welsh Office. (1988). *The Interim Report of the National Curriculum design and technology Working Group.* London: HMSO
65. Department of Education and Science and the Welsh Office. (1988). *The Interim Report of the National Curriculum design and technology Working Group.* London: HMSO
66. Department of Education and Science and the Welsh Office. (1989),. *Design and Technology for ages 5 to 16,* (The Parkes Report). London: HMSO, page vii
67. National Curriculum Council (1989). *Consultation Report on Statutory Proposals for Technology.* York: NCC, pags 3 - 4
68. National Curriculum Council. (1989). *Consultation Report on Statutory Proposals for Technology.* York: NCC, page 12
69. 'Handwork in secondary schools' in *The Book of School Handwork.* Ed. Holman, H. (1913) London: Caxton
70. Kimbell, R. (1982). *Design Education: The Foundation Years.* London: Routledge & Kegan Paul, page 160
71. Hilsum,J. (1976). 'Assessment of design based project work'. in *Studies in Design Education and Craft.* Volume 8,2
72. Porter, D.J. (1967). *A School Approach to Technology - Curriculum Bulletin No. 2.* The Schools Council. London: H.M.S.O., page 23

73. Schools Council Working Paper 26. (1969). *Education through the use of Materials.* London: HMSO, page 11

74. Grady,D. (1976) 'Are we teaching bad design' in *Proceedings - Conference of the Institute of Craft Education.* London: Batsford

75. Schools Council - Examination Bulletin 23A (1971). *Common Systems of Examining at 16+,* London: Evans/Methuen

76. Hilsum,J. (1976) 'Assessment of design based project work' in *Studies in Design Education and Craft.* Volume 8,2

77. Kimbell, R. (1993) 'Assessment of design and technology' in *Teaching Technology,* Ed. Banks,F. Milton Keynes: The Open University, page 161

78. Schools Council, (1974) *Design and Craft Education Project, Education through Design and Craft,* London: Edward Arnold, page 38

79. Hilsum,J. (1976). 'Assessment of design based project work', in *Studies in Design Education and Craft.* Volume 8,2

80. Kimbell,R. (1994). 'Assessment of design and technology' in *Teaching Technology* (1994). Ed. Banks, F.J. Milton Keynes: The Open University Press, page 162

81. Kimbell, R. (1982). *Design Education: The Foundation Years.* London: Routledge, Kegan & Paul, page 172

82. Beloe Committee

83. Schools Council Report (1975). *A Common System of Examining at 16+.* Examinations Bulletin No 23, London: Evans Methuen

84. Rees, D. (1987). *Teaching GCSE Craft, Design and Technology.* London: Hodder and Stoughton, page 15

85. Schools Examination Council. (1986). *General Criteria for the General Certificate of Secondary Education.* London: HMSO

86. Schools Examination Council and The Open University (1986). *Craft, Design and Technology - GCSE A Guide for teachers.* Milton Keynes: Open University Press. page 13

87. Schools Examination Council and The Open University (1986). *Craft, Design and Technology - GCSE A Guide for teachers.* Milton Keynes: Open University Press, pages 29 - 30

88. Schools Examination Council (1986). *General Criteria for the General Certificate of Secondary Education.* London: HMSO

89. Kimbell,R. (1994). 'Assessment of design and technology' in *Teaching Technology* (1994). Ed. Banks, F.J. Milton Keynes: The Open University Press, page 166

90. Rowntree,D. (1977). *Assessing Students: How shall we know them?* London: Harper & Row, page 118 - 119

91. Task Group on Assessment and Testing (1989). *The National Curriculum - A Report.* London: Department of Education and Science, para. 48

Chapter 3

The context surrounding the research and development of statutory tasks

Synopsis

This chapter looks chronologically at the context in which the research and development took place; commencing in September 1989 with the award of contracts and finishing in July 1993 with the industrial boycott of the first statutory assessment. This period is reviewed primarily to illuminate the way in which decisions were taken, the reasons for taking those decisions and the effect they had on the nature of the evolving assessment procedure.

During this period there were two formal specifications against which contracts were awarded. The first specification operated for a two year period and the work consistently attempted to meet this specification and was evaluated against it. The second specification was, as the first statutory assessment approached, far less tangible to grasp. It was constantly subject to changes and modification to meet the rapidly changing political perspective. All these factors are related to the effect which they had on design and technology which because of its novelty lacked the secure foundations of the core subjects, although, as with these, all aspects were subjected to intense scrutiny.

As the key stage three end of key stage assessment attracted ever increasing media attention, the way in which views polarised is reviewed. The way in which the first statutory assessment of design and technology eventually succumbed to the power of the English Teachers' subject association in uniting the opposition to statutory assessment, and the increasing disillusionment with the technology Order are analysed.

This chapter is divided into the following sections

The context
Developing standard attainment tasks - the development agencies
The steering procedure
Main steering September: 1989 to August 1990
Main steering September: 1990 to August 1991
Subject steering December: 1989 to August 1991
The Summer of 1991
Research and development: September 1991 to August 1993
National curriculum design and technology: Blue Peter or Micky Mouse?
The first statutory assessment
Confrontation with the unions

The context

It might seem unusual in a thesis of this nature to devote a chapter to describing the background against which the research and development took place. However, this was an exceptional situation and one which may prove to be unique. The task of conducting research in relation to the curriculum and its assessment had previously been left to Examination Boards, Local Educational Authorities and other regional and local educational bodies. Never before had there been a national examination, one which all pupils in an age cohort would by law be required to take. Consequently, the political involvement was intense as both the reputation of government ministers and the implementation of government policy was at stake. A national cohort of pupils in England and Wales varies between 600,000 and 650,000 and the numbers of teachers involved in a single subject, such as design and technology, could be as high as 40,000. Individual results were to be reported to parents and average school performance would be norm-referenced in league tables. There were very few in the country who would not have either a direct or an indirect experience of, or interest in, the process. Testing and reporting was the central plank in the Government's objective to raise standards in school. That design and technology was at the forefront of these developments was in itself novel but the reasons for this have been explained in the previous chapter. The unfolding of the story serves to illustrate that change in education cannot be effected solely by legislation and enforcement. The educational system is of dinosaur proportions, it takes a great deal of time for the commands from the head to reach the tail and even when they do the head has moved on, so it appears as if the tail has a mind of its own and is acting in an entirely independent and irrational fashion. Perhaps the most important lessons from this experience are that change in such a gross system as education can only be effected if goodwill and reason are maintained by all concerned and that political timescales, and the demand for instant results, must be modified when dealing with a process which takes a minimum of 11 years, namely the formal education of every member of our society.

Developing standard attainment tasks - the development agencies

The development of statutory assessment tasks officially commenced on the 1st September 1989. From that date seven agencies were commissioned by the Schools Examination and Assessment Council to develop tasks. The seven groups were responsible for five National Curriculum subjects namely: English, maths, science, technology and Welsh. In both English and technology two independent groups had been awarded contracts whilst in maths, science and Welsh a single group had been charged with the task. The seven agencies represented four

organisations. The Centre for Assessment and Testing in Schools (CATS) had gained four contracts, one in each of the subjects with the exception of Welsh. This consortium was centred on the University of London and had close links with the University of London and East Anglian Assessment Group (ULEAG). The agencies within the consortium responsible for each subject were based at constituent colleges of the University. For example, the maths group was based at King's College and the technology group at Goldsmiths' College. The other contract in English had been gained by the East London and MacMillan Assessment Group. This group was based at The Polytechnic of East London which worked in close collaboration with the Macmillan Publishing Group. The contract for Welsh was awarded to the National Foundation for Education Research (NFER) in collaboration with The University of Reading and the Normal College, Bangor. The remaining contract for technology had gone to the Midlands Examining Group (MEG). In common with all the GCSE Boards MEG had bid for four or five contracts, but had only been successful in the case of technology. The bid was made under the auspices of the Midland Examination Group National Assessment Project (MEGNAP) and in collaboration with the Technology Education Centre at Middlesex Polytechnic. The author of this thesis was a director of this agency. It is important to note that although examination groups and publishers were linked to most agencies all of the development teams were based in institutions of higher education and all the Project Directors were academics with a significant and relevant experience. The research and development associated with standard assessment tasks had attracted highly respected educationalists, all of whom were prominent within their subject community.

Arriving at the starting point had taken six months. The advertisements inviting organisations to tender had been placed in February and the announcement of the awards was made on the 22nd June 1989. The awarding of contracts had been a demanding process as they had attracted a significant number of bids. In total 15 organisations had submitted tender documents and the majority of these organisations was bidding for several contracts. This resulted in 47 separate subject bids. The tender documents were required to respond on a number of key issues which had been detailed in a specification document. In particular this document described the nature of the required SATs:
'The central task of the development agencies will be to develop the SATs to be given to all pupils reaching the end of the third key stage in 1992 having commenced the programmes of the National Curriculum in Autumn 1989. These SATs should be of the kind proposed by TGAT - that is, packages of tasks administered through a range of modes as discussed in paragraphs 47-49 of the main TGAT report, but restricted in the first instance to the assessment of

achievement in mathematics, science, English, Welsh and design & technology. Their purpose will be formative and summative. The SATs applicable at the end of the key stage should be constructed so that between them they enable pupils' performance to be ascribed to any of the ten levels associated with national curriculum ATs. Any given SAT should be capable of ascribing pupils to one of a range of contiguous levels. Some overlap between SATS of the levels covered in each will be required. A possible package for any particular range of ATs and profile components (PCs) would be of eight SATs covering levels 1-3, 2-4, 3-5, 4-6, 5-7, 6-8, 7-9, and 8-10 respectively. Teachers will have a responsibility to decide initially , in the light of their knowledge of their pupils, which range of levels to use in the case of each pupil, but the arrangements should permit pupils to demonstrate achievement up to the highest level of which they are capable.'[1]

The first stage of the selection process was the submission of a tender document. This was required to address in detail each of the issues identified by SEAC, consequently they were substantial documents (the MEGNAP tender document was in excess of 220 pages [2]). This was followed by an interview undertaken by HMI at the headquarters of the bidding agency. Eight of the initial bidders for the technology contract were interviewed and five survived this first scrutiny and were invited for a final interview at Newcombe House, the home of SEAC. The interviewing panel represented all interested parties, numbering with SEAC officers 21 in total. Each agency was allowed three representatives. The panel was chaired by Philip Halsey, the Chief Executive of SEAC. Each interview lasted thirty minutes and consisted of a ten minute presentation followed by questions, all asked by the chair of the panel. Competing agencies did not meet and the operation was carried out with military efficiency in a highly formal fashion. It is understood that the panel made its decisions prior to the lunch break before proceeding with the science interviews in the afternoon. There then followed a two week period in which successful candidates were vetted and ministerial approval of the awards sought. Following the Prime Minister's approval, successful agencies were asked to confirm that they would accept a contract; this was followed by a press release.

This procedure is outlined as it indicates the national importance being attached to this work and it established the style of relationship between the awarding authority and the contractors. The Council wished to impress from the outset that the agencies would work within the remit provided by the Council and that they would be steered as and when the Council felt it necessary. In the words of the Contract: 'In carrying out this project, the contractor shall follow the guidance given by the Council through the Steering Group.'[3]

The contract also established the confidential and secretive nature of development which the Council wished to employ:
'The Contractor agrees to keep and to ensure that its personnel shall keep all information concerning the project which is by its nature confidential, secret and confidential and shall not at any time for any reason disclose such information or permit it to be disclosed to any third party except strictly for the purpose of enabling the contractor to carry out its duties under this agreement .....'[4]

The contract was in practice, similar to that issued to any agency undertaking contract research for the Government in any field. The level of secrecy was, however, unusual in relation to education research which is normally open, informative and collaborative in nature. Confidentiality was interpreted by teachers as covert, many felt alienated from the process and consequently, in due course they came to see themselves as unwilling agents rather than trusted partners. All agencies were given contracts for three years, terminating on the 31st August 1992. This was an interesting date as it would cover the first full statutory assessments in maths and science, due to be held in the Summer of 1992, but would not cover the first statutory assessments in English and technology which did not occur until the following year. The total funding to meet the cost of this research and development was approximately £12 million, £1.9 million of which was allocated to the MEGNAP technology agency.

The steering procedure

Following the notification of the award of a contract agencies had six weeks in which to establish a research base and recruit staff. For some this presented no problem as a pool of permanent research staff already existed and they could be seconded to the project for its duration. For others however, including MEGNAP, this was a much longer process, but by November all teams were fully staffed and development and trialling in schools had commenced. The research and development process adopted by each team was overseen by two steering committees. The first was the overarching key stage 3 (KS3) Development Subject Committee. Its remit was to oversee all the agencies, deal with policy issues of general concern and to harmonise the work. Excluding SEAC staff it had a membership of 24, six of these being representatives of the agencies. The remainder representing, for example, The Department of Education and Science (now the Department for Education), The Welsh Office, Her Majesty's Inspectors in England and Wales, The National Curriculum Council and the Council of SEAC. There were normally five members of SEAC staff servicing the Committee.

In addition to the main committee a subject steering committee was established for

each of the five subjects. The membership was smaller but included the staff inspectors for both design and technology and information technology and six representatives of the SEAC subject specific committees. This committee was concerned with subject issues. Those established in English and technology were, from the outset, different from those for the other subjects. These two subjects each had two agencies developing tasks, consequently and inevitably there was a degree of competition not experienced in the other subject committees. Both the main and the various subject committees met at least three times in each calendar year; the meetings being phased so that the main steering meeting could receive reports from each of the subject steering committees. In addition SEAC established other committees to deal with specific issues. An 'information technology across the curriculum' group met to look at the implications this would have for assessment; a special needs and English as a second language group was formed to coordinate research in this area and project directors met to discuss specific issues such as aggregation, trialling, INSET and confidentiality. Each of these meetings required written reports prior to the meeting and generally, follow-up investigation and resubmission. For each agency this imposed a considerable workload in addition to the research, development and trialling required in the production of Standard Assessment Tasks.

The context of the KS3 development work must be set alongside the parallel work taking place at key stage 1 (KS1). This work had commenced several months earlier and consequently had established approaches and procedures which influenced some aspects of KS3. The work of the KS1 agencies entered the public domain, via trialling and piloting, 12 months prior to the KS3 developments. Public opinion and teacher reactions evolved from what occurred at KS1 and this in turn greatly influenced ministerial decisions in relation to both key stages. SEAC staff, because of their involvement in KS1 development, also had pre-conceived ideas of what they wanted and these they attempted to impose on development agencies. This was evident in SEAC papers relating to 1990 trials:
'It is necessary for agencies to test whether the principles being applied at key stage 1 are appropriate at key stage 3.....Similarly decisions already taken for KS1 concerning the combination of SAT and TA, aggregation to PC and whole subject levels should be examined for their appropriateness for KS3.'[5]

Main steering - September 1989 to August 1990

The initial steering meetings dealt with issues concerning policy and procedure. These were important from the SEAC perspective because their insistence on protocol assisted them in establishing their position over the agency directors. An important issue which quickly emerged and one which would prove to be a

continuing concern, was interpretation of assessment statements. The strategy to be employed to address this issue was exemplification. The technology agencies as early as May 1990 reported to the Steering Committee that there were significant problems in obtaining successful exemplification of statements of attainment. The agencies were having to initiate design and technology activities because as yet schools were not sufficiently familiar with the Order to initiate genuine design and technology activities. Another aspect specific to design and technology concerned teacher intervention:

'The Directors of the technology agencies suggested that technology raised special questions about teacher intervention. Since SATs in D&T were extended activities, there had to be a recognition that teaching and advice to pupils would inevitably and properly continue.'[6]

The agencies were instructed to train teachers so that they might appreciate what level of intervention was appropriate.

Two key issues common to all agencies concerned the piloting procedure and the laying of assessment Orders. The Order would establish the legal framework for statutory assessment. It would describe the number and type of assessment activities in each subject and when they would take place. Agencies carried out research via either a trial - controlled and organised by the agency - or a pilot - organised by SEAC and based on a semi-random selection of schools. All agency work undertaken in English and technology would be deemed to be trialling until the Summer of 1991 when the first national pilot would take place and followed in the Summer of 1992 by the principal national pilot. As the first statutory assessment in these subjects did not take place until 1993 there would be plenty of opportunities for both trials and pilots. In relation to assessment arrangements the key issue was the relationship between SAT assessment and teacher assessment. HMI felt it important that:

'...members recognise that to fulfil the requirements of the ERA, end of key stage assessment would need to combine three elements: written terminal examinations (short SATs); course work constrained by formal guidance (long SATs) and continuous teacher assessment.'[7]

SEAC would need to advise Ministers of the proportions of each in the assessment procedure for each subject. The minister would use this advice when formulating each assessment Order.

Within six months of the first steering meeting all the agencies had produced their first SATs which were to be trialled in the Summer of 1990. It quickly became apparent that the members of the main steering meeting would find it extremely difficult to adequately review the quantity of materials being produced. Consequently an additional meeting was arranged at which each agency was

required to present the tasks which it had developed and trialled. Following the Summer trial all agencies were required to produce an extensive report to a structure determined by SEAC. This was the first evidence that SEAC and their steering groups could not maintain control. The agencies were submerging them with materials and statistics and consequently beginning to take control of the process. When they wished to, the agencies collaborated over specific issues; jointly they presented a powerful cabal and could, when required, present a united front. The strength of argument, based on genuine expertise and growing research evidence, could only be countered by an autocratic steering process.

Main steering - September 1990 to August 1991

The first steering meeting of the second cycle focused on the trialling reports produced by the agencies. It began to emerge from these reports that there would be significant problems relating to the time taken to administer the tasks. Schools had commented on the problems which this created in relation to Y9 pupils and the way in which they took decisions about their key stage 4 programme. The proposal was that the assessment tasks should be completed before May 31st, leaving time for assessment, moderation and reporting before the end of term. This however, might create a legal problem as the Educational Reform Act had stated that SATs should be taken at or near the end of the key stage; legal advice would need to be taken on what constituted near the end of the key stage. This meeting reviewed the relationship of SAT and TA assessment and decided to recommend:
'that the SAT result should be preferred to TA where there are two results for the AT.'[8]
This was the start of the move to give priority to SAT scores. In addition the need to adopt common approaches across agencies and to keep instructions to teachers to a minimum was emphasised. Manageability was clearly an issue common to all subjects. The issue of stranding was also discussed. All agencies has used strands to make assessment more manageable, however, in common with the Orders and the non-statutory guidance, their use varied from subject to subject. Consequently it was concluded that a common policy on the use of strands could not be produced. The variety of approach led to SEAC reaching the following conclusion:
'The KS3 agencies' report on the 1990 trial exercise have made it clear that the existing specification (issued in March 1989) is rather too general and contains too many imponderables to be of much direct assistance to the agency required to produce SATs for the first full assessment (national pilot) in 1992.'[9]
This started to create an air of uneasiness amongst the development agencies as a new specification would mean, at the least, a re-negotiation of contracts.

The second main steering meeting concentrated on the final evaluation of the

Summer trial. The representative from the Department of Education and Science commented that:
'...current ministerial thinking is that SATs should meet requirements of public confidence and manageability, which was taken to include economical use of teaching time. They expect assessment to be in the form of short sharp questions, unless clear arguments exist for an extended assessment. The arguments for an extended assessment in design and technology are generally accepted. Nevertheless he looked to the pilot to investigate approaches employing short tasks wherever appropriate, and especially when addressing knowledge and understanding.'[10]

This sharpening of attitudes was apparent when the recently appointed Secretary of State for Education and Science, The Rt. Hon. Kenneth Clarke, responded in his first interview with The Times newspaper:
'I thought 'tasks' was a typing error for tests, I propose to go on using the word tests.'[11]
The Secretary of State elaborated on this view when addressing The Society of Education Officers in January 1991:
'Like key stage 1, assessment should combine some of the teachers' own accumulated judgements of pupils' classroom work as well as the results of end of key stage tests. I am persuaded however that the process will be manageable and the results will command more confidence if the tests are mainly in the form of short written tests. Tests on this model will provide a fair and objective measure of pupils' abilities. They will be straightforward for teachers to conduct and mark.'[12]

This was confirmed in a response to a Parliamentary Question from James Pawsey, MP for Rugby and Kenilworth. He asked the Minister:
'What local arrangements does he envisage for administering the assessments of 14 year olds.'[13]
In his response, The Secretary of State, The Rt. Hon. Kenneth Clarke stated:
'...We have also signalled to SEAC that, in the interest of rigorous and manageable assessment, the end of key stage tests which contribute to the overall assessments of 14 year old pupils should take the form of written terminal examinations except where the nature of the subject matter can clearly be shown to justify some element of practical or project work.'[14]
The Department of Education and Science press release (22/91) gave examples of practical work:
'.... technology or exploratory aspects of mathematics and science - clearly demand some element of practical work.'[15]

By March, when in the case of design and technology the pilot was already underway, SEAC held a review meeting at which officially all SAT material would be vetted prior to production. The impossibility of completing this task in a day was apparent to the agencies but obviously not to others. The SEAC Chief Executive, Philip Halsey, also had the task of presenting the pilot material to The Secretary of State and the Schools Minister, The Rt. Hon. Timothy Eggar; he had a thirty minute meeting in which to accomplish the task. The SEAC Council and steering groups understood that complex subject Statutory Orders required complex assessment procedures, and realised that simple tests would trivialise the whole notion of the attainment targets and criterion-referencing. The Secretary of State and his Minister did not. It was apparent that whatever occurred during the pilots, the process was, from the political perspective, out of control and firm action would need to be taken. At the end of June, in the middle of the pilot analysis and the production of the report all agencies received official confirmation that the clause which allowed either party to terminate the contract in three months without reason was to be invoked by SEAC. A new specification would be produced which would be advertised nationally. The Secretary of State in announcing these changes emphasised the need for simple paper and pencil tests and indicated that the expertise of the GCSE Examination Boards would be vital to future success.

Subject steering December 1989 to August 1991

Subject steering meetings had more impact on the style and content of the SATs developed and the interpretation of the standing Orders. The initial meetings took place prior to the publication of the technology Order and this forum was therefore important as it allowed NCC representatives to inform agencies of the final changes taking place. These meetings also indicated from the outset the difficulties and confusion which the Order was going to create. During this phase there were eight technology subject steering meetings. As there were two agencies much of the time during meetings was concerned with reporting and presentation of material. Although both agencies expressed a desire to collaborate and share information, as work progressed a competitive element entered into the proceedings. This was fuelled by the belief that only one team could produce the first statutory assessment so inevitably a choice would need to be made at some point. The agency based at Goldsmiths' College was still in the process of completing the final report for the Assessment of Performance Unit. This contract provided them with a firm base and a proven track record. Their philosophical approach to the subject and its assessment were well known via reports and other writing. Their reputation placed them firmly in the 'process driven' school of thought, designing was the key, skills and knowledge should be acquired on a 'needs only' basis. The team based at Middlesex were more practical in their

approach to the subject. Their reputation was based on the development of practical teaching resources and complementary INSET. This difference between the two, as commonly perceived, was captured by HMI Hicks at the first steering meeting:

"I have great confidence that this work will be successful as we have the best thinkers (CATs/Goldsmiths) and the best doers (MEGNAP/Middlesex) in design and technology; together they will make a success of the job.'[16]

Even in the first meeting the topics discussed would be those which would keep recurring during the next two years. How long should a SAT be? The time allocated for a SAT varied considerably; CATs first tasks were based on nine hours whilst MEGNAP had opted for extended activities of up to twenty hours. When should the SAT take place? Legally at the end of the Summer term, but was this feasible if a task lasted twenty hours? Another focus was the need to produce tasks which covered all ten levels, even though the programme of study only covered levels 4 to 7. Politically there was a view that some children would be capable of achieving the highest levels even though they had not been taught or experienced the substantive skills and knowledge contained in the programme of study. The problems of assessment activities using criterion referencing, problems arising from the different number of statements at different levels and the nature of the exemplar material were all issues which would re-emerge on a regular basis. No conclusion was reached on any of these matters.

The second meeting started to articulate the need for the statements to be exemplified, if teachers were to be able to make sound and consistent judgements. In the discussion relating to exemplification it was minuted that:

'.. all SATs should be designed specifically to test design and technological capability as a whole.'[17]

This clearly pointed agencies to single tasks from which assessments of all four attainment targets could be obtained. The issue of exemplification was central at the third meeting. The problems concerned with this issue were summed up as follows:

'... SoAs can be categorised into two groups - 'active' and 'reflective'. The 'active' statements can be identified as those requiring students to 'carry out', 'discuss' etc. a certain task whilst the 'reflective' statements require students to 'review' and 'justify' an activity'..... evidence of the 'reflective' statements is problematic given the ephemeral nature of the statements. This difficulty is further magnified by the fact that the number of 'reflective' statements far outweigh the 'active' statements.'[18]

In addition it was noted that the demands of Te1, a novel approach for the majority of teachers, meant insufficient time was being spent on the remaining three

attainment targets. Poor interpretation of the statements by teachers and confusion over how they should be applied, along with these factors, contributed to the lack of good examples which could be used to exemplify the statements. A procedure for focusing pupils' activity so that they naturally addressed the statements of attainment and provided evidence was required. Both agencies reported the development of assessment 'probes' as a means of achieving this. The approach however was very different. CATs were employing reflective question papers which were answered at regular intervals during the task whilst MEGNAP used an active approach requiring pupils to incorporate a set of labels (headings) in their project folder.

Independently both groups had deconstructed the attainment targets into strands of competencies, although the strands in relation to Te1 and Te4 were different. Although the procedure advocated was different there was agreement that the focus of any assessment task should be the 'levelness' of a pupil's capability. Both agencies had produced an assessment matrix as a means of organising the statements for teachers. MEGNAP had in addition modified and revised some of the statements to create clearer and smoother lines of progression. This was clearly contentious. Members were sympathetic to the logic behind the proposals, but were unsure of the legality of the approach. They decided to remain undecided. By the fourth meeting MEGNAP had also introduced another approach, one which collapsed all of the statements at a level to produce a single coherent level descriptor. Hence three approaches were being investigated: the raw SoA, the stranded matrix with modified SoAs and the level descriptors. The validity and reliability of these three approaches was an important aspect of the Summer trial. Following the trial the agency reported that of the three the raw SoAs produced the least reliable results and was the most difficult to use. However, at the sixth meeting the NCC raised objections to any assessment procedure which was not based totally on the legal statements. Between the sixth and seventh meeting negotiations took place between the agency and the NCC. The outcome, which was used in the 1991 pilot was a stranded matrix which only used the statements of attainment and did not modify them in anyway. The agency also sought permission to pilot the single level descriptor; this was reluctantly agreed. It was, however, this approach which was eventually used in the first statutory assessment.

The steering committee assembled by SEAC had a distinctive CDT bias, consequently, SATs which offered opportunities to other partners in the design and technology federation were questioned. An example being the SAT 'Public Place' developed by MEGNAP. A position adopted by some members was that the result of a design and technology activity should be an artefact and that models should

not be encouraged. The thrust behind this position was that Te4 - evaluating - could only be addressed if the outcome could be properly tested. This discussion was never fully resolved. Because of issues such as this it was apparent, even by mid-1990, that the statutory Order could not be operationalised without greater clarification. To solve the immediate difficulties the agencies were given the task of defining the terms used in the Order because they were open to interpretation; the assumption being that the agencies would be able to agree!

The fifth meeting, following the first large scale trial, started to focus on the technical issues relating to assessment and in particular aggregation. How would a subject score - profile component - be arrived at from the four attainment target levels and prior to that how would statements be aggregated to produce an attainment target level. Comparability was the key issue for SEAC, consequently they wished to employ the rules which had been adopted at key stage 1 and employ these for all subjects. Strong, but different, arguments were mounted by the two agencies why this should not be the case. CATs wished to employ procedures which:
'must have the effect of focusing teachers' attentions on pupils' weaknesses and not their strengths.'[19]
In practice this meant applying a 'trailing edge' principle which would allow the PC level to be only one level above the lowest AT level. MEGNAP were employing both combination and aggregation rules, see appendix 3.1, page 263, with the objective that strengths in a particular AT should be reflected in the PC. In addition they stated strongly that an n - 1 rule (key stage 1 rule) for determining attainment target level was totally impracticable given the wide variation in the number of statements at various levels.

The final meeting, prior to the agencies receiving their termination of contract notice, dealt with reviewing the materials for the 1991 pilot. It was evident that the approach adopted between the two had widened. CATs' approach had become more academic and consequently imposed more management demands on teachers. The objective was to produce more concrete evidence of performance but it was clearly distorting the activity and shifting the balance away from practical achievement towards recording. MEGNAP had attempted to streamline the process by simplifying and eliminating superfluous material. The Director of the MEGNAP agency was minuted in the following terms:
'SAT materials needed to be accessible, manageable and realistic in the amount of reading matter that could be reasonably considered by teachers. For reasons of reliability a simple, unambiguous approach was required.'[20]
This clearly reflected the mood of the time:

'However by mid-1991 it became clear to both design and technology teams that the sophisticated, complex SATs they were developing were unlikely to be acceptable.'[21]

The Summer of 1991

The Education Reform Act had been master-minded by the Rt. Hon. Kenneth Baker, but he had departed to be Home Secretary in 1990, being replaced by the Rt. Hon. John MacGregor. He adopted a relatively low profile and allowed the introduction of the National Curriculum and the accompanying assessment arrangements to proceed without interference. His stay at the Department for Education and Science was short lived. Following the resignation of Sir Geoffrey Howe, The Prime Minister, Margaret Thatcher, was forced into a mini-reshuffle of her cabinet. Surprisingly, not least to himself the person sent to Education was Kenneth Clarke. He told The Times in an interview:
'I told my staff (in the Department of Health) the night before that the two cabinet ministers who would not be moved, because of the reforms going through, were John MacGregor and me.'[22]
His approach was the opposite of MacGregor's, he appeared determined to distance himself from the reforms of his predecessors. As Duncan Graham (1993), Chief Executive of the NCC noted
'Clarke came to the Education Department prejudiced against the National Curriculum Council and the Schools Examination and Assessment Council, viewing them rather quaintly as creations of an opposition government certainly not his own.'[23]

The Secretary of State's desire for change reached fruition in July 1991. Within days of writing to SAT development agencies terminating contracts, Philip Halsey at SEAC left his post. This has been preceded, by only a few days, by the departure of Duncan Graham from the NCC. Both Chairman and several council members at both organisations also departed in the same month. The Secretary of State now had the opportunity to make his own appointments and to try, like others before and since, to take a firm grasp of the curriculum. In relation to the development of statutory tasks it now appeared that the work undertaken in the first two years was just, 'water under the bridge.' The process could begin afresh with new teams working to a new specification. This specification embodied the right wing approach to testing - short sharp pencil and paper tests. No mention was made of the Task Group on Assessment and Testing, the ground rules had been rewritten rather than revised.

The GCSE Examination groups believed that the Secretary of State wished them

to play a more central role in the development work. They also believed that he wished the work to be taken away from educationalists in departments and institutes in higher education. Consequently, the examining groups severed all their links with their former partners and created their own consortium. Each group bid for one of the five subjects, namely: Northern Examining Association (NEA) - English; University of London and East Anglian Group (ULEAG) - mathematics; Midland Examining Group (MEG) - science; Southern Examining Group (SEG) - technology, and Welsh Joint Education Committee (WJEC) - Welsh. The specification required leaner agencies working to far more specific goals and within a far tighter structure. It was also quite apparent that no more than one agency would be working on each subject. Following the submission of a tender document (appendix 3.2 provides extracts, page 268), short-listed agency directors were interviewed. In the case of technology 5 directors were interviewed. The complexity of the issues in this subject area made it very difficult for agencies not previously involved in the development work to offer a viable proposition, given the short period of time in which trialling materials had to be developed. Consequently the selection committee in effect had to decide between the CATs consortium based at Goldsmiths' College and the previous MEGNAP group based at Middlesex Polytechnic. The contract went to Middlesex, the sole Director of the successful agency being the author of this thesis.

There were possibly a number of factors which affected this decision. The author believes that the main reason was that the more pragmatic, practical approach adopted by MEGNAP, based on their view that the manageability of the procedure was crucial if it was to be reliable, was the key to the award of the contract. During the final interview the Director had also articulated the view that the agency would, if successful, adopt the same procedure as a design consultancy. That is, it would conduct the research, present the results to the client, advise the client on what in their view would be the most successful approach but naturally allow the client to make the final decisions about the parameters of the product. The interviewing committee had warmed to this model of an open working relationship. It may have been in contrast to others who had strongly held philosophical views about the nature of testing which might lead to conflict between the agency and SEAC.

The events of this three month period were borne out of politicians frustration that an education system does not respond immediately to the passing of legislation. Politics and education operate on very different timescales, education is notoriously slow to change and teachers are frequently reactionary in their approach. It is difficult to convince them that the tried and tested approaches are not the best and should not be tampered with - possibly their scepticism was right! For ministers the testing of 7 year olds (key stage 1) had resulted in a great deal of

hostility from teachers which had been fanned by stories in the tabloid press of 'test stress' amongst this age group of children. The response of teachers to what they saw as over complex classroom tasks which 'only told them what they already new' allowed ministers to distance themselves from the National Curriculum and the assessment arrangements. As Graham (1993) noted:
'Delivery or ridicule faced NCC and SEAC at every turn. Perhaps the staff suffered even more. They were given little time to settle in and adjust before having to produce detailed reports and proposals in the full glare of publicity.'[24]
Both organisations realised that they were now on trial. It was being openly discussed that they should be merged and this was their last opportunity to deliver what the politicians required.

The political and media spotlight now shifted from key stage 1 to key stage 3. The new chairman of SEAC was Lord Griffiths of Forestfach, a former Professor at City University, and a close confidant of the Prime Minister. This was a political appointment as he undoubtedly shared the views of the Secretary of State. It was clear that his task was to ride roughshod over the educational opposition and fully implement the Government's testing and reporting policy. The award of the new contracts left secondary school teachers in no doubt that the remorseless advance of the National Curriculum assessment procedure would soon reach them. From September 1991 onwards there was scarcely a day when some aspects of the Government's educational reforms was not mentioned in the daily press.

The successful key stage 3 agencies were informed on the 27th September and this was followed by a press release on the 16th. October. By the 21st October schools were embarking on the first trials under this new contract. Of the five new contracts, those for English and science went to examination groups, CATs retained the mathematics contract and NFER the Welsh. The total contract value was £3.161,851 approximately one quarter of the total awarded in 1989. All these agencies proceeded through until the first statutory assessment in their subject with the exception of the English agency whose contract was terminated after 12 months. The press release gave the following information:
'The tests will be timed and taken under controlled conditions and at the same time throughout England and Wales. The tests will cover most, if not all, the attainment targets for each subject, and at least half of the content within each attainment target. All questions will be compulsory. They should not require pupils to answer questions significantly below their ability but should enable them to demonstrate their best achievement against the 1 - 10 national curriculum scale.
They will be marked by teachers against clear guidance, which should ensure that standards will be nationally comparable.'

Subject specific information		
	Attainment targets covered	**Length of tests**
Technology	1, 2, 3, 4	11/2
	5	1

In relation to Welsh, the press release noted that there would be longer oral and written tasks spread over a three week period. No special mention was made of technology, so clearly the need for a longer, statutory practical task in design and technology had, for the time being, been dismissed.

Research and Development September 1991 to August 1993

The research in design and technology which attempted to meet this specification was based on the premise that all pupils would firstly undertake a prescribed extended long task which was non-statutory and that the statutory test would be based on pupils' experience of completing this task. Unhappy with the notion that the end of key stage level of capability would be based on a written test the agency, via trials, attempted to provide evidence which would question the validity of such an approach. SEAC on entering this new phase of the work had decided to abolish the steering apparatus. A new assistant chief executive had been appointed specifically to oversee key stage 3. He had formerly been the history subject officer at the NCC. His approach was less formal and involved meetings between agency staff and SEAC officers, decisions taken were then included in a letter of the meeting which noted the action to be taken. The chairman of the SEAC technology subject panel was also involved in the process for the first time via a number of working/social meetings with the agency. The first formal meeting did not place until December when many decisions had already been taken by both parties. In particular at the beginning of December the Chairman of SEAC wrote to the Secretary of State offering advice on the assessment of fourteen year olds in technology. This advice, given without consultation with the development agency, was fully accepted. It radically changed the assessment procedure and undermined the philosophy of design and technology capability enshrined in the Statutory Order. The letter contained the following advice:

'We have been giving some thought to how an assessment Order might refer to these two elements (statutory written test and non-statutory long task). We now believe it might be preferable to make the long task statutory. A sharply focused long task, prescribed in an assessment Order, would ensure that pupils were prepared for the test on the basis of appropriate activities. It would make possible an overall score for design and technology based on rigorously assessed practical activities as well as on a written test.

'Our recommendation therefore is that the 1993 key stage 3 assessment Order refers to a long task in design and technology, and a short test. The task would be based on Te2 (designing) and Te3 (making). It would involve real materials (sic) and the making of an artefact. It would take 10-12 hours and be done during normal lessons.

The test would assess pupils' ability to evaluate their own designing and making (Te4) and, on the basis of this, identify new needs and opportunities (Te1). It would also test the knowledge elements of the attainment targets and the programme of study.

The SAT as a whole would ensure that end of key stage assessment at 14 is based on the kind of rigorous design and technology activity which we are all keen to promote.'[26]

The reply from the Secretary of State included the following passage:
'I believe strongly that tests for 14 year olds should take the form of written tests unless there is a strong case for requiring pupils to undertake a more extended practical activity. I agree with you that there is a strong case for this in design and technology, and therefore agree with your proposal that such a task should be included as part of the statutory tests in 1993. I am happy therefore for you to proceed as you propose, on the understanding that the task takes no longer than the 10 to 12 hours envisaged in your letter.'[27]

(It is worth noting that SEAC had been threatened with legal action for using the acronym 'SAT'. So from December 1991 onwards the term ceased to be used in publications, however the term is still in common usage.)

Without this knowledge the agency had established their trialling based on a different premise. Even in the first trial the problems associated with pencil and paper tests were beginning to emerge. For example, despite the remark to the contrary in the letter of the SEAC chairman to the Secretary of State, the attainment targets have no knowledge requirement. This resulted in questions being framed which it appeared could be answered by general knowledge. This was particularly the case in relation to Te1 and Te4, those which were now to be the focus of the written test. It had also become apparent that pupils taking the test and teachers marking them recognised that as the highest level of achievement would be the recorded level, there were significant advantages in working backwards through a paper. In relation to 'tick back' when marking, this might result in a high percentage of pupils' work remaining unmarked. This would further reduce the tests' educational credibility for being formative in addition to

summative. The problem of pupils 'dropping off the bottom' of a tier and failing to record a level was also an important issue.

The decision to split the assessment of the attainment targets remained confidential until March of the following year. Then the correspondence outlined above was leaked to the Times Educational Supplement, which ran a main feature titled, 'Technology test leak raises fears.' The article pointed out how this decision contradicted advice from the Assessment of Performance Unit and SEAC's own document, 'Pupil's work assessed.' Prominent design and technologists were quoted in the article, for example, Professor Richard Kimbell:
'... this decision will be desperately damaging to the development of technology in schools.'
Professor John Eggleston:
'SEAC have been influenced by the recent debate in design and technology about returning to traditional values.'[28]
The genuine concern expressed in this article was that dividing theory and practice would hasten a return to the practice of labelling certain subjects academic or non-academic.

Design and technology in schools - Blue Peter or Mickey Mouse?

The debate referred to by Eggleston commenced during the Autumn of 1991 when concerns begun to surface about how the National Curriculum in design and technology was being implemented. National papers were beginning to express concern, The Guardian amongst others reported:
'Blue Peter-type activities involving cutting up cardboard are taking the place of lessons in the use of technology and industrial machinery under the Government's National Curriculum'[29]
These fears were based on a report of the National Institute of Economic and Social Research. The report claimed that:
'the new curriculum, designed to improve technological skills, is concentrating on "pretentious design briefs" and getting pupils to write about solving problems rather than on teaching them the basics of technology.....In contrast to the current emphasis in Britain on pupils' activities in 'problem solving' and 'design' in the teaching of technology, the emphasis on the Continent continues to be, distinctively and deliberately, on craft skills in an industrial context: problems in design, it was often said to us on the Continent, can only sensibly be tackled after mastering the properties of the relevant materials and tools.'[30]
This was just the opening salvo in a debate which is still continuing about the nature of technology.

These concerns had a direct impact on the nature of statutory assessment. As a result SEAC was becoming extremely sensitive about the substantive content of practical tasks. For example, the task being trialled in the Autumn of 1991 involved batch production. Through five different tasks, in five different materials, this theme was explored in relation to raising money at a charity fair. The theme was taken further in the written test by asking pupils to explore the issues relating to the production of up to a 1000 badges for a flag day. The SEAC response to this test was:

'There is concern about the use of badges as a theme; it might appear to some to be trivial.'[31]

In response the agency director replied:

'The notion that the theme is badges is actually incorrect. The theme is batch production i.e. manufacturing items in quantity. The test puts pupils into the situation of having to produce items in large number - possibly 1000s if they are to be sold during a flag day. One has to consider the complexity of an item which can be produced by 13 to 14 year olds with reliability and precision in these quantities... Alternatives could be no more complex because the problem revolves around the manufacturing issues.'[32]

This example serves to highlight concerns and sensitivities about the necessary rigour attached to any task. The need to motivate children became secondary to the need to satisfy those advocating a return to a curriculum focused on skill competence. This was summed up by Eggleston (1991) writing in the Times Educational Supplement. After comparing Technology to the brightest, fastest growing flower in the National Curriculum, he continued:

'Yet as the Summer passes the flower may be fading on the stem. Faced with the cold blasts of a new right winter there are fears that the first blossom may be its finest.'[33]

The hardening of attitudes as to what kind of activity was suitable for long tasks in design and technology continued. In February 1992 the agency made its first proposal for the first statutory assessment in 1993. Following the procedure which had been accepted nationally, an over-arching theme and a context were proposed along with five tasks, one in each of the material categories. The theme was 'keeping-fit' and examples of the individual tasks were a device for measuring lung capacity as a means of comparing fitness; an adjustable timer for recording exercise repetition; a garment for carrying exercise weights and a high energy food bar. The minuted response was as follows:

'The design and technology context 'keeping-fit' has not been well received. We need to give further consideration to the choice of an appropriate context. As you are fully aware, the choice of contexts raises fundamental questions about the

nature of D&T which we will need to address. We need to ask if there is a need for a context?[34]
This is yet another example of how the assessment process was modifying the non-statutory guidance, which was central to the way in which schools were attempting to deliver National Curriculum design and technology. Schools were pursuing what they believed to be a legitimate approach whilst policy makers were surreptitiously changing the ground rules.

During 1992 the main protagonists, politicians and the teaching unions, started to raise the heat of the debate. The first statutory tests in maths and science had been downgraded to a pilot in which all schools were invited to participate. This situation resulted from the necessity to revise the statutory Orders in these subjects, the outcome of which was to reduce the number of attainment targets in each from 16 and 17, respectively, to 5. This represented the first slippage in the National Curriculum implementation, but from this point onwards the ministers and civil servants were on a downward slide. Compromise followed compromise in relation to both the assessment procedures and eventually the content of the curriculum.
When the plans for the 1992 tests were announced they were greeted in the following terms. Nigel de Gruchy, Secretary of the National Association of Schoolmasters and Union of Women Teachers commented:
'... schools should not volunteer to take part in the tests unless they are given cast-iron guarantees about financial and staffing support. Teachers will be marking the scripts of public examinations, and when this happens at GCSE they will be paid.'[35]
The National Union of teachers claimed that:
'...the tests will take up an extra 22 hours of each teachers time with no benefit to pupils or their parents.'[36]
Jack Straw, Labour's educational spokesperson observed:
'Almost every secondary school in this country already conducts tests, and there is a danger of duplication between those used to determine which GCSE options pupils will follow and these new national tests. The new attainment targets that will be tested have not yet reached schools, so this announcement points to chaotic administration by the Government.'[37]
Very few of those involved in the administration and marking welcomed any aspect of the testing procedure.

Following the general election which returned the Conservative Party to Government for a fourth term, the Rt. Hon. John Patten was appointed Secretary of State for Education in succession to Kenneth Clarke who became Home Secretary. The legacy left Patten by his predecessor was of policies failing amidst increasing teacher confrontation. His period of office lasted just over two years

during which time he managed to exacerbate many of the tensions and conflicts which he inherited. Clarke's failure was to reduce complex matters to the grotesquely simple as a means of gaining support for Government reforms against the 'trendy teachers'. The interpretation, by Clarke, of the first national tests at 7 was that one in three of the age group could not read; this was clearly not the case, but it was an effective slogan for urging acceptance of the Government's reforms. Every new Secretary of State on taking up the post insists that their first task is to listen to teachers' concerns before reaching decisions. John Patten took this position to the extreme, making no comment in public for several months after his appointment.

A month after the election the debate concerning the nature of design and technology reached a new dimension with the publication by The Engineering Council of a report entitled, 'Technology in the National Curriculum - Getting it Right'. The report was based on research conducted by Professor Alan Smithers and Dr Pamela Robinson at Manchester University. The report started with a phrase which was quoted extensively by the national press:
"Technology in the National Curriculum is a mess.'[38]
The headline in the Daily Telegraph was typical: 'There's a spanner in the curriculum.'[39] This view was supported by the HMI annual review of secondary schools which reported that 40% of lessons observed in design and technology were either poor or unsatisfactory. This was based on a survey of 884 schools in 95 authorities. The report expanded on this poor quality as follows:
'... standards were lower in classes using the national curriculum, which was introduced to improve academic performance, than in those where it had not been implemented.' In relation to 10% of schools in the sample, it stated: 'Low standards went unchallenged, work lacked rigour, progression was weak and pupils were poorly motivated.'[40]
Again this captured headlines in the national press. Indeed if press coverage was a means of judging the changing status and importance of the subject then during this period technology was second to none. The National Curriculum Council produced a document outlining the arguments for a review of the technology Order.[41] This was based extensively on the evidence produced in the HMI report. It could equally have used evidence produced by the SAT Development Agency.
The pressure mounted and on the 2nd June, The Secretary of State announced the order which had long been expected to the House of Commons:
'Her Majesty's Inspectors will conduct an urgent review of technology in the National Curriculum. The Inspectors will review the technology Order of the National Curriculum with a view to increasing teachers' expectations of pupils, specifying more clearly the skills and knowledge which they should acquire, giving more emphasis to the practical, and improving how the subject was managed in the classroom.'[42]

This announcement was the first of many to create confusion in technology education. In effect the Secretary of State's announcement recognised that the National Curriculum Orders were flawed, yet in four months time new GCSE courses were to be introduced based on this deficient Order. The extent of the revision would become apparent as a stream of reports and draft orders followed over the next three years. The review was intended to be short and sharp but other events overtook the parochial needs of technology. Many saw this step as a triumph, particularly The Engineering Council. They believed that they had been instrumental in this decision and in a press release their Director, Denis Filer, commented:

'This should put an end to the Micky Mouse technology that has manifested itself in schools.'[43]

The Times Education Supplement echoed the fears of many teachers in an article entitled 'Technology U turn starts'[44]. This predicted a return to the traditional craft approach and the removal of food, textiles and design from the Order. This article created an immediate response in high quarters. The next week's letters page contained denials of this intention from David Pascall, Chairman of the NCC and from The Baroness Blatch, Minister of State at the Department for Education. Additionally Dr John Williams of The Engineering Council defended the Council's position and reiterated that it did not wish to return to a craft focus in technology. [45]

The first statutory assessment

These events took place months before the first statutory assessment at key stage 3 in design and technology got under way. This assessment had to be based on an Order that was in the process of being revised as this had been the basis of teaching and learning for these pupils. The dissenters, regardless of their view, now had a focus for their criticism. Whether they were craft-based traditionalists or anti-assessment unionists, they could jointly object to the assessment procedures having little value if it was seeking to assess something that was so seriously flawed that it required an urgent review. Against this backdrop hard decisions were finally beginning to be made about the nature of the first statutory assessment.

By the end of May the first proposals were being discussed. The informal meetings with SEAC officers had, within less than 12 months, been replaced by formal meetings with HMI and DfE representatives. As a result of one of these meetings SEAC steered the agency to abandon the notion of theme and context. They also requested the agency to:

'...consider the need for ensuring effective differentiation in the long task.'[46]

This would require activities which differentiated by task rather than just by outcome, and this was a radical departure from design and technology practice

and one with which teachers would be unfamiliar. As a consequence, it would require teachers to enter each pupil for one of four tiers, a task which they found difficult and unpalatable. A draft assessment Order was also circulated during May. This again challenged the commonly held concept of design and technology as it weighted the attainment targets. The proposed, and adopted weightings, were as follows: Te1 - 15%, Te2 - 25%, Te3 - 40% and Te4 - 20%. This firmly placed the emphasis on designing and making and the long task as it would now provide 65% of the end of key stage subject score. If the attainment targets were no longer equally weighted it would be difficult for teachers not to interpret this in terms of relative importance and time allocation in their teaching. This weighting had also been adopted for GCSE courses where its effect would be far more significant.

Following the national pilots of 1992 there was a concerted effort to unify the approaches being used across all the subjects. Principles were established for the setting of both questions and mark schemes. In reality other subjects were now wishing to follow the lead set in design and technology by the long tasks which were focused on the level of attainment rather than satisfying individual statements of attainment as had been the policy. For test papers the notion of mastery at a level was introduced. At each level a number of marks would be available and a number of these marks would be required to demonstrate achievement at that level; these marks could be accumulated in any way. The agency had also proposed a solution which would ensure differentiation from the practical task. There had been pressure to devise totally different tasks for each tier (task focused on only a range of levels, e.g. 3 to 6). This approach was resisted by the agency which, following research and trials, devised a system based on a generic design specification which increases in complexity band by band but essentially involves all pupils in tackling the same task. In addition there had been pressure to have tasks only in one material area, construction materials. This was to appease those critics of the Order who believed it lacked rigour and promoted design and technology through inappropriate materials. But a decision had to be made concerning the number and nature of the tasks. Concerns relating to reliability dictated that there should be as few tasks as possible. Yet manageability dictated that there should be one in every material area. However, this was counteracted by the required linkage between the long task and the test. If there were five long tasks and four bands, there would be twenty different test papers. It was already widely acknowledged that technology was by far the most complex subject in relation to assessment but this would have exacerbated the situation beyond reason. Politically there would have to be a task in construction materials. As end of key stage assessment was a legal requirement, a food task was essential as many all girl-schools only had facilities in this area. It was apparent that only two tasks, one in the traditional boys' area and the other in that of the girls' would not

be acceptable. Graphics media had, quite wrongly, been dismissed as paper and card technology. Textiles was politically more acceptable than food, but piloting statistics indicated that only 9% of pupils would take a long task in this material. The decision was taken to offer a third task in control materials. This term was not used in the statutory Order and had been introduced by the agency for the 1992 pilot to redress the lack of work taking place in this aspect of design and technology. It was also very in-tune with the demand that design and technology should have a more scientific bias. Following the decision to have three tasks the agency provided two possibilities in each area and SEAC officers took the final decision as to which should be fully developed into the first statutory tasks.

The long task material received Ministerial approval in October and was ready for despatch to schools at the end of November. Associated with this were three National Conferences aimed at informing Local Education Authority Inspectors and Advisers about the assessment process. By December all schools had received the documentation and the protests started. Initially the most vociferous lobby centred on textile teachers. They believed that the failure to include a textiles task signalled the imminent demise of their subject. Indeed the very opposite was true. During December the HMI proposals on technology were published and these provided textiles with a more secure position than food [47]. However, one can only sympathise with the mixed messages which teachers were undoubtedly receiving. A key issue throughout the past four years had been the level of teacher intervention whilst a pupil was engaged in a statutory task. This reached a head when schools received the long task material. The tasks had been based on particular aspects of the programme of study, for example, the construction material task was focused on mechanisms. Teachers wanted to know if they could revise or teach these aspects the week before the task commenced or whilst the task was in process. Many believed the task, lasting 10 to 12 hours had to be taken in examination conditions whilst others felt they should assist pupils to achieve the highest possible level. The reasons why were summed up by Shaw (1993):
'There is a built-in incentive on teachers to cheat because of the pressure of reported results and league tables.'[48]
This view was supported by a headteacher:
There would be no problem if teachers were allowed to get on with the Assessments which SEAC advise. The problem is created by the requirement of reported results nationally and the publication of them in league table form. That puts teachers under intolerable pressure to teach to the test which is the antithesis of what we want to happen.'[49]

Accusations were made by schools against other schools. The SEAC hotline, established to answer teachers' questions, dealt with as many complaints about

what other schools were doing as with genuine questions. Schools were accused of pre-task practising, providing children with drawings/circuit diagrams, producing highly informative wall charts and work sheets and of making jigs and moulds. In other schools teachers were pedantically using stop watches to insure pupils had every minute of their 12 hour entitlement, refusing to answer any questions other than with a yes no response, removing all visual information from the walls and refusing to allow pupils to consult reference books when seeking information on, for example, the performance of components or the nutritional value of foods. All these local responses were compounded by the production by commercial organisations of task kits. These were being advertised and sold through the education press by the time schools returned from their Christmas vacation. The issue of their legitimacy was raised by the national press. The Times Educational Supplement, front page headline read:
'Probe into pass kits for technology tests.'[50]
The article started:
'Government advisers are investigating the publication of a £25 "SAT Pack" for teachers designed to help their pupils pass the first national technology practical tests for 14-year-olds in June. They are also looking into a £20 kit containing teacher support material including circuits and resistors for the electronics practical tests in technology.....The breaking of the rules.....was called desperately unprofessional by a Schools Examination and Assessment Council source this week.'[51]

This episode serves to illustrate the anxiety, apprehension and suspicion which these tasks had engendered. In reality there were no rules to be broken, schools had been given guidance on how to conduct the tasks but it had been left to their professional integrity to operationalise this advice. This advice was broad and offered schools a degree of latitude so that the tasks could be accommodated with as little disruption as possible. The majority found this advice sensible but the minority appeared to want a far stricter and more formal approach. However, these were not 'life opportunity' examinations and they were not certificated, but some teachers appeared to believe that they carried far more status than GCSE.
'SEAC's unclear instructions and complete ignorance of the procedures normally carried out in an examination task of this type have created confusion.'[52]
Others perceived that this confusion might be interpreted differently by Ministers who would use it as evidence that practical tasks could not be operated with any degree of reliability:
'STEP (Staffordshire Technology Education Project) apparently sold £10,000 worth of material (SAT kits) up to the end of last week - a third of a salary of an advisory teacher. Has it been worth it when it has almost certainly contributed to the demise

of the long task and the use, instead, of more controlled written tests?'[53]
Standard tasks had been required as the statistics which they produced would clearly have been more reliable for comparative purposes. However, such tasks rely heavily on the integrity and goodwill of those conducting the tasks, and goodwill towards 'Government tasks' was in very short supply. The press both national and subject specific set out to trivialise the tasks. For example, the control task which required pupils to make a miniature warning beacon was dubbed: '...an illuminated tent peg...'[54]
Perhaps it is best to leave the final comments on the long task to a 14-year-old pupil who wrote to the Times Educational Supplement:
'...This exciting task is part of our National Curriculum technology test. It's fairly harmless, as far as I'm concerned; I can do it.....It's an unrealistic situation; no talking is allowed so no one can discuss ideas, offer constructive criticism or help anybody. Cries of 'he's copying me' ring around the otherwise silent classroom - we are silent because either we are proving how wonderful we are, or we haven't got a clue what is being asked of us.....but I really feel sorry for my CDT teacher. The thing is, he actually isn't allowed to teach us anything. 'Erm, yes it's like this, look, no! I can't tell you, but watch this, watch!' sketching wildly with his hands..... So the position is, we are spending 12 hours, which amounts to five weeks of CDT lessons not being taught. It's probably happened before, this 'not being taught' and maybe it won't do us any harm. But I'm not so sure. Oh and another thing. If anyone can find an export market for thousands of clamps 18mm long which can hold an object 1mm thick, please let me know. We could be on to a winner there.'[55]
The sophistication of the letter makes one suspicious of its origins! Indeed, it also over estimated the capability of 14 year olds as only a handful of those who completed the task produced a clamp which half worked.

Conflict with the Unions

The long task got under way in most schools and the majority completed it. However events in other subjects, notably English, had a far reaching effect on the first statutory assessment. English had always been the 'hot spot' of the National Curriculum. There was significant disagreement over what should comprise the programme of study. Teachers in general approved of the Order developed by a working group chaired by Professor Brian Cox, but it did not satisfy the reactionary, right wing politicians who demanded a far greater focus on basic skills. The first statutory tests proved to be the battleground. Once the sample test questions had been published The National Association of Teachers of English orchestrated a campaign against the proposed tests. They criticised both the quality and purpose of the questions. The questions were hastily withdrawn and new ones published. This was the catalyst which united the anti-SAT campaign. One by one the main

teaching unions conducted ballots amongst their members and one by one they advised their members not to cooperate in either the invigilation or marking. The reasons for the union boycott differed from union to union. The NUT were philosophically opposed to the notion of National Curriculum testing, whilst the NASUWT opposed the testing on grounds of workload. This reason for non-cooperation was tested in the courts by the London Borough of Wandsworth. When the courts found in favour of the Union it effectively brought to an end the first statutory assessment in all subjects.

By the beginning of June, when the tests were meant to take place, the confrontation between The Secretary of State and the unions had spread even to the previously moderate National Association of Headteachers. When he spoke at their annual conference on 2nd June his speech was delivered to:
'...a chorus of disapproving laughter, cries of "shame" and "rubbish" and muted hissing.' [56]
This was at its loudest when he announced:
' ...next year's tests will take place in May.....testing and publication of results must be a factor in justifying education expenditure. Without a testing regime or publication of results, neither government, taxpayers, employers or the wider community will have any idea of what is happening in schools.' [57]
In response the President of the association spoke on behalf of many teachers and not just head teachers:
'The future is in our hands whether we can change the legislation or not, because by resisting, we will refine; by subverting, we will redirect; and by protecting we will create; and in doing so, we will remain loyal to the principle that the learner comes first.' [58]
The cost of the largely abandoned testing procedure was now released:
'More than £35 million is being spent on this years's crippled school tests for 7 and 14 year olds, a government adviser said yesterday.....£23 million has been spent by central and local government to ensure consistency in marking.....Materials for the tests have cost £6.5 million, while up to £6 million has been earmarked for auditing the 14-plus mini-exams.' [59]
The eventual outcome was the Secretary of State's decision to establish the Dearing review of the National Curriculum and National Testing procedures.
Whilst this story unfolded the design and technology practical tasks had continued. The majority of teachers, however, decided not to report marks to Audit authorities, if indeed they decided to assess their pupils' work. Consequently, the standard of work achieved was never objectively established. HMI conducted an internal inspection of the long task and produced a report which has never been released. However an MP, the Rt. Hon Tim Boswell, revealed some of the content when

writing to a school in his constituency:
'..... the preliminary evidence from OFSTED is encouraging. On their inspections so far, the inspectors have found that most pupils are motivated by a clearly defined task and that some pupils are already exceeding their teachers' expectations.' [60]

SEAC were also charged with the production of tests despite the likelihood that they would never be taken. Consequently, trialling continued through to March 1993 and the scrutiny and modification process became even more rigorous. All those concerned were determined that the final products should, within the context, be exemplary. The research and development undertaken to produce these statutory tasks was detailed and extensive even though the conditions were frequently fraught. It is hoped that the evidence which follows records in a permanent manner this experience and that this research and development of assessment procedures in such a context and for these purposes will prove of value to others tackling such a task.

Conclusion

The involvement of the Government in testing schoolchildren was unique. For the first time the responsibility lay with a single authority, the Schools Examination and Assessment Council, rather than with a number of independent examining bodies. Consequently, schools had no scope for choice as all aspects of the procedure were compulsory. The Government's approach was mirrored by the organisation it charged with the task, which had no previous experience as an examining body. The examination groups, which had a proven record, were only involved indirectly in the process. Initially the task of producing the statutory tasks was seen as a one for academic researchers. In this climate of compulsion and polarisation it is doubtful if any organisation could have fully achieved the Government's expectations.

The total cost of the research and development, for all aspects of the assessment process, has not been revealed. But between 1989 and 1993 it was probably in excess of £90 million, and this figure is additional to the cost of implementing the National Curriculum. This seems a large sum but it is approximately, £3 per school pupil per year. The four year process was undoubtedly a formative period for all those involved. The administration of the time, and ones for the foreseeable future, will have realised that it is extremely difficult to impose, control and assess the curriculum centrally without the goodwill of those who have to implement the legislation. It is equally important that the parents of the pupils involved have been convinced of the benefits. The Government was more than surprised when it became apparent that for once public opinion, and eventually the force of law, was on the side of the teachers.

References

1. Schools Examination and Assessment Council (1989). 'Specification for the Development of Standard Assessment Tasks for pupils at the end of the third key stage of the National Curriculum'.
2. Midland Examining Group National Assessment Programme (1989). 'Proposals for Development of Standard Assessment Tasks for pupils at the end of the third key stage of the National Curriculum'. Cambridge: Midland Examining Group
3. Schools Examination and Assessment Council (1989). Contract CGRM/36/1, Issued to Middlesex Polytechnic Higher Education Corporation
4. Schools Examination and Assessment Council (1989). Contract CGRM/36/1, Issued to Middlesex Polytechnic Higher Education Corporation
5. Schools Examination and Assessment Council (1990). Arrangements for the 1990 Trial Exercise
6. Schools Examination and Assessment Council (1990). Minutes of the second meeting in 1990 of the KS3 Development Steering Committee.
7. HMI Marshall, C. (1990). Minutes of the second meeting in 1990 of the KS3 Development Steering Committee, Schools Examination and Assessment Council
8. Schools Examination and Assessment Council (1990). Minutes of the fourth meeting in 1990 of the KS3 Development Steering Committee.
9. Schools Examination and Assessment Council (1990). Programme of work for KS3/DSC - 27, 1990/91
10. Schools Examination and Assessment Council (1990), Minutes of the fifth meeting in 1990 of the KS3 Development Steering Committee.
11. Clarke, K., Secretary of State for Education (November 1990). Interview with *The Times*: London
12. Clarke, K., Secretary of State for Education (January 1991). Speech to the Society of Education Officers: Harrogate
13. Pawsey, J. (January 1991). Question to the Secretary of State for Educatrion, House of Commons, London: *Hansard*
14. Clarke, K., Secretary of State for Education (January 1991). Reply to tabled question, House of Commons, London: *Hansard*
15. Department of Education and Science (January 1991). *Press release 22/91*, London: Department of Education and Science
16. Schools Examination and Assessment Council (1989). First meeting of the key stage 3 technology - subject steering committee
17. Schools Examination and Assessment Council (Feb. 1990). 'Minutes - second meeting of the key stage 3 technology - subject steering committee - point 5'
18. Schools Examination and Assessment Council (Feb. 1990). 'Minutes - third meeting of the key stage 3 technology - subject steering committee - point 3.4.1'
19. Schools Examination and Assessment Council (Oct. 1990). 'Minutes - third meeting of the key

stage 3 technology - subject steering committee - point 6.6'

20. Schools Examination and Assessment Council (January. 1991). 'Minutes - eight meeting of the key stage 3 technology - subject steering committee - point 5.6.1'
21. Eggleston, J (1992),. *Teaching Design and Technology*. Milton Keynes: The Open University, page 52
22. Clarke, K., Secretary of State for Education (November 1990). Interview with *The Times*: London
23. Graham, D. (1993). *A lesson for us all*. London: Routledge, page 109
24. Graham, D. (1993). *A lesson for us all*. London: Routledge, page 113
25. Schools Examination and Assessment Council (1991). 'PR 91/7 - *SEAC awards contracts for fourteen-year-olds tests*. London: Schools Examination and Assessment Council
26. Lord Griffiths of Forestfach (2nd December 1991). Letter to the Rt. Hon. Kenneth Clarke QC MP, Secretary of State for Education
27. Rt. Hon. Kenneth Clarke QC MP, Secretary of State for Education (undated, received at Schools Examination and Assessment Council 12th December 1991). Letter to Lord Griffiths of Forestfach, Chairman of SEAC
28. Blackburne, L. 'Technology test leak raises fears', London: *The Times Educational Supplement*. (20th March 1992)
29. The Guardian. 'Blue Peter lessons cost pupils dear'. London: *The Guardian* (21st October 1991)
30. Beadle, E., Prais, S.J. (1991). *Pre-Vocational Schooling in Europe Today*. London: National Institute of Economic and Social Research, Report Series, Number 1
31. Tate,N. (11th December 1991). Technology SAT review meeting, letter recording main outcomes, Schools Examination and Assessment Council
32. Tufnell, R. (8th January 1992). Response to letter from N. Tate
33. Eggleston, J. 'A fading flower', London: *The Times Educational Supplement*. (22nd November 1991)
34. Tate, N. (25th February 1992). Technology SAT review meeting, letter recording main outcomes, Schools Examination and Assessment Council
35. The Guardian. 'Tests for pupils of 14 to start in June'. London: *The Guardian*. (7th January 1992)
36. The Guardian. 'Tests for pupils of 14 to start in June'. London: *The Guardian*. (7th January 1992)
37. The Guardian. 'Tests for pupils of 14 to start in June'. London: *The Guardian*. (7th January 1992)
38. Robinson, P. and Smithers, A. (1992). *Technology in the National Curriculum - Getting it Right*. London: The Engineering Council, page 5
39. Education Correspondent . 'There's a spanner in the curriculum'. London: *The Daily Telegraph* (7th May 1992)
40. HMI Report (1992). *Technology - Key Stages 1, 2 and 3; a report by HMI on the First Year 1990-91*. London: Department for Education

41. National Curriculum Council (May 1992). *National Curriculum technology: The case for revising the Order.* York: National Curriculum Council
42. Patten, J. (2nd June 1992). 'Statement to The House of Commons', London: *Hansard*
43. Filer, D. (3rd June 1992). Press Release, London: The Engineering Council
44. Blackburne, L. 'Technology U turn starts'. London: *The Times Educational Supplement* . (26th June 1992)
45. Correspondence page. London: *The Times Educational Supplement.* (10th July 1992)
46. Minutes (27th. May 1992). Technology SAT review meeting, Schools Examination and Assessment Council
47. Department for Education/The Welsh Office. (December 1992). *Technology for ages 5 to 16.* London: HMSO
48. Shaw, B. Chief Adviser for Bolton. Reported in the *Times Educational Supplement.* (5th. February 1993)
49. Duffy, M. Headteacher, King Edward VI School, Morpeth . Reported in *The Times Educational Supplement.* (5th. February 1993)
50. Blackburne, L. 'Probe into pass kits for technology tests'. London: *The Times Educational Supplement.* (12th February 1993)
51. Blackburne, L. 'Probe into pass kits for technology tests', London: *The Times Educational Supplement.* (12th February 1993)
52. Hutchinson, P. Deputy Head, Moulton School, Northampton (26th February 1993). Correspondence Page, London: *The Times Educational Supplement*
53. Parish, T. Correspondence Page, London: *The Times Educational Supplement.* (26th February 1993).
54. Henessey, S., McCormick, R., Murphy, P. (1993). 'The truth about problem solving in design and technology'. *Design and Technology Times*, Autumn. Manchester: University of Salford
55. Hardcastle, J. (12th March 1993). Correspondence Page, London: *The Times Educational Supplement*
56. Melkie, J. 'Heads heap derision on Patten', London: *The Guardian.* (3rd June 1993)
57. Patten, J. Secretary of State for Education (2nd June 1993). Address to the Association of Head Teachers Annual Conference, Newcastle
58. Partington, P. President, Association of Head Teachers (2nd June 1993). Presidential address of the Association of Head Teachers Annual Conference, Newcastle
59. Melkie, J. '£35m cost of tests 'a waste of paper". London: *The Guardian.* (5th June 1993)
60. Boswell, T. (13th May 1993). Letter to the Head Teacher Daventry William Parker School, Northamptonshire

Chapter 4

Developing the Research Instruments

Synopsis

The development of the statutory assessment instruments occurred over a four year period. A significant focus of the research centred on the assessment devices which the teachers would use. If the assessments made were to be standard and reliable the criteria used by teachers would need be clear, unambiguous and capable of consistent use. The assessment tools had to empower teachers to act as examiners, a role with which only a few had experience. End of key stage (1,2 and 3) national assessment is a school based activity in which an individual's performance is not subject to external scaling as in the majority of award bearing examinations. External examinations, such as GCSE, do not require the final standards to be set prior to the test or task taking place. Post test or task calibration was never envisaged as part of the National Curriculum assessment process. The context in which these developments took place is reviewed followed by an analysis of the statutory Order and the difficulties which it posed in relation to assessment. The development of assessment devices took place in three phases and in relation to two different types of assessment activities. Due to the very different nature of the procedures, the research and development of assessment criteria for practical tasks is detailed separately from the research and development of assessment criteria for pencil and paper tests. This is followed by a review of the development of the pupil material. Research focused on establishing effective strategies which prompt pupils to provide appropriate evidence to meet the statements being assessed. This section is also divided between practical task material and test material.

Chapter sub-sections:

The development of Assessment Instruments for Statutory Assessment
The suitability of the statements of attainment for assessment purposes
Research and development - the assessment instruments for practical tasks
Research and development - the assessment instruments for tests
The development of the pupil material - practical tasks
The development of the pupil material - tests

The development of Assessment Instruments for Statutory Assessment

The context

Prior to any large scale trialling or piloting it was essential to investigate and understand the way in which the National Curriculum Order for technology evolved, its structure and content. The Order was developed by a group of eminent practitioners under the chairmanship of Lady Parkes. It was based largely on their experience and expectations. The recommendations of this group were then subject to consultation under the auspices of The National Curriculum Council as outlined in chapter 2. The outcome from the consultation process resulted in the NCC final proposals, however, responsibility then passed to the Department of Education and Science. Consequently the text of the final Order, laid before Parliament, was the result of senior civil servants working with the Department's legal branch. The curriculum document laid before Parliament on the 6 March 1990 under Section 4 of the Education Reform Act 1988, consisted of both statutory statements of attainment and programme of study. To this was added non-statutory guidance, developed by the NCC in conjunction with subject specialists, to assist schools in implementing the final statutory Order.

The design and technology profile component delineated a new subject, as Layton (1991) has observed,

'....there is no obvious general version (of design and technology) which can serve as a model for what goes on in schools. This is in contrast to subjects such as mathematics, science and English, each of which has a long history in the curriculum...'[1]

Yet despite its novelty, the prescribed characteristics of the subject were not based on the result of substantial research; rather, it represented the culmination of a phase of continuous evolution, reaching back thirty years. It was, therefore, an extremely bold step to define levels of achievement and stages of progression with little substantive evidence to support the conclusions The working group had been well informed of the work of the Assessment and Performance Unit for design and technology based at Goldsmith's College. Although the APUs findings did influence the group's thinking in many ways, it had not been the APUs role to define levels of achievement or progression in the manner required by the National Curriculum. In retrospect, in taking this imaginative leap it is apparent that the Order's architects failed to appreciate the fragile nature of the foundations. It was apparent from the outset, that although the Order was philosophically coherent and all embracing in its coverage, its interpretation as an assessment device, capable of being used both for day to day and end of key stage assessment would present significant problems.

The Order represented a substantial re-orientation from product to process. A product-oriented approach, focusing on skills and knowledge, is more easily assessed by traditional procedures. TGAT (Task Group on Assessment and Testing) had given priority to assessments which gave direct information about pupils' achievements in relation to objectives. This requirement to assess via criterion referencing complemented the thinking of the design and technology working group. Their most radical step was to describe achievement in the subject solely in levels of operational capability. The emphasis was placed on how things were done rather than the final outcome. The process was characterised by terms such as identify, investigate, communicate, plan, select, make and evaluate - the ability to take decisions and follow them through was at the heart of the subject. This was an aspect of the group's work which concorded with the conclusions of the Assessment of Performance Unit final report.
'It is what pupils do with their understanding that counts; and the attainment targets have to enshrine this activity - capability. We must learn to live with the fact that content is not the master of technology; it is the servant.'[2]

The result, was a statutory Order which presented teachers with a complex assessment model. As with all National Curriculum subjects it was based on attainment targets which consisted of statements of attainment, but many of these statements were capable of being interpreted in a wide variety of ways. To many it seemed the attainment targets and the statements of attainment described curriculum objectives and ideals rather than key performance indicators which children would exhibit and teachers could observe in a clear and unambiguous fashion. The legalistic nature of the Order prevented many from gaining a clear vision of what would be the pupils' experiences of the subject. The Order also failed to establish ground rules for determining achievement. For example, no clear rules were provided on how a level was achieved - how many statements needed to be satisfied. This lack of clarity, within the context of a statutory document, created widespread confusion which had not been resolved by the time of the first statutory assessments. These fundamental shortcomings were only apparent to those working with the Order on a day to day basis. In contrast the optimism which surrounded the launch of the Order was fulsome. For example, the then Director of the National Curriculum Council, Dunchan Graham, commented,
' technology is the jewel in the crown of the National Curriculum.'[3]
In this climate critics were perceived as Luddites, lacking in vision and labelled, for example, as traditionalists. The majority of teachers attempted to respond to the process driven nature, developed via subject integration which the Order promoted. Many however, lacked a genuine comprehension of the integrative nature of the attainment targets. Projects and schemes of work were instigated in

which teachers, misguidedly, interpreted the Order in a way which allowed pupils to determine the nature and extent of the learning outcomes. These were frequently based on youth culture as a means of providing pupils with a relevant context. This approach failed to tackle the key issue of assessment as assessment goals were not established at the outset. This resulted in an unstructured approach to learning outcomes which was all to often deficient in rigour. Teachers were not wholly to blame. There was a lack of national guidance and consequently a wide variety of approaches to assessment was developed. This existed at regional, LEA, school and teacher level. Complex schemes were devised, a few for financial gain, for example NDTEF (National Design and Technology Education Foundation), but the majority resulting from the commitment and enthusiasm of teachers attempting to make practical sense of the Order.

Despite the conscientious and committed way in which many embraced the Order, not all were convinced of its message. Some practitioners were genuinely concerned about the direction and motives of the Order. Others focused on the practical problems of realising the Order's objectives and many were just sceptical of change. Within this context, the task of determining if the Order could be assessed in a valid and reliable way was one of the most challenging aspects of implementing NC design and technology. Research and development had to take place with pupils who had not yet followed the programme of study. Before assessment could be undertaken the trials and pilots had to instigate design and technology activity. This was a lengthy process and by comparison the development of the statutory Order appeared in relative terms straightforward, if judged by the speed with which the working groups produced the statements of attainment. The key issue, which could only be established by operationalising the Order, was had they got it right?
' We do not yet know whether they are in the right order, and some will have to be described in more detail to allow teacher assessment to take place.'[4]

In design and technology the issue of precision versus breadth, in relation to assessment, also raises a fundamental issue. Assessment is more reliable if it focuses on the specific but the danger of this approach is that knowledge and understanding will be divorced from a genuine context. Many, such as Nuttall (1989), warned of this,
'school learning is disconnected enough from real life as it is, without fragmenting the decontextualised skills as well.'[5]
In this context valid and reliable end of key stage assessments would need to:

- be based on a practical, pragmatic approach, so that those carrying out the assessment process could do so in a standard and consistent fashion;
- demonstrate unequivocally the relationship of the assessment to the Order to ensure that the assessment procedure remained within the legal and

constitutional framework;

- develop a system of assessing design and technology which reflected the nature of the National Curriculum subject as interpreted by the majority of the design and technology community.

As described in chapter 3, the requirements of statutory assessment became increasingly driven by political dictate during the development period. However, if the statutory assessment procedure was to maintain the support of teachers and maintain educational credibility it was essential that the system:

- reflected the holistic and process-driven nature of the subject;
- produced valid and reliable attainment target levels which supported and informed teacher assessment;
- provided teachers with diagnostic information about each pupil;
- was realistic given the resources available - especially teacher time;
- reinforced good practice in relation to teaching and learning.

The suitability of the statements of attainment for assessment purposes

The Statements of Attainment as described in the report of the Working Party, the NCC consultative document and the final Statutory Order are difficult to interpret as an assessment tool. It was essential, via research and development, to devise an assessment instrument which interpreted the statements accurately and comprehensively yet produced a system which met the above objectives. An analysis of the statements in the 1990 Order revealed structural difficulties:

1. The different number of SoAs within ATs created inconsistencies in terms of the number of statements at different levels. The demand at each level is not directly related to the number of statements, and it is perfectly acceptable that there may be fewer statements at higher and lower levels than in the middle. However, it is recognised that this creates problems in promoting consistent assessment and even progression. Moreover some of the differences in the numbers of statements, in the ten levels, seem extreme (see table below). The number of statements, by AT in the 1990 Statutory Order is shown below:

Level	Te1	Te2	Te3	Te4		Total
1	2	1	1	2	=	6
2	3	1	3	2	=	9
3	2	5	4	2	=	13
4	6	4	5	4	=	19
5	2	5	4	3	=	14
6	3	4	6	5	=	18
7	4	3	3	1	=	11
8	3	3	3	2	=	11
9	2	2	3	1	=	8
10	3	1	2	2	=	8
Total	30	29	34	24	=	117

average number of statements per AT = approx. 29

2. In relation to elements of each attainment target the statements and levels do not necessarily provide a complete hierarchy (given the nature of the subject this would be difficult). It is essential that where progression exists it can be clearly demonstrated through the assessment process. In practice pupils who might not be capable of completing all the statements within a certain level, should still have the opportunity to move forward in aspects which they have mastered. The opportunity to progress must be provided in the interests of good teaching, learning and motivation. An assessment instrument should take into account a path of progression. How the level of capability achieved is arrived at is a separate decision. Each level must be clearly differentiated from the others by distinguishable activities and the levels of achievement required. The statements, although suitable in some instances, create problems where progression, on a level by level basis, is not possible. For example, in Te2 the ability to communicate ideas is fundamental, yet at several levels there is no statement which characterises this aspect of capability.

3. Many statements are multi-faceted creating problems in relation to assessment, for example Te2 level 4b, 8a and 10a. The number of assessment items by level and attainment target is shown in the table below. This results in a 174 assessment items a 50% increase on the number of statements.

Level	**Te1**	**Te2**	**Te3**	**Te4**		**Total**
1	3	1	1	4	=	9
2	4	2	4	4	=	14
3	2	5	4	4	=	15
4	7	5	8	6	=	26
5	4	7	6	6	=	23
6	6	5	8	8	=	27
7	5	6	7	3	=	21
8	3	5	5	8	=	21
9	2	3	5	2	=	12
10	3	8	5	3	=	19
Total	39	47	53	48	=	187

average number of items per attainment target = approx. 47

4. The statements are difficult to apply as criteria for assessment as they include value judgements, for example Te3 level 7a..... "to achieve the desired quality". Desired quality is clearly open to interpretation. Such statements needed to be exemplified across a range of media to resolve the issues they create. Only if terms such as "desired quality" are represented by practical examples of pupils' work which meets this standard, will teachers be able, by comparison, to apply the criteria consistently.

5. Certain statements, given the time constraints, are difficult to assess in a single design and technology activity. In the context of statutory end of key stage assessment these statements need to be identified and placed in the province of teacher assessment.

Research and development
- the assessment instruments for practical tasks

The tradition in design and technology is to assess practical tasks by outcome - to present the same task to all pupils and then award credit on the basis of their outcome. This approach was initially endorsed by all parties as the only sensible way of proceeding. As all tasks were to include work on and assess all four attainment targets, pupils were responsible for identifying needs and opportunities (Te1) within a prescribed context. This in practice resulted in a wide variety of tasks. Consequently assessment criteria could not be task specific and had to be general in character. Following a range of small scale investigations two possible routes appeared worthy of further investigation, namely:

- defining strands within each attainment target;
- compiling level statements which consolidated SoAs into a single coherent descriptor.

The concept of strand assessment

Careful analysis of the statements revealed internal strands of competency/ capability in all of the Attainment Targets. These strands were used as the basis of an assessment matrix with the strands arranged as vertical columns and the levels of attainment as the horizontal bands. The strands of competency identified were:

Strands	
Identifying Possibilities Carrying out investigation Recognising implications	attainment target 1
Developing designs Decision making Communicating and modelling ideas	attainment target 2
Competency with materials Manufacturing capability Organisation and planning	attainment target 3
Evaluating own technological activity Appreciating and appraising technology	attainment target 4

If the SoAs are sorted into these strands the Attainment Targets appear as shown in the stranded matrix, shown in appendix 4.1, p. 273. When statements appear to include elements common to two or all of the strands in the Attainment Target, they have been shown as bridging across the strands.

The strands are generic components of design & technology capability whilst SoAs, in isolation, are not. The eleven strands encompass and structure the key capabilities of the holistic activity which the attainment targets seek to assess. The aim of the matrix was to encourage teachers to view achievement within the ATs across the whole activity, which might not be the case if the statements were seen as independent assessment goals. The matrix provides a commonsense framework and a realistic basis for assessment.

There was sound evidence that (if a task was structured correctly) most pupils produce some examples of performance in each of the eleven strands whilst engaged in a typical design and technology activity. Some pupils may not produce their best performance level in every strand in a single activity but this can be accommodated by the design of aggregation rules. This matrix, by design, sampled in a fair way without attempting an exhaustive coverage of the whole repertoire of SoAs.

In terms of descriptive validity, evidence from trialling strongly suggested that teachers and inspectors who examined the stranded matrix and used it for assessment purposes found it a more direct and manageable instrument than the statements which it subsumed.

The development of a coherent stranded assessment matrix

The Statutory Order and experience gained during trialling in the Spring of 1990 were used to improve the original matrix. These improvements were incorporated into the matrix used in the 1990 Summer trials. This model incorporated the following principles:

- The attainment targets described the attributes of a level of performance whilst the matrix established a pupil's operational level by stranding and subsuming statements. The matrix developed the attainment targets into operational criteria.
- The matrix was designed to assess the level at which a pupil could operate rather than the specific SoAs which the pupil had achieved.
- The matrix had to produce operational levels equivalent to the levels of attainment as characterised by statements in the Standing Orders.

Two key criteria were observed in its design:

- the statement should keep as closely to the Statutory Order as possible;
- a continuous series of statements was essential in each strand.

To achieve this the following rules were applied.

1. Statements or parts of statements were moved up or down one level when there appeared to be an excessive demand in one strand at a certain level. This does not affect levels achieved because of the rules of aggregation; it simply established a more realistic learning curve;
2. If there was a gap in a strand the first procedure was to split the statement at a higher level in a hierarchical way . If this was not possible a reasonable statement was produced to describe appropriate performance at that level;
3. If there were more than three statements, appropriate ones were amalgamated. If this presented problems, the least appropriate in terms of establishing operational 'levelness' was dropped;

The matrix used in the Summer 1990 trial was produced by following these guidelines. In practice, there are few instances where changes have been made and these did not alter the intention or meaning of the Statutory Order. The matrix used in the 1990 trial is shown in appendix 4.2, p. 227.

Developing an attainment target 'levelling' assessment instrument

In exploring the concept of AT 'levelness', an instrument was developed that characterised each AT level by means of a single definition. The main reason for doing this was to make assessment practicable for teachers. This development was a response to issues observed during the final assessment phase of the 1990 Spring trial. An assessment instrument of this nature might also result in reliable assessments being made more efficiently. For the four ATs the SoAs at each level were deconstructed and synthesised into a single definition encompassing the meanings of the original SoAs and their attributes. In the course of this development the guiding principle was to make new single definitions functionally equivalent to the subsumed originals as criterion reference points.

In the operational form of an assessment instrument, part of each single definition was highlighted to show the essential difference in levelness between it and those definitions immediately above and below. This assists in making a holistic judgement of levelness within an AT, consequently there is no requirement for a mastery rule to be used to determine level of achievement. Each decision therefore called for a broader qualitative judgement based on the concept of a single definition best describing performance. An AT level could be awarded when there

was evidence in a pupil's folder and the teacher records that the majority of aspects within the sentence defining the level had been achieved, This matrix is included in this thesis for information, see appendix 4.3, p. 283.

Level Indicator/Progression Guide

Another aspect of the research focused on the notion of identifying the key progression which took place from level to level in each attainment target. In the 1990 Spring trials considerable difficulty was reported in differentiating between levels when judging pupils' work. This problem appeared to stem from the SoAs not providing any obvious markers pointing to differences. A 'level guide' was developed that detailed, level by level, the essential differences in pupil expectation between them.

This guide emphasised the developmental progression and gradual sophistication that can be traced logically through each attainment target. Teachers used this guide to steer pupils in appropriate directions and encourage them to address those aspects which would enable them to demonstrate their highest level of capability. During assessment the guide helped all teachers to clarify what was required for levels to be achieved. It did not contain sufficient detail from the statements of attainment to be an accurate assessment instrument in itself. Its purpose was simply to offer a clear overall grasp of what was required in each attainment target at every level. This guide received a highly favourable response from teachers and is detailed in appendix 4.4, p. 289.

Operational trials of these devices

These three instruments were used as the basis of all trials between 1989 and the Summer of 1992.

1990 Trial

These approaches were the focus of a detailed evaluation at the conclusion of the 1990 Summer Trial. This revealed that both approaches were more acceptable than the statements of attainment to teachers and produced more reliable results than using the statements, see appendix 4.5, p. 291.

The 1991 Pilot

The developments undertaken during 1990 demonstrated that, given a suitable structure, assessment procedures could be developed which would help teachers to carry them out in a standard and reliable fashion. This pilot was designed to investigate key issues in relation to two assessment instruments. The two selected

were the AT Instrument and the original stranded matrix, which used the statutory statements of attainment. All teachers were also provided with the level guide to assist teacher assessment and initial assessment of the end of key stage task. SEAC had instructed development teams that a full national statutory assessment would need to identify clearly the statements being assessed. The matrix used in 1990, which attempted to overcome the deficiencies of the statements, would not meet the legal requirements as it had modified the language of some of the statements. (This matrix was however valuable in establishing principles which were adopted by the HMI task group charged with producing proposals for a revision of the Order in 1992. The proposal included stranded attainment targets which, with the exception of levels 1 and 2 provided a coherent, progressive and hierarchical structure.)

Coverage of ATs, SoAs and items within ATs

In any statutory assessment process it is important that legal requirements as established in the specification are fully met. The holistic nature of the assessment tasks dictated that any genuine design and technological activity required pupils to demonstrate their capability in each of the attainment targets. This, at the time, was accepted as the only valid context for the assessment of design and technology capability. In design and technology it is the application and use of skills, knowledge and values in relation to a particular task which constitutes capability and this is what any end of key stage task should seek to assess. Assessment should be in the context of the whole activity, any attempt to focus on statements or items would invalidate the notion of capability. The 1991 assessments were designed to prompt pupils to address statements as their response to a situation developed. Any teacher intervention should be appropriate to the task, issues and the developing situation. Compulsion to address all statements would have resulted in tokenism. Each pupil consequently compiled his or her own agenda of evidence whilst undertaking a practical task. No one statement might be fully met but items from a wide range might be adequately evidenced. From the evidence, the assessment focus was to determine the level at which the pupil was operating.

The assessment process must therefore reflect the sampling procedure which each pupil's approach to the project determined, fairly and reliably. The two approaches used for the pilot did this in different ways. The stranded matrix is essentially an analytical/systematic approach based on rules whilst the AT Instrument allowed an holistic/lateral approach which placed more emphasis on a judgement of levelness.

Reporting in this trial was at attainment target level. Aggregation rules were therefore required for the stranded matrix, as teachers would have to determine a

level for each of the two or three strands within an attainment target. The AT Instrument, however, provided data at this level. The stranded matrix required achievement in each of the strands in an AT for that level to be awarded. Typically this meant that if there are three or less statements, all needed to be satisfied. If a level is characterised by 4, 5 or 6 statements, a minimum of three must be evidenced - 75%, 60% and 50% respectively. In practice this meant that in levels 1 and 2, and 8, 9 and 10 (on average 2 statements per level per AT) all statements must be met to achieve a level. In levels 3 to 7 - the focus of key stage 3 (on average 3.5 statements per level per AT) - sampling ensured that the task asked of each pupil was sensible and realistic.

The AT Instrument asked teachers to make judgements about the levelness of a pupil's achievement. Teachers were asked to judge which of the level descriptors in each AT most typically described the pupil's work, hence the teacher had some discretion to decide what could be expected as sufficient evidence, given the nature of the task.

Evaluation of the assessment procedure

Teacher responses to the assessment procedures used in 1991 were mixed. Irrespective of the assessment instrument used, they revealed three underlying problems:

Language of the Order - Very few teachers appeared to be familiar with the Order or to appreciate its statutory force. In commenting on the assessment procedure, several teachers even went so far as to question the difficulty of the language used (for them) and the appropriateness of the statements themselves, apparently believing that these had been devised for the pilot.

Criterion referencing - The difficulty teachers found in coming to terms with a new subject was for many compounded by the equally new concept of criterion referencing. Considerable debate took place before any kind of consensus view emerged on levels achieved across a wide range of work.

Communication - The benefits of the INSET were lost in the transmission from the teachers in each school receiving training to the other teachers not receiving training. Important messages became confused or were simply not passed on.

The schools which successfully cascaded INSET adopted thorough and systematic procedures for end of key stage assessment. In some schools, for example, specialists from different disciplines were paired off to assess work; in others, initial consensus was arrived at by group discussion. In these schools all specialists came together at least once to discuss common issues and internally moderate the work of the cohort. A range of positive outcomes was reported including:

- improved communication and understandings between specialists;
- consistency of and confidence in making assessments;
- advancement of curriculum planning for key stage 3.

The stranded matrix and assessment instrument were compared in detail by means of a re-mark exercise. Teachers' preferences and comments were evenly mixed and it was impossible to draw more than general conclusions. Remembering and consistently interpreting the statements proved to be a problem whatever assessment procedure was used but some were clearly in favour of making atomistic judgements while an equal number were happier to take a more holistic view on the basis of lengthy compound statements.

The stranded matrix, more so than the assessment instrument, produced 'quantum leap' effects whereby teachers were tempted to award a higher level in an AT but could not justify the level immediately below it - and so settled for the one below that. It was generally agreed, however, that this problem stemmed from the original Order and not the assessment devices. A full evaluation of the effectiveness of the assessment devices appears in appendix 4.6, p. 297.

The 1992 Pilot

As detailed in chapter 3 the specification in 1992 changed on several occasions as the focus of statutory assessment was reappraised. Initially, it was indicated that practical tasks would only provide a context on which the formal statutory, end of key stage tests would be based. Participating schools, therefore, undertook a practical task but did not receive any formal assessment advice on how to judge the outcomes. Many schools adopted their internal school procedures, others devised systems for the first time, whilst many obtained copies of either the stranded matrix or the AT Instrument from a variety of sources. As a result this pilot produced no substantive evidence which could be used to inform the development of assessment procedures relating to practical tasks. Further political decisions recognised the role of practical tasks in design and technology and led to a decision being taken to assess the attainment targets by different methods, Te2 and Te3 being assessed by a practical task.

The 1992 Pre-Statutory assessment trial - practical task

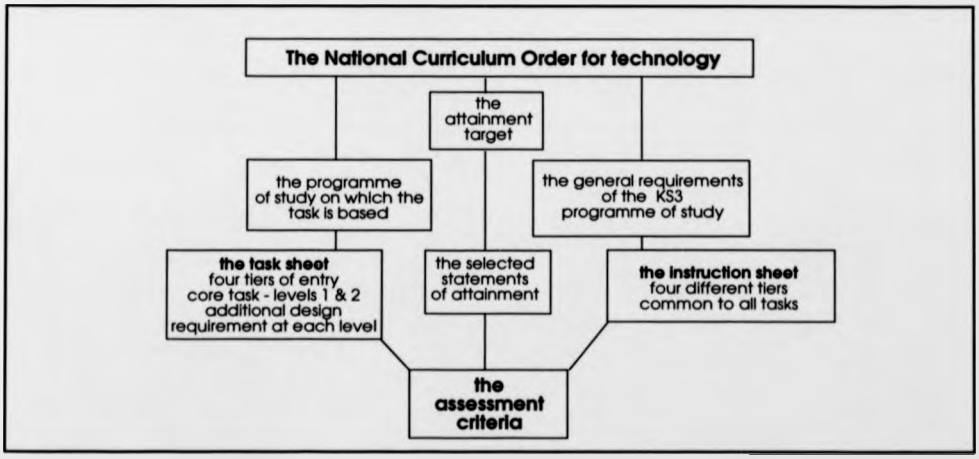
The assessment procedure used in this trial was, with only minor modifications, the same as the one adopted for the first full statutory assessment. The research undertaken in previous pilots and trials was pivotal in the development of the material and in convincing Schools Examination and Assessment Council that this approach was the most likely to produce valid and reliable assessments. The key principle adopted was that the assessment should focus on the level of a pupil's

performance rather than the statements of attainment which were evidenced. Previous trials had shown that this approach resulted in assessments which were as valid and reliable as any other method and removed the requirement for a complex aggregation rule to arrive at an attainment target level. This was essential as the assessment order prescribed a complex weighting of attainment targets to reach a subject score - two complex procedures (one to reach an attainment target level and a second to determine a subject score) would have proved indefensible in the prevailing climate, which was becoming increasingly hostile.

The 'levelness' approach in the practical task was the first example in statutory assessment where the statements of attainment were subsumed and interpreted into task-specific criteria. The argument was won on the basis of manageability, however the approach established a model example of how to assess design and technology activities and one which has been adopted by others including GCSE groups, (see Welsh Joint Education Committee GCSE syllabus 1993 - Design & Technology). In addition the approach of the level guide, which was so successful with teachers, was also adopted. The requirements of the progression from one level to another were clearly highlighted in the assessment criteria at each level for each attainment target. The first proposal is shown in appendix 4.7, p. 306.

Developing assessment criteria

The assessment criteria, for the first time, were highly specific as they interpreted statements of attainment and elements of the programme of study through the prescribed tasks which pupils were undertaking. The assessment criteria did in fact serve a dual purpose as they established clear unambiguous goals, detailing the requirements of each level for both the pupil being assessed and the teacher carrying out the assessment. The process outlined in the diagram below, is then detailed in a practical context.



Development of assessment criteria - a practical example

Te 2 - level 5 - Construction material

Te 2 - Pupils should be able to generate a design specification, explore ideas to produce a design proposal and develop it into a realistic, appropriate and achievable design.

Statements of attainment

5a. record the progress of their ideas, showing how they have clarified and developed them.

5b. extend their first ideas by combining various aspects of them to be formulate a design proposal and explain why some ideas were not used.

5e. specify what they intend to do and what they will need by using simple plans and flow diagrams.

The programme of study

- select and use simple mechanisms, including linkages and gearing making prototypes;
- present their design and technological ideas and proposals using modelling techniques and specialist vocabulary.

The task

Design and make a clamp to hold thin objects. The clamp should not be longer than 180 mm. The clamp must be easily moved around. If you push down on the clamp with your hand it should lock and stay locked when you take your hand away. It should be possible to unlock the clamp using just one hand.

The assessment criteria - Te 2 level 5 - construction materials

A record of ideas shows how the design has been **developed and refined.** Possible solutions to the various aspects of the design have been **integrated to produce a sound proposal** which is clearly presented using either detailed annotated drawings or a working model. The design proposed will operate successfully as a clamp. **Unlocking of the clamp has been tackled, though the solution does not need to work well. A plan is provided which describes how the clamp will be made.**

This process was undertaken for the three different tasks for the two attainment targets - this resulted in 60 assessment criteria - 20 per task. It was then important to ensure that the criteria clearly described progression so that each statement was examined in relation to both the preceding and proceeding statements. As a result of this process some modifications were made. Progression was indicated by highlighting additional requirements at each level in bold text. An example of three statements which illustrate this are shown over the page.

Progression in the assessment criteria

Te 3 - level 3 - Food - A compact bar has been made which shows **a basic understanding of ingredients and some attempt to achieve a quality finish,** though the content may be high in sugar and fat.

Te 3 - level 4 - Food - A compact bar has been made which **achieves a quality finish** and shows that **thought** has gone into **providing energy in a form that would be palatable and portable**, though not necessarily healthy, for the user.

Te 3 - level 5 - Food - **A plan for making** has been provided and the finished bar, **containing some protein, fat, vitamins and minerals, as well as carbohydrates**, provides a compact solution. There is an attempt to enable the user to **divide the bar into smaller pieces, using divisions which remain after the bar has cooled or set.**

Finally the assessment criteria at each level for each attainment target were reviewed across the tasks to ensure comparability of demand. If modifications were made the new criteria were re-examined to ensure that progress, within the task, was not affected. An example of a cross-task criteria is provided below.

Comparability of the assessment criteria

Te 3 - level 3 - Construction materials
A clamp has been made which **works satisfactorily. Materials and equipment have been chosen** such that the clamp has been made **to a satisfactory level of accuracy and quality which enables it to be used.**

Te 3 - level 3 - Control
A beacon board has been made which **works reliably. Materials and equipment have been chosen** and used, such that the board has been made **to a satisfactory level of accuracy and quality.**

Te 3 - level 3 - Food
A compact bar has been made which shows a **basic understanding of ingredients and some attempt to achieve a quality finish**, though the content may be high in sugar and fat.
(From level 3 onwards, a bar should be defined as something which maintains its shape and in which all the contents adhere together.)

In the trial teachers commented favourably on the assessment procedure. Levelling was commonly felt to provide a fairer assessment of each pupil's capability. It was also far more manageable than identifying the statements of attainment which should have been evidenced.

Research and Development of the assessment instruments for tests

Issues relating to the statements

The decision to introduce pencil and paper tests was a political one. Politically it was justified by the teacher evaluation of the 1991 pilot (KS1) which stated that the assessment procedures were unmanageable. This decision was not based on the suitability of this subject for this style of assessment. In design and technology the statements of attainment which these tests sought to assess were not designed as assessment criteria for pencil and paper testing. The ambiguity, lack of clarity and clear progression make the statements difficult to use as assessment tools in any assessment context. These issues are even more apparent in the context of pencil and paper tests. The statements describe attributes of what is essentially a practical activity. The ability of fourteen year olds to address such statements in a theoretical context, drawing on their experience of design and technology was from the outset open to question. With experience, guidance and planned teaching, pupils' capacity to answer tests of this type will improve. However, this will only be achieved at the expense (in terms of time) of the unique practical experience which this subject offers.

August 1991 - August 1992

The initial research was based on a model which sought to assess all four attainment targets. But following a decision of the Secretary of State both trials - held in February and March 1992 - and the Summer Pilot assessed only attainment targets 1 and 4 via the test. Although the Autumn trialling was therefore superseded by a different model, some important conclusions were reached on the basis of its outcome. The first test papers assessed performance by outcome, questions were general and teachers "levelled" a pupil's response by applying, "what to look for" exemplification based on a band of specified statements. It was decided not to proceed with this approach for the following reasons:

- the questions developed for this approach needed to be broad and general in nature to accommodate the wide range of statements and levels which each test covered;
- pupil responses were bland and uniform, which resulted in a lack of clear differentiation - teachers found it difficult to discriminate confidently between answers;
- assessments made by teachers were extremely varied and inconsistent. The statements of attainment, because of their lack of clarity and clear progression, are open to wide interpretation. These characteristics are particularly extreme when teachers are attempting to determine the worth of a response in relation to four or five levels;

- as only a single statement could be assessed at each level the evidence on which levels were being awarded was superficial (this is dependent on the number of statements at a level in an attainment target - variation from 1 to 6).

It was concluded that questions needed to be far more focused and would, therefore, have to assess specific statements. This would allow assessment criteria to be more focused. Consequently all subsequent trials of pencil and paper tests were based on the principle of differentiation by task. A rationale for an assessment procedure dealing with only these two attainment targets was developed. This was based on the following criteria:

- the test should be common to all pupils regardless of the practical task undertaken;
- the test should draw on the experience of the practical task;
- the test should also assess the programme of study;
- the two attainment targets should be given equal weighting in relation to aspects such as, number of statements assessed and time;
- the test should be as coherent and logical as the specification allowed.

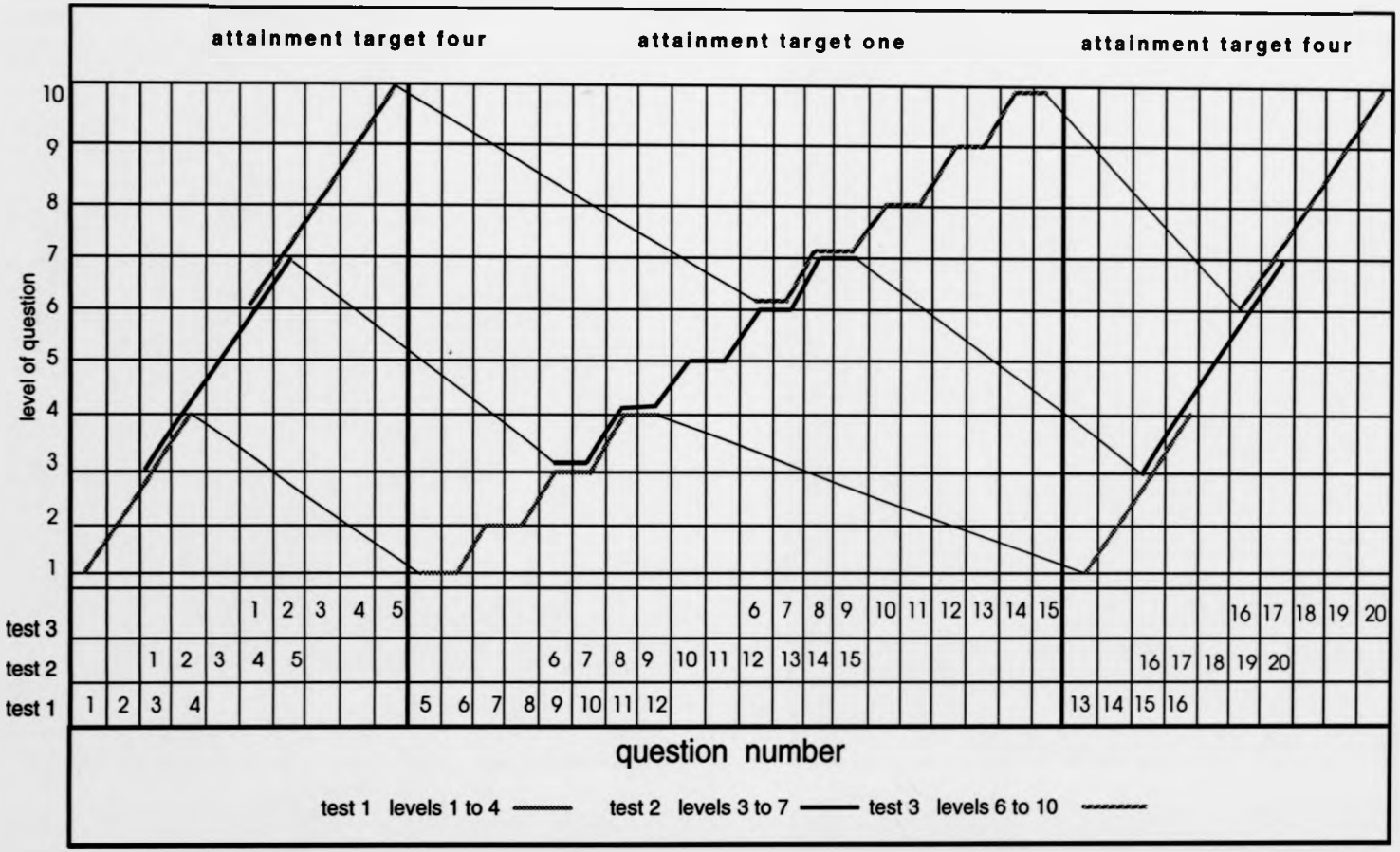
The following model, for the 1992 Summer Pilot, was developed to meet these criteria.

section	**pupil context**	**assessment**
Section A	review the long task	Te 4
Section B	identifying needs and opportunities in an extension of the long task	Te 1
Section C	evaluate the work of others in response to the context extension in section B	Te 4

As only two attainment targets were being assessed by this model, the number of statements addressed at each level was increased by 100%. This provided far more evidence of achievement and offered greater confidence in the resulting assessments. Section A assesses one statement at each level in relation to Te 4, whilst Section C assesses another statement (at some levels the same statement is assessed, e.g. Te 4 level 7). Section B assesses two statements at each level. The pattern of the three tests is shown on the diagram overpage. Appendix 4.8, p. 312, shows the test instructions and an example in relation to one AT.

Selection of statements of attainment

The selection of statements to be assessed via the test had to ensure that coverage was sufficient to be legally valid. These were selected on the basis that,



The pattern of the three differentiated tests in design and technology

at each level one of the following two cases applied:

- If two or fewer statements occur at a level both were selected;
  (11 out of 20 possible cases - 4 levels in Te 1 and 7 levels in Te 4.)
- If more than two statements occur at a level the following criteria were applied:
  (9 out of 20 possible cases - 6 levels in Te 1 and 3 levels in Te 4.)
  - can the statement be tested by a pencil and paper test?
  - does the statement allow a logical question to be asked within the sequence?
  - does the statement allow a different issue to be addressed within the test band(s) in which it falls?

In reality this resulted in a situation which gave no scope for choice in selecting statements. Indeed in case 1, statements might have been rejected if the case 2 criteria had been applied. It was impossible, in a test of this nature, to assess fully all aspects of each statement. The questions sought to probe aspects of a statement so that a final decision/legal point of arbitration in each case should be the statement.

In addition to the statement the mark scheme provided, for teachers, "what to look for" in marking a question. This information was based firmly on the statement being addressed. In reality the "what to look for" became the assessment criteria which were used by teachers. Consequently the validity of the question's ability to probe the statement became central to the research. Examples of possible responses were also provided which were generally seen as useful, but did create issues of interpretation. As is often the case with examples, some teachers, quite wrongly, believed they were the only correct answers. The nature of the questions, mainly discursive, meant that professional judgement had to be exercised in order to reach an assessment decision. An example of a, "what to look for" is given below.

Two valid reasons are given for the choice of materials. Answers may refer to visual properties such as colour or pattern, or the effect of combining materials. A personal statement such as, "because I like them" is not adequate at this level.

The use of the word "valid" created assessment problems as, like the SoAs , it relied on teachers interpreting "valid" in a consistent fashion. What to look for statements were more general in Section A because of the linkage to the practical task. In the other two sections - B and C, "what to look for" criteria were more specific, but as none of the questions had a straightforward, unambiguous, right or wrong answer assessment was always more than just a matching process. This lack of precision is a consequence of the style of statements of attainment being assessed and is directly related to their process characteristics.

A key issue in this trial was the application of the assessment criteria. The question posed in assessing a single statement at a sufficient level of both depth and coverage could be quite complex. It might consist of several parts and consequently, several different assessment decisions had to made by the assessor. How many of these needed to be answered satisfactorily to accept that the pupil had successfully satisfied that statement? As there was a requirement to operate the principle of criterion referencing it was decided that a pupil must satisfy all aspects of the test in relation to a statement, to be assessed as having satisfied the statement. Asking pupils to get every aspect of a question correct was undoubtedly a severe rule. The principal outcome from this pilot was a conclusion that:

> mark schemes should be constructed to allow for a margin of failure within a question or level - compensation should be available when the majority of an answer is correct.

Another aspect of concern was the way in which the test had assessed Te 4 - Evaluation. As detailed in chapter 3 a political decision had created the split between the practical task and its evaluation. The lack of focus of the practical task in this pilot meant that the question in the test had to be of a general nature, to cover all possible activities and outcomes. If the test had to be more specific so that the marking criteria could be more focused then the practical task would have to be prescribed. In a statutory context there were sound arguments for this decision being taken.

The 1992 Pre-Statutory assessment trial - test

The two principal findings from the 1992 pilot required that a different structure be established for the written test. Te 4 has two generic strands - evaluating one's own work and evaluating the work of others. The test of Te 4 was now linked exclusively to a focused task undertaken in a prescribed material context. To assess this aspect securely the decision had to be taken not to assess the other strand. The test therefore had two sections.

section	**pupil context**	**assessment**
Section A	evaluate the practical task	Te 4
Section B	identifying needs and opportunities in a new context	Te 1

Section A would need to be developed in three different contexts so that pupils could sit a test which related to the practical task which they all took. Section B would be common to all.

The focus of assessment

As with the practical task there were persuasive reasons why the assessment should focus on the level at which a pupil was operating rather than a mastery of specific statements. Teachers were undoubtedly more comfortable with a process which involved: reading an answer; deciding if it met the assessment criteria; awarding a mark if yes, not awarding a mark if no. Research focused on designing a test which accommodated this principle within a criterion referenced framework. Based on pre-test trials undertaken the following model was developed.

A decision is taken about the number of mark points available at each level.

At each level
- the statements to be assessed are identified;
- questions are devised which probe these statements;
- on the basis of trialling a mastery level is set;
- questions are modified to ensure progression and hierarchy between levels;
- trialling confirms reliability of the model.

As manageability was perceived to be a major factor two other strategic decisions were taken.

1. The questions would be designed to create the same number of mark points at each level.
2. The questions would be calibrated to achieve the same mastery requirement at each level.

In addition the statement being assessed by each question was identified so that it would also be possible to identify, for formative purposes in relation to teacher assessment and reporting to parents, the statements which a pupil had satisfied. An additional construct required that at levels where more than one statement was being assessed (this was at all levels except where there was only one statement in the Order) the mastery level could not be achieved on assessment evidence relating to only one of the statements.

The Assessment Instrument

Following extensive trials and modifications an assessment instrument or mark scheme was developed for the first full statutory assessment. The previous research outcomes resulted in an assessment instrument which was a compromise between a true criterion referenced approach and a traditional norm-referencing mark scheme. This was necessary for three reasons.

1. To promote higher levels of reliability by providing teachers with a procedure with which they would feel confident and familiar.

2. To accommodate political dictate in relation to the method of assessment.
3. To devise a system which would be within the legal requirements of coverage and would enable both summative and formative information, at a variety of levels, to be readily gathered.

The mark scheme provided assessment criteria at three levels of specificity:

1. legal - the statement being assessed;
2. professional - general characteristics of a satisfactory answer
3. examples - a range of pupil responses which had been categorised into acceptable and unacceptable; these ranged from answers on the cusp to those which were clearly in one category or the other.

Selection of statements of attainment

The test was required to assess half or more of the SoAs and consequently the following pattern emerged.

Level	Te 1	Te 4
1	1	1
2	2	1
3	1	1
4	3	2
5	1	2
6	2	3
7	2	1
8	2	1
9	1	1
10	2	1

Statements per Test

Test 1 - 12
Test 2 - 15
Test 3 - 14
Test 4 - 11

The following statements were selected to be assessed via the test:

Te 1 1a & 1b; 2b & 2c; 3a & 3b; 4d & 4e; 5a & 5b; 6a & 6c; 7a & 7c; 8a & 8b; 9a & 9b; 10b & 10c

Te 4 1a & 1b; 2a & 2b; 3a & 3b; 4a & 4b; 5a & 5b; 6a & 6b; 7a; 8a; 9a; 10a

The questions addressed either a whole SoA or part of an SoA . This approach was essential because the design and technology statements have no equality in terms of demand. Some are relatively straightforward, whilst others are extremely complex. At each level it was decided that eight marks should be available. In some instances all eight marks were targeted on a single statement (Te4 level 7) but in the majority of cases they were distributed between two statements (Te1 all levels). The main purpose of the trial was to establish an appropriate level of mastery. The questions could be fine tuned following the trial to ensure satisfactory rates of achievement but there was no established definition of what was satisfactory, i.e. what percentage of pupils should achieve certain levels. The mastery level did, however, have to be decided prior to the first statutory assessment on the basis of the trial.

The development of the pupil material - Practical tasks

1. 1990-91

As discussed at the beginning of this chapter, before assessment could take place pupils had to be provided with a realistic task which would generate appropriate assessment evidence. Initially these tasks asked teachers to operate with pupils in a style and fashion with which they were unfamiliar. To ensure that the tasks which were developed promoted the philosophy of the Order the following principles were established; each activity should:

- offer pupils the opportunity to undertake a task encompassing the four attainment targets, which would enable them to demonstrate their capability in an holistic fashion;
- provide a structure in which pupils could identify needs and opportunities which would form the focus of their task;
- allow pupils to undertake a task in any of the disciplines which fall within the design and technology federation;
- provide pupils with a valid learning experience which drew on the key stage 3 programme of study;
- provide pupils with a satisfying and successful experience which promoted progress and achievement.

The tasks also needed to gain the support of teachers, so in addition each activity should:

- offer teachers the opportunity to take ownership of the activity and adapt its delivery to suit their own circumstances and teaching styles without compromising the standard nature of the activity.
- be sufficiently robust to withstand the variety of teaching styles, curricula organisation and staff collaboration evident in schools.

Research into structured activities

Three models/structures of delivery were devised which offered sound approaches to design and technology and operationalised the above principles. A number of models was investigated to explore the effect of the structure on the assessments and the consequences for both reliability and validity.These models were described as 'broad', 'focused' and 'guided'. The rationale governing each of these structures was as follows:

Broad - this structure interpreted Te1 in its broadest sense, consequently it was the most demanding for the teacher, as the possible outcomes were varied and diverse. Pupils were given the task of identifying needs and opportunities within the boundaries of a selected theme.

Focused - this structure focused pupils' initial response to Te1, by defining certain parameters relating to the theme. The defined parameter used during trialling was the specification of the location, but there are many other possibilities.
Guided - this structure guides the task by starting from a particular attainment target and guiding the pupil through the activity back to the starting point. During trialling the task started at Te4. Pupils started by identifying potential design and technology tasks after evaluating the work of others; following the resolution of the identified situation, pupils returned to Te4 to reflect on their own activity.
The three themes on which the research was based during 1990 and 1991 were: Broad - Exhibition; Focused - Public Places; Guided - Measurement.

Common features of the tasks

An essential aspect of any statutory assessment procedure is standardisation. To ensure consistency in relation to the evidence produced by pupils, all tasks were provided with a common level of resourcing, both for the teacher and the pupil. The three common aspects are described below:

The Video
Each task commenced with a video which established the context. The video aimed to motivate the pupil by illustrating the breadth of the theme and creating a focus for initial thinking and investigation. It provided a standard input and a baseline of information common to all, from which each pupil could progress. It provided teachers with a quality resource which would lighten the preparation load at the outset of the activity, allowing the teacher to pay due attention to delivery and assessment requirements.

(Other forms of resource packs were investigated and found to be inflexible and prohibitive in relation to cost. The intention would have been for the task video to be broadcast annually via national schools' television and recorded by school or LEAs. Television networks had given outline agreement and may even have undertaken the production. The videos produced typically cost about £15,000 (1991 prices) or approximately 2.5p per pupil (national cohort of 600,000). It is difficult to imagine any other resource with the richness of a video which could be produced for this cost.)

Quest
Following the video each pupil undertook a Quest activity. This was an assessment device which could be used either to focus or guide the pupil to examine the context from a particular perspective. This enabled the pupil to get inside the theme by directing their thinking through his or her own experience. The quest included both a research activity and a short task of up to 2 hours. The objective of

the short task was to sensitise pupils to issues related to the task context, illustrate the scale of the task they might tackle, allow pupils time to consider what their self-determined task might consist of and give teachers the opportunity to discuss these ideas with pupils and negotiate tasks appropriate to their capability. Short tasks promoted active engagement rather than reflection. The quest also provided additional assessment evidence and could be designed to focus pupils' attention on specific statements.

The Labels

Once pupils embarked on their personal activity they were asked to use label sets to organise their project folder. These labels served several functions. They involved the pupils in the assessment process by informing them of the aspects they need to address whilst aiding them in creating a structure for their work. They assisted teachers in carrying out assessment in that they helped to standardise responses and show the aspects which each pupil had identified as being of importance. Pupils only use the labels they required. Clearly those labels not used also provided important assessment information. The labels offered a flexible assessment tool which allowed each pupil to demonstrate his or her capability in a way which was most appropriate to the task they had in hand. In all trials the labels were found to be extremely effective in assisting teachers in making assessments.

The labels empowered pupils. Pupils will only evidence their capability if they know what is required of them - they need to be aware of the assessment goal posts! The labels provide them with this information. The labels cover comprehensively all assessment items from level 3 to 7. They help children structure their work and provide signposts which will assist them when a teacher is unavailable. The labels, in the main, do not address specific SoAs, they address issues which impinge on a number of statements. The importance of the labels cannot be underestimated. A Polish proverb says, "it is easy to turn an aquarium into fish soup but try reversing the process!." Pupils' design and technology work, without the labels, is like the fish soup. It may in itself be excellent but it can be demanding to identify the component parts, if indeed they existed! It is difficult to assess an outcome if it has no framework or structure and it is not cross-referenced with the statement of attainments. Unlike the soup, it is possible to reverse the process but it is time consuming; it is far easier spotting fish in water!

The use of the labels did not necessarily imply a statement has been achieved, that judgement was still one a teacher had to make on the quality of the evidence. However, the labels helped teachers identify the relevant evidence. The labels provided a flexible tool which enabled each pupil to demonstrate their capability in the way which was most appropriate to the task which they had in hand.

The labels indicated the tasks and operations with which pupils should engage. They were drawn from levels 3 to 7 - the KS3 range. Many of the words on the labels were ones which some 14 year olds had difficulty in understanding. They represented the language of design and technology, they are concepts which teachers must start introducing pupils to from Y7. To use only "child speak" would trivialise the concepts and prevent pupils from positively addressing the real issues required to secure a level.

The video, quest and the label system provided a structure which when undertaken within a time window created a standard format for a task which could be repeated to make valid comparisons. It is the structure of the task in the context of design and technology which provided standardisation not the chosen theme. If the structure proved to be robust and reliable themes could be, 'poured in' at will.

1992

Research carried out during 1991 and 92 resulted in a revised specification which recognised some of the difficulties certain of the practical tasks had revealed. All tasks created logistical challenges which not all schools could overcome. A national assessment had to be capable of being undertaken by all schools. Schools still working towards implementing the Order had to be able to carry out the activity as well as those that had fully embraced the design and technology approach. The activity also had to assist the traditional schools to come, 'on stream' without unduly compromising those which were successfully delivering genuine design and technology. Pragmatic decisions had to be taken to insure comprehensive participation. The two key ones were that:

- all pupils would undertake tasks which had a common theme, which would be placed in a context, thus pupils' responses would be focused;
- to assist organisation and management tasks were prescribed in relation to specific categories of materials, e.g. construction materials, textiles and graphics.

Hence the practical task for the national 2% pilot established a framework for a curriculum activity. It was designed so that all areas of the design and technology federation could participate. It was firmly based on the Statutory Order. For legal requirements, it was essential that positive linkage to the Order could be demonstrated. The theme selected was batch production, and tasks were set in the context of a school fair being run to raise money for a charity. The task bank was based on the materials with which pupils should work during the key stage - as defined by the Order (p 19 National Curriculum Order for technology) - with the addition of control, providing five tasks in total. The tasks were designed to be non-

level specific, differentiation being determined by outcome. To ensure pupils operated at appropriate levels and set realistic goals, two strategies were developed.

- The programme of study to be taught during the task was provided in two forms. The overarching key stage 3 statements offered teachers general guidance whilst the relevant, "working towards level" statements were listed so that teachers could determine learning objectives for individuals;
- Each pupil was provided with a structure chart. This consisted of a clear set of issues/goals which should be considered in producing a solution. The charts were produced in three bands: levels 1 - 4, 3 - 7 and 6 - 10. These charts were developed from the label sheets which had proved so successful (label sheets were ruled out because of cost implications).

All five tasks consequently had two common objectives. Pupils were asked to:

- undertake a design and make task associated with a school fair which might be held to raise money for a charity;
- the outcome should be capable of being produced in a limited quantity. Although there was no specific Business Studies task, the nature of the activity was such that the programme of studies elements dealing with business could be seen to permeate all the practical tasks.

1992 Pre-statutory assessment trial

Political decisions became increasingly more prominent in determining the style and nature of assessment procedures. The task which evolved was distinctly different from any previous activity as it was the first to assess only Te 2 and Te 3 (Secretary of State's decision January 1992). Proposed tasks were submitted to SEAC for scrutiny during the first phase of development. Initially these materials were based on the 1992 pilot practical task. The tasks all had the same theme with a common context; there were five different tasks (construction materials, control, food, graphics media and textiles) and differentiation was solely by outcome.

The first major development involved devising activities which would, in part, differentiate by task. This was a novel development in relation to task setting in a practical context. Several approaches were investigated in attempting to devise a workable solution. The following objectives were used to evaluate these approaches. The task(s)

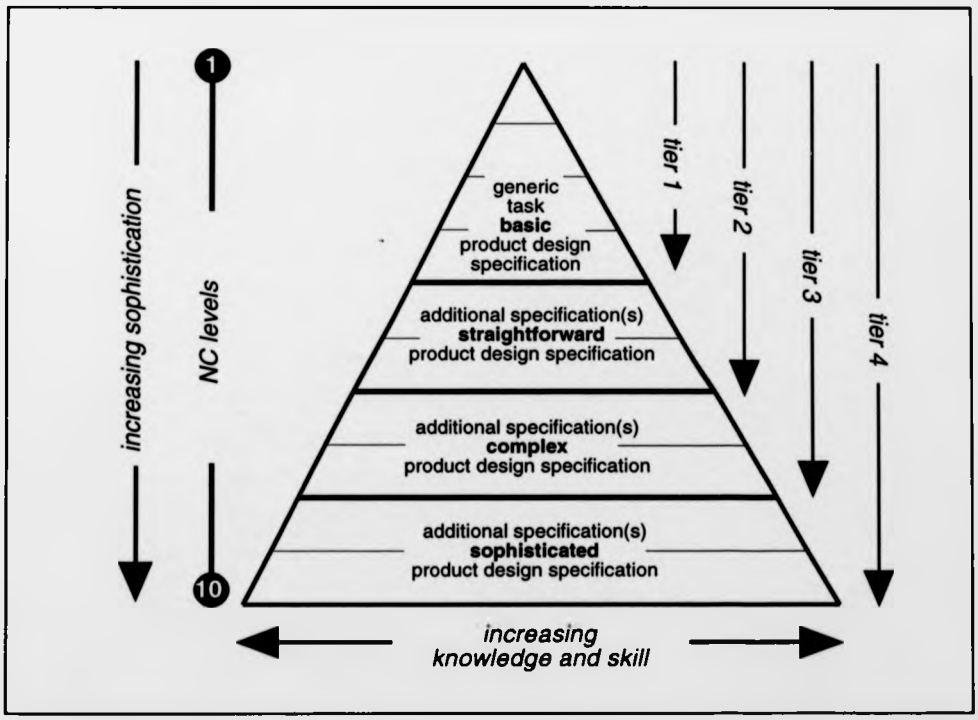
- should allow all pupils, undertaking the task through a particular material, to do the same task at a level appropriate to their capability;
- should be comparable across each band, regardless of the material(s) in which the pupil was working;
- should focus on assessing the level of achievement in relation to an attainment

target rather than the statements of attainment evidenced;

- should be assessed using criteria which reflect the task taken by the pupil; these criteria should be comparable across the tasks;
- should clearly indicate progression which should be stated in the assessment criteria.

The final model did achieve these objectives. Each task was based on a core activity which required pupils to: "design and make an artefact.'
As differentiation was based both on the complexity of the task and each pupil's likely performance in relation to it, teachers had the task of deciding in which tier to enter each pupil. To ensure there was sufficient overlap four bands were advocated: 1-4, 3-6, 5-8 and 7-10.

Differentiation by task was achieved by adding an additional design requirement at each band of entry. This made the task, at each band, more demanding and sophisticated, both in relation to designing and making. This was not because the additional demand at each band was necessarily the most complex, it was the cumulative effect and the relationship of all the design requirements which made a task level 4, 6, 8 or 10. Each of the three tasks was therefore prescribed in four levels of complexity, early proposals are shown in appendix 4.9, p. 315. This structure is detailed in the diagram below.



The selection of materials in which tasks were developed

The model for the practical task was initially developed in relation to five material areas. External expertise expressed concern that standardisation across five tasks would create difficulties in relation to reliability and validity. A single task would have produced the most standard context for assessment and this approach was advocated by Civil Servants and non-subject specific experts. However, a single task would have created a situation which could not have been managed by the majority of schools in the time window which would be allowed for the statutory assessment. The political wish was for a single task in construction materials which allowed possible solutions via control. Data from previous trials was used to justify the increase in the number of tasks to three. The reduction of possible tasks from five to three was however the most controversial aspect of the first statutory assessment. This reduction created significant issues in relation to both morale (teachers of textiles felt excluded) and organisation and management.

Previous research was used to establish that Food offered the best opportunity to alleviate some of the management issues (in previous trials/pilots, between 25 - 30% of pupils were entered for a task in Food). There were also good reasons for including a task based on Control - a term used to encompass references in the programme of study to both 'components' and 'Control.' The argument was eventually accepted that there would be three practical tasks, one each in Construction materials, Control and Food. This decision was seen by many teachers as arbitrary and provocative.

Two tasks were developed in relation to each of the three identified areas. From these three were selected and these tasks were developed into a form in which they could be trialled. Tasks which specified the design requirements prompted pupils to achieve appropriate levels because clear goals were established. In addition instruction sheets were required to ensure pupils provided adequate evidence of the process and decision making aspects of the task. These were evolved from both the flow-diagrams and labels used in previous trials. The instruction sheets were based on the over-arching programme of study statements for key stage 3. The instruction sheets were based on the same rationale as the tasks - at each tier additional instructions were introduced. Hence the instructions were cumulative from the lowest tier to the highest. The instructions also linked to the statements of attainment being assessed and helped pupils to provide appropriate evidence. This linkage would have been clearer and more obvious if the Order had been more explicit in this respect.

The development of the pupil material - Tests

1. 1991-92

Following political decisions, described in chapter 3, the notion of assessing design and technology capability became an urgent issue towards the end of 1991. Limited trials were conducted, which attempted to produce assessments in relation to all four attainment targets via written tests. Some strategies were identified worthy of further investigation. However, following the Secretary of State's, January 1992 decision on statutory testing in technology the focus changed. The assessment Order for 1993 established that only Te1 and Te4 were to be assessed via a written test. It was therefore essential that this pattern of testing was adopted for the 1992 national pilot. Because of the requirement to have written tests which addressed all ten levels, it was inevitable that there would need to be different tests aimed at different levels. Three tiers of entry were established: levels 1-4, level 3-7 and 6-10. This was felt to provide the most straightforward approach for teachers in making entries. The majority of pupils would be entered for the middle tier which covered the key stage range. The lower and upper tiers ensured that pupils with learning difficulties or high achievers would be offered tests at an appropriate level.

It must be stated that the statements of attainment which tests of Te1 and Te4 sought to assess were not designed as assessment criteria for this style of assessment. It has been noted earlier in this chapter that the statements are ambiguous, lack clarity and clear progression and are unsatisfactory as assessment tools in a statutory assessment context. The statements describe attributes, of what is essentially a practical activity. The ability of current fourteen year olds to address such statements in a theoretical context, drawing on their experience of design and technology was doubtful. It is likely that with experience, guidance and planned teaching pupils' capacity to answer tests of this type will improve, however, this will only be achieved at the expense (in terms of time) of the unique practical activities which this subject offers.

Criteria for test development

The tests were developed to conform to the following features:

- questions were targeted on specific statements of attainment;
- the same question addressed the same statement regardless of the test, i.e. question 4 on test 2 and question 1 on test 3 are identical and both address statement 6c in Te 4;
- questions in each section, started at the lowest level in the band and work towards the highest level;

- each test had three inclines of difficulty;
- where the Order allowed two statements are assessed at each level in each attainment target.

Question banks were developed covering the ten levels, resulting in 40 distinct questions, 4 at each level. The questions in a particular section, across all three test papers, were written to have a coherence and logic. Entry points to the sequence were provided at levels 1, 3 and 6. The question bank allowed logical question sequences to start from each of these three levels. Issues are not assessed twice in a test band, but the same issue may re-occur at a level outside a particular test band. This imposes considerable constraints on those setting the questions.

The questions in the test were designed to extend pupils' thinking from the long task activity and to draw on their experience from across the key stage. The test assessed Te 1 and Te 4, the aspects of design and technology which are peripheral to the mainstream activity of designing and making, the aspects which receive a lower priority in learning objectives and which rely, to a large extent, on pupils' experience, perception and awareness of the world in general. The type of question which could be used to probe the statements did, in the majority of cases, require a written response. Given these conditions the tests were inevitably, as much an assessment of a pupil's general level of intelligence and literacy as about aspects of design and technology which might have been either taught or experienced. Each Section of the test paper posed a different set of issues for those writing the questions, each Section had its own distinctive flavour.

Apart from Section A , a decision was taken to ask children to apply their knowledge and understanding to new situations rather than report previous experiences. Phase A trialling clearly demonstrated pupils' lack of motivation when faced with recounting an activity which, for many, may not have been stimulating the first time round.

Questions were developed by taking the programme of study elements identified for the practical task and interpreting them through the statement of attainment being assessed into the identified context and theme of the test. All the questions were underpinned by two key issues:

raising money for a charity;

issues relating to batch production.

The creation of a distinctive incline of difficulty in the questions was essential even though in some instances this was clearly not the case in the statements of attainment.

Section A

The questions in this section asked pupils to evaluate their long task activity, and it is therefore, primarily reflective. As long task experiences would have been extremely varied, the questions were less specific than in the other two sections. For example, the question addressing Te 4.6c.

a) Give TWO reasons why the materials you used were suitable for the manufacture of your design.

b) Give TWO reasons why the materials you chose gave your design a pleasing appearance.

Pupils' answers to this question could obviously be quite different. This style of question had implications for a mark scheme and for moderation.

Section B

A new context was established at the start of this section, which was also pursued in Section C.

"The charity fair in your school was a success. The charity still needs more money. You and your friends have decided to raise more money by holding a badge day. The badges you make will be sold in your local high street."

The new context allowed pupils to be assessed in relation to identifying needs and opportunities. The notion of selling badges is offered as an example in the programme of study. It occurs in the, "satisfying needs and addressing opportunities" strand in the section detailing, "pupils working towards level 10 should be taught to.....". Questions in this sections were generally of two types, children being asked either to use design and technology knowledge:

to describe and give reasons for how they would go about doing something;

or

to interrogate information e.g., pictorial or statistical, reach a conclusion and justify their decisions.

An attempt was made to balance these approaches at each level, but the nature of the statements did not always allow for this approach. A good example of a level where this balance was achieved was level 5.

statement Te4. 5b

How would you take the following into account in developing a design for a badge? Give ONE example under each heading.

A functional aspect

A visual aspect

An environmental factor

statement Te4 5a

From a questionnaire you discover the following:

(see appendix 4.10, p. 318 for an example of the test paper, ref. Que. 11)

a) Which information would be the most useful in deciding on the selling price of your badge?

b) Give a reason for your choice.

Section C

This section of the test paper essentially assessed pupils' critical and evaluative capacity by making judgements about the work of others. Prior to answering questions in this section pupils were asked to look carefully at a picture sheet. This provided a range of information about a set of possible solutions to the scenario established in Section A. The sheet provided information on the design of two badges, sequential pictures describing how they were made and a picture of the badges at the point of sale.

Decisions concerning the intentions of the programme of study had to be taken. The lack of clarity in identifying precisely what pupils would have done by the end of the key stage posed significant problems. Expert judgements had to be taken in relation to equipment, processes and techniques. These judgements could have been at a relatively low level of expectation, however, it was decided to illustrate a level of rigour believed to be appropriate to 14 year olds.

In addition decisions had to be made about the materials in which the example badges were produced. Selection is surely a key aspect of any testing procedure. Such selections are bound to create dissatisfaction amongst the range of subject experts contributing to design and technology. This will continue to occur whilst teachers perceive themselves as subject experts first and design and technologists second. The questions ask children to give reasons and justifications for why things might have been done, identify areas of possible improvement and to offer and justify alternative approaches. A good example of this approach occurs at level 7.

Statement Te4. 7c

The photograph on the accompanying sheet shows a point of sale dispenser for the badges.

a) Identify THREE aspects which you think could have been improved.

b) Describe, using labelled drawings, how you would improve TWO of these aspects. Your answer should provide details relating to choice of materials and manufacturing procedures and techniques.

1992 to 93 - The pre-statutory assessment trial - tests

The test materials for the first statutory assessment were developed on the basis of a full evaluation of the previous year's pilot. The following recommendations were adopted:

- the tests would be offered in four tiers of entry: 1-4, 3-6, 5-8 and 7-10;
- the assessment of Te4 would be closely linked to the practical task.

The first recommendation reduced the number of levels in a test paper, allowing statements to be probed in greater depth. The second allowed more precise, probing and rigorous questions to be set in relation to Te4. The combined effect of these two decisions, four tiers of entry and three different practical tasks, resulted in the requirement for 12 test papers and accompanying mark schemes. The following approach was developed for the tests:

Section A -Te4

Pupils were asked to evaluate the work they produced in response to the practical test. Each set of questions assessed the same statements from Te4, but in the context of the programme of study identified for each of the practical tasks.

Section B - Te1

This section was totally independent of the practical task and common to all pupils. It established a new situation/scenario and assessed the selected statements in this context. Aspects, to be tested, were selected from the whole of the programme of study for the key stage. There were no material specific questions, as Te 1 does not view design and technology in this way, appendix 4.11, p. 321 provides an example.

The development of questions -Te4

All of the questions, in the assessment of Te 4, asked the pupil to reflect on the practical task and evaluate the way in which they had tackled the task and the final outcome. To ensure comparability, the three sets of questions designed to assess achievement in relation to Te 4 were developed on the basis of the following common criteria:

- the questions reflected the programme of study identified at each level in relation to each task;
- the same statements were assessed regardless of the material context;
- the same items within statements were assessed regardless of the material context;
- wherever possible, the questions were framed in the same style;
- wherever possible, the response required was similar, e.g., a drawing.

Issues in constructing questions

Even though the pupils had tackled one of three focused tasks, it was still difficult to produce precise questions. In framing questions, an assumption had to be made about the kinds of ways in which the task was tackled and the range of outcomes. For example, should the questions assume the pupil completed the task, produced an outcome and had the opportunity to see if it worked? Some statements require comment on issues such as possible improvements, overcoming problems, and changes and modifications. Questions concerning such issues either had to make assumptions about what took place or the mark scheme had to recognise a very wide range of valid responses. For example, if a pupil answered the question:
Describe one improvement which you would make to your ..... (Que. 10d - Te4.5a)
by responding, 'it could not be improved in the time,' was this acceptable? Because the response is valid, ways had to be found of asking questions which would produce a more positive response. There was also a difficulty, with four overlapping bands, of creating four plausible starting points for the tests. Each had to make sense to the pupil, regardless of the level at which the test being taken started. The four starting points were level 1, 3, 5 and 7. Consequently, to avoid repetition, an aspect focused on in a question at level 3, on the tier 1-4 test, cannot be readdressed until level 7, on the tier 5-8 test. Within these constraints, the questions had to be credible whilst probing systematically a pupil's capacity to evaluate what they had done.

Comparability of questions

To be fair, questions had to make comparable demands on pupils regardless of the task which they took. In many cases this was straightforward, as a very similar question could be asked in relation to the three tasks (questions at level 4). In others this was not the case, for example, Te4 statement 6b, this statement requires pupils to be able to:
"devise and carry out ways of testing the extent to which the product satisfies their design specification."
The questions devised to test this statement illustrate how this was resolved.

On the construction materials test, the question asked was:
If you had to test three different clamps, how would you find out which one held a 1mm piece of material most tightly?
a) What things would you need to use to test the clamps?
b) Describe how you would use these things to do the test.
c) Give two limitations of the test you have described.

On the control test, the question asked was:
Circuits do not always work first time, if your beacon board did not work:
a) Describe one thing that you would do to test if the cells were faulty.
b) Describe one thing you would do to test if the cell holder was faulty.
c) Describe two things you would do to test if the circuit was faulty.
On the food test, the question asked was:
Imagine your snack bar is to be produced for pupils from your school going on a sponsored walk. You would need to carry out tests to see if the finished product needs to be improved. Describe a test you would carry out for each of the following:
Test for texture; Test for appearance; Test for taste, Test for smell.

In each case the question probes the key requirements of the statement, devising and carrying out tests, through the context of the practical task and the material context in which it was set. Each of these questions had four distinct mark points, half the marks available at this level. The questions were framed in a hypothetical fashion but all dealt with issues which pupils should have tackled or have experienced. Consequently pupils who had been encouraged to test their work were able to reflect on first hand experiences when answering these questions. Many pupils however, would not have been asked to evaluate the final product by testing it, so the question had to be framed in a way which would still make it accessible to such pupils. These similarities established fairness in relation to structure and approach. Level of difficulty in relation to subject matter was more difficult to determine and could only be successfully established by analysing correct response rates from a significant population.

Mode and relevance of response

A pencil and paper test was clearly not the way in which some of the statements were intended to be assessed. Statements such as:
Te4.3a - discuss their design and technological activities and their outcomes with teachers ..... Te4.6b - devise and carry out ways of testing .....
are either based on verbal interaction or are practical and purposeful in their intent. In addition the majority of pupils was asked to carry out the Te4 test on evaluation, well after they had finished the task and without the folder containing the work relating to the task. As Bruner (1960) commented:
'It would seem much more sensible to put evaluation into the picture before and during ..... as a form of intelligence operation.'[6]
The way in which Te4 was assessed would probably have been equally pointless and unhelpful to the majority of pupils as it was not even connected to the actual activity. As an educational experience many of the questions would have been extremely pertinent and positive, if asked whilst pupils were engaged with the task.

To assist pupils in engaging with the questions and to prevent the tests being based only on memory and recall, questions were frequently framed to help pupils recall what they had done before, probing the reasons for taking decisions and evaluating the consequences. For example at level 4.
a) What was the most difficult decision you had to make about the design of your snack bar? b) Why was this a difficult decision? c) Now you have made the snack bar, would you still make the same decision. Answer YES or NO and then give one reason.

Although marks were available for answering each part satisfactorily, it was impossible for a pupil to achieve mastery at a level without gaining marks from parts of question, such as c) in this example. At level 4 there were three parts of the questions which required these types of evaluative responses.
4b) Give one reason why you decided not to use this idea
5c) Now you have made the snack bar would you still make the same decision. Answer YES or NO and then give one reason.
6b) Give one reason why your snack bar would be suitable instead of a meal.
As the mastery level was 5 out of 8 no pupil could achieve level 4 without providing a satisfactory answer to one of these parts of the questions asked at this level.

Wherever it was reasonable, in the context of the statement, pupils were asked to use an appropriate skill for communicating the answer, such as drawings. Traditionally pupils have not been encouraged to draw or sketch in some aspects of the subject, notably food. However there are many statements in the programme of study and statements of attainment which legitimised this approach. For example, at level 5.
"Develop styles of visual communication which take account of what is to be conveyed."
(Programme of study, working towards level 5 - Developing and communicating ideas)
This programme of study statement linked to statement Te4.5a:
"Evaluate their product in relation to the design intentions and to the original needs....."
allowed the following question to be asked.
You were asked to design and make a clamp which could be unlocked using just one hand.
a) Use a drawing to show how you did this.
b) Label your drawing clearly to explain how you did this.
On the control test the question was asked in relation to designing and making a switch and on the food test, designing and making a snack bar which could be divided into small pieces.

The development of questionsTe1

The specification for 1993 created a very close link between the practical task and the assessment of Te4. This was achieved by having no overarching context for the three practical tasks. Consequently, there was no common aspect to the practical tasks which could be revisited in the assessment of Te1, as in the 1992 national pilot. The test, in relation to Te 1, was therefore quite independent of the practical task and was based on a situation of which pupils were unaware, prior to taking the test. All pupils took the same test of Te1, at the appropriate tier, regardless of the practical task taken. The criteria for developing these questions were as follows:

- the question could draw on any aspect of the KS3 programme of study;
- each question was focused on a statement, or items within a statement.

A context needed to be established in which questions could be framed, in relation to identifying needs and opportunities. The following factors were taken into account when selecting a context. The context should:

- not be seen to favour pupils who had taken one of the three practical tasks;
- be as neutral as possible;
- offer opportunities for questions to be asked in relation to any aspect of design and technology;
- present identical information to the pupil, regardless of the band of entry.

Several potential solutions, based on a range of contexts were investigated. Following discussions with SEAC a choice was made and questions were developed around the context of the interior of a waiting area. Following the practical tasks, which were product-orientated, a decision was taken to focus the Te1 questions on the needs of people. Care was taken not to indicate what the people were waiting for and to keep the context neutral, i.e. area rather than room. Pictures were selected as the stimulus material, as trials had indicated that pictures were most accessible to pupils at every band of entry.

A model was developed based around four pictures. The questions at each level focused on one of the pictures. The result was that each pupil answered four sets of questions, each set relating to a different picture. (It was observed that following the test pupils, regardless of the band of entry, felt they had done the same test because they had answered questions in relation to the same pictures.) The question pattern in relation to pictures is shown below.

level	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
picture	A	B	C	D	A	B	C	D	A	B
tier	1	1	1/2	1/2	2/3	2/3	3/4	3/4	4	4

The pictures covered the following issues:
anthropometrics, economics, ergonomics, health and safety, materials, refreshments and storage.
The questions were designed to draw on pupils' understanding of:

- questionnaires, interviews and observations in helping to identify needs and opportunities;
- how to take decisions about what might be worth doing;
- thinking about the implications of choosing what they might do;
- resolving conflicts views and opinions.

Questions were then framed to assess the statement. For example, at level 6 statement 6a requires that pupils should be able to:
"Explain how they have identified needs and opportunities for design and technological activities and give a justification of the conclusions they have reached."
The question told pupils to start by looking at a picture. The picture showed people sitting in a small cluttered waiting area, the chairs are clearly not comfortable. The question then posed was as follows:
You have been asked to design a better chair. Describe two important things about chair design for each of the following.
a) The comfort of the people sitting on the chairs.
b) The convenience of the cleaner of the waiting area.
c) Describe two important design features for a chair which would meet the needs of both the people sitting on the chairs and the cleaner of the waiting area.
The question focused on identifying needs in relation to two different users and then resolving conflicts.

Mode and relevance of response

As the context of the questions was new to the pupils the questions tended to be longer as they needed to be fully informed of the situation. Consequently the time taken by pupils to read and understand the question and relate it to the picture at which they had to look, had to be considered when considering the response mode. This was an especially important factor at the lower levels. Multiple choice questions were used to overcome this issue. Significant problems were encountered in finding satisfactory distractors which did not trivialise the topic yet were not too close to create doubt in the minds of markers. The facility of the distractors could only be established by trialling with pupils of the appropriate age. This was deemed the only reliable method as subject experts had difficulty in reaching agreement. The following example is the question set in relation to

Te1.1a

Describe to others what they have noticed in familiar surroundings or visualised about imaginary situations

Look at picture A. Why is the woman in the picture finding it difficult to choose a magazine?

Tick the boxes besides the two best reasons.

- because the magazines cost too much ☐
- because the magazines are untidy ☐
- because the table is too low ☐
- because she has been waiting too long ☐

Picture A showed an elderly women in a waiting area having problems selecting a magazine from an untidy pile on a low table. The middle two responses were the correct ones and the fourth option the close distractor. The purpose of this question was to differentiate between pupils working towards level 1 and those having achieved level 1. Consequently in determining the difficulty of the distractors all pupils who achieved level 2 and above were disregarded. It could then be observed that this question was at an appropriate level as the most frequent reason for failing was an inability to follow the instructions, for example, ticking more than two boxes. This indicated that the pupil could not discriminate and make the judgement required. In Te1 all responses were made either by selection or in writing. It was not deemed appropriate to ask pupils to draw or sketch.

It was important that pupils' answers were focused so that the mark scheme could be as specific as possible. Consequently questions identified particular aspects, in relation to an issue. For example at level 7, the question written to assess both statement 7a and 7c could have been:

Look at picture C. Many of these people would like something to eat or drink, while they wait. Imagine that the person responsible for the waiting area has asked you to look into the possibility of starting up a small refreshment area, what would you do?

(Picture C showed a number of people who have clearly been waiting a long time.) However, the question in this form would have produced answers which could not have been marked in either a standard or consistent manner. The question used was as follows:

Look at picture C. Many of these people would like something to eat or drink, while they wait. Imagine that the person responsible for the waiting area has asked you to look into the possibility of starting up a small refreshment area, but is concerned about the following:

a) the cost of starting up; b) litter; c) health and safety; d) what food to provide.

Describe two things you would need to investigate about each of these to provide enough information for the owner to decide whether to have a refreshment area. The question now has eight specific mark points and the criteria for acceptable answers can be given to the markers along with acceptable and unacceptable responses.

General language 'rules' resulting from trialling

It was essential in developing pupil material relating to tests that the language used was appropriate to the age of the pupils taking the test. From trialling, rules were established which were then used in the writing of the questions, which were used in the first statutory assessment. The rules established were as follows.

- Removal of conditional clauses: sentences starting with 'If you had to...' followed by an instruction within the same sentence can frequently be rewritten in the form of a brief factual or contextualising sentence, then a further sentence with a direct instruction. This is part of a more general process of separating contextualising information from the question or instruction itself. Where possible, information should be placed at the beginning of a question in a separate sentence or sentences, followed by the instruction.
- Where possible, conditional tenses should be replaced by the present tense. Past conditionals are especially likely to cause problems and are fairly easily avoided. These tend to occur in Te4, when pupils are evaluating their practical work. 'If your beacon board had not worked.... what would you have done to test.....' is quite reasonably replaced by 'If your beacon board did not work.... what would you do to test....' (This is an example where a conditional cannot always be entirely removed but shifting out of the past version of the tense gives a much more direct meaning.)
- Wherever possible, complex sentences should be divided into shorter sentences.
- Passive forms of verbs should be replaced in almost all instances. e.g., 'Using drawings, show ....' should be 'Use drawings to show ...'
- Other examples of passive or indirect language should be made more active and direct, e.g., ' strong enough in use' has become 'strong enough for someone to use' 'Which club has a meeting on a Wednesday?' should be 'Which club meets on a Wednesday?' This removes confusion caused by abstraction.
- Abstract phrases in common usage but with other concrete meanings should not be used e.g., the phrases 'stand up to' and 'carry out' should be replaced by language that does not conjure up misleading graphic images of some kind of resulting action. The phrase 'in each case' should be replaced by a direct

reference to the object in question.

- Quotations from the task sheet for the practical task are confusing in the context of the test and add nothing to the meaning. It is preferable to use plain text and direct description, e.g., 'You were asked to design and make a switch fitted to the board....'
- Direct prompts should be used to introduce a question and focus the pupil. 'Think about the snack bar which you made.'
- Further to this, subordinate clauses should not be used. Shorter, direct sentence structures should be used wherever possible.
- Vocabulary should be as simple as possible, for example replace Latinate words with Anglo-Saxon monosyllables. For example, in place of 'criteria' use 'points', and instead of 'employ' substitute 'use', etc. Equivalents do not exist for all words which might be difficult. In these case words should be made available in the form of a glossary.
- Abbreviations should be avoided, as they create unnecessary difficulties for pupils with language difficulties or those for whom English is a second language. e.g., St. as an abbreviation for street.
- With visually impaired pupils in mind, reference to visual action should not be used, e.g., 'Look at the information on the screen' should be replaced by 'Read the information on the screen'.

General 'rules' concerning layout

It was essential in developing pupil material relating to tests that the layout used was clear and straightforward and applied systematically. From trialling, rules were established which were used in the design of the questions used in the first statutory assessment. The rules established were as follows.

- Illustrations and other information relating to a question must be contained within the question and should not, for instance, be located inside the cover of the test book. Cross-referencing is thus avoided, since this can cause confusion for pupils.
- Where complex information is provided in the form of screens or illustrations, a note of the identity or purpose of the provided information is helpful. For example, two similar illustrations should be labelled, 'frame A' and 'frame B'.
- Connecting questions have been arranged ***as a), b), c), etc.,*** rather than expecting pupils to carry information over from one question to another.
- Where discrete elements are required within a question, the question should be divided into constituent parts, each task being completed wherever possible before the next one is requested.
- Where a question falls into a number of parts, it might not always be possible to avoid some follow-on from a stem. Prompts should be added to each part to

add emphasis to the instruction in the stem.

- Prompts in the layout should reiterate instructions.
- Locational clues should be built into instructions where these are likely to help all pupils find their way around the paper.

Conclusion

Perhaps this chapter more than any other exemplifies the significant issues with which this research had to grapple. It would have been a sufficient challenge if design and technology had been a single coherent subject with a homogeneous teaching force, or had there been no political interference accompanied by changes to the specification, or had the mode and manner of testing not been regularly changed; however, as all these pertained to a greater or lesser degree throughout the four year period the task became unduly complex. The variety of research tools described in this chapter is a direct result of the external pressures which were brought to bear on this research. Without these pressures a more coherent approach could have been adopted and consequently a more refined product may have emerged at the end of the process. However these pressures have resulted in a wide range of ideas and strategies which, it is hoped, will be of value to those investigating the assessment of design and technology in the future.

The research involved in operationalising the Order was essential if the final assessments were to be valid and reliable. The shift from the precision of statements of attainment to broader level descriptors was recognition of the limitations of statutory assessment. The eventual adoption of a best-fit approach resulted in teachers only having to make a single decision in relation to each attainment target. This obviated the need for complex aggregation rules which were and are in conflict with criterion referencing. This approach to assessment may initially seem cruder, but it offers a more practical and realistic route to obtaining a fair and acceptable means of statutory assessment. It is particularly relevant to practical tasks where, this research has confirmed, teachers prefer to make holistic judgements rather than a large number of atomistic decisions.

The range and variety of pupil materials developed is also testimony to the complexity of this research. Each specification required a review and modification to both practical tasks and, in the later stages, tests. Consequently, a wealth of pupil materials was produced which found their way both officially and unofficially into a large number of schools, permeating the curriculum with varying degrees of permanency. The evolution from assessment by outcome to an approach which also defined progressively difficult tasks was extremely influential on the way in which GCSE boards approached syllabus development. The tasks described in this chapter were the first of their kind in this subject area and represent a model

on which future developments can be based.

The complexity and variety of research instruments had serious implications for their trialling and piloting, as is apparent in the next chapter. This had serious implications for their evaluation and refinement given the pressures under which this research was undertaken.

References

1. Layton, D. (1991). *Aspects of National Curriculum Design & Technology*, York: National Curriculum Council, page 2
2. The Assessment of Performance Unit project in design and technology (1991). *The Assessment of Performance in design and technology*. London: Schools Examination and Assessment Council, page 23
3. Graham, D. (1989). *The Introduction of the National Curriculum*, lecture, Cambridge Institute of Education, 16th. May
4. Gipps, C. (1992). *Assessment - A teacher's guide to the issues*. London: Hodder & Stoughton, page 95
5. Nutall, D. (1989). 'National Assessment - Will reality match aspirations?', Paper delivered to the conference 'Testing Times', Macmillian Education
6. Bruner, J. (1960). *Towards a theory of instruction*. Cambridge, Mass: Harvard University Press, page 30

Chapter 5

Pupil Performance in Major trials and Pilots

Synopsis

In determining which aspects of the 1990 Statutory Order could be assessed in a valid and reliable way it was essential to conduct systematic trials and pilots of potential assessment materials. In this chapter the results of principal trials and pilots are reported and analysed using similar statistical procedures (in total 15 significant trials were conducted). The data presented in relation to pupil performance relate mainly to the assessment framework but also to the design of the instruments against which pupils were assessed. The aspects of pupil performance being tested varies from phase to phase, as the process changed both in response to data gained from previous trials and political decisions. In each phase however key indicators, such as attainment target performance, are detailed. Following the introduction, which outlines some of the issues involved in relation to the nature of the data to be collected and how it should be presented, performance is related firstly, in Section A, to practical tasks from 1990 - 92/3. (Initially these tasks covered all four attainment targets but by the final trials only two, Te2 and Te3, were being assessed via a practical task.)

In Section B performance is detailed in relation to the pencil and paper tests trialled in 1992 and 1993. The inclusion of tests resulted from the political decision that certain aspects of all subjects should be assessed via this method. This created a unique situation for design and technology. Using criterion referenced tests to assess a practical, 'process-based' subject did, however, offer another strategy for conducting statutory testing on a national scale which was worthy of investigation. The tests were focused on two of the four attainment targets, Te1 and Te4. This provided an opportunity to compare pupil performance, in relation to these attainment targets, from three different perspectives: teacher assessment, practical activity and written testing. The main focus of this investigation took place in the 1992 national pilot. Proposals, and subsequent results, for aggregating performance are detailed both in relation to attainment target and profile component level.

This chapter is divided into the following sections:

The method of analysis

Section A - Practical tasks

The 1990 Trials

The 1991 Pilot

The 1992 Pilot

The 1992/3 Pilot

Section B - Pencil and Paper tests

The 1992 Pilot

The 1992/3 Pilot

The Methods of analysis

It is worth remembering that educational measurement is not necessarily a precise activity and measurement error is likely in the majority of situations, as Thorndike (1951) confirms:
'Whenever we measure anything.....that measurement contains a certain amount of chance error. The amount of chance error may be large or it may be small, but it is universally present to some extent. Two sets of measurements of the same features of the same individuals will never exactly duplicate each other ..... '[1]

When collecting data in relation to the questions posed by this study a number of factors had to be reconciled. The nature of the data required special consideration. The National Curriculum levels which report pupil achievement are discrete categories. Each number (from 1 to 10) represents a level of achievement. Each number is not a 'score' in the normal sense because it is not part of a continuous scale between 1 and 10. It is not possible for example to score a National Curriculum Level of 3.2 or 4.6, but only 3 or 4 or 5. Data which 'categorises' in this way is referred to as 'ordinal' data.

However, many of the traditional methods of psychometric analysis depend on scores and continuous distributions in which the interim points between whole numbers are meaningful. These kind of data are referred to as interval data, and are referred to as being of a 'higher order' than ordinal data. In a criterion referenced system where a level is achieved or it is not achieved at the formative stage, reporting interim positions between the discrete levels in any analysis requires special consideration.

The decision needed to be made therefore, was whether or not it was valid in this context to use parametric methods of analysis, many of which are based on calculations involving the mean. (Which will nearly always be some interim point between the discrete levels).

Some would argue that analysing the National Curriculum Levels as interval data using higher order statistics would not be appropriate and that a (less-extensive) range of non-parametric analyses only should be used. However, it can also be argued that treating the National Curriculum Levels as interval data opens up a wider range of parametric analyses and that the higher order statistics are robust enough to produce valid results, as long as the approach is declared.

In undertaking this research it was decided to include the use of parametric statistics in the analyses. Reporting mean National Curriculum Levels for the

purpose of analysis was a valid and enlightening procedure. Statistical purists can be referred to various sources which support the robust nature of the parametric statistics in being able validly to deal with ordinal data. For example, Hagedon and Labovitz (1971) state:

'Although some small error may accompany the treatment of ordinal variables as interval, this is offset by the use of more powerful, more sensitive, better developed, and clearly more, interpretable statistics with known sampling error.'[2]

Educationally it seemed important to adopt both parametric and non-parametric approaches. For example non-parametric Spearman rank-order correlations are used which deal with the pupil achievement levels as discrete categories in terms of rankings, and deal with tied-ranks, of which there are a large number because the ratio of cases (ie, number of students in the trial) to categories (ie, 1 to 10) is high. But the analysis specification also included reporting mean National Curriculum levels per strand, attainment target and profile component; and mean group level analyses to look at the effects of various features of the design model, as well as straightforward frequency tables and contingency tables.

Possible Analysis Variables

The data for the trials undertaken were gathered using pro forma (frequently optical mark read) for teachers to complete and, in addition to results, school context variables were also collected. The pupil results were recorded either by SoA, strand (a coherent set of statements of attainment within an attainment target which describe progression in an aspect of an attainment target), attainment target and design and technology profile component. The results can be interpreted in terms of cross-referencing informationsuch as:

Centre Type:	Comprehensive/High School; Single sex; Special.
SAT version:	Exhibition, Measurement and Public Places School charity fair - five material specific tasks locking device, warning beacon and high energy snack bar
Ability Index:	Selective/non-selective . If selective the pupils were recorded as average, above average or below average.
Gender:	Gender balance overall and by trial for school and teaching group (although the latter is only reported in the context of 1990).
Teacher assessments/estimates:	Teacher estimates for attainment targets. These present the only validity criterion possible in the absence of a programme of study. These assessments, logically, should have become more secure as the research progressed.
Pupil Attendance:	Problems such as teacher absence were also reported.

In addition, information was collected on other context variables which included:

- the specialist backgrounds of the teachers;
- any specialist facilities available;
- any additional materials used during the trial/pilot;
- timetabling details for art, business studies, CDT, home economics, information technology and design and technology (integrated);
- the teaching styles for D&T subjects (circus, team teaching, pupil swapping or other);
- any special circumstances which disrupted the trial/pilot.

Teacher Evaluation and Feedback

It was considered a priority that as much qualitative information as possible should be gathered from the teachers involved in any of the trials, a large part of which was to inform materials development on an on-going basis. Questionnaires were completed by teachers indicating their views on the pilots and trials in terms of materials, pupil engagement and interest, and on the assessment system in terms of evaluating the implementation of a statutory assessment activity in the classroom/workshop, these are reported in chapter 6.

Section A - Practical Tasks

The 1990 Trials

The Sample

The 1990 trials aimed to replicate the situation which would ensue when statutory assessment occurred for the first time. Consequently, it was mainly based on whole cohorts of pupils taking the Exhibition SAT. Two additional SATs, namely Measurement and Public Places, were also trialled, but by smaller groups.
A break-down of the trial cohort is given below.

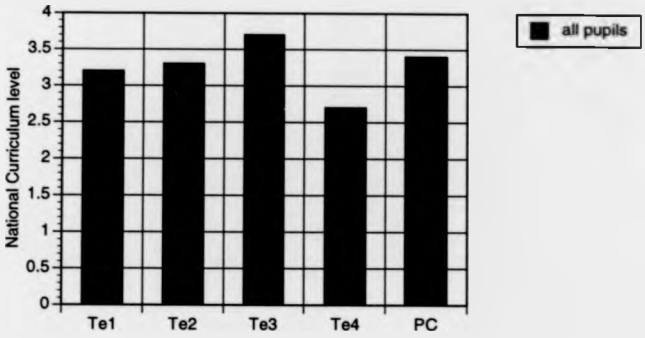
School type	No. of Schools	No. of teaching groups	No. of pupils fem.	male	No of teachers
Comp./HS	24	78	546	698	82
Selective	4	4	50	30	5
Special	4	4	18	24	5
Single Sex	3	6	67	30	6
Total	31	88	651	772	95

n.b. total is not the sum of the columns as some items appear in more than one category, for example, all selective schools were also single sex schools.

Teacher Assessment

Prior to the start of the trial, teachers were asked to make assessments in relation to all the pupils taking part in the trial. For the majority this was the first time they

had made an assessment using National Curriculum procedures. Teachers used the stranded assessment matrix to make these judgements. These assessments are shown below, by mean performance levels for both attainment targets (ATs) and profile component (PC).



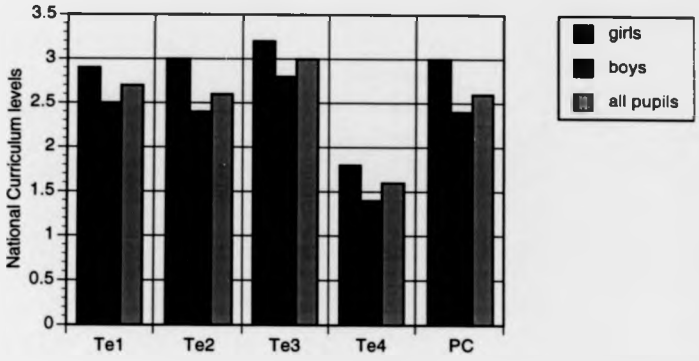
Achievement by level - teacher assessment

	% of pupils achieving levels			
	Te1	**Te2**	**Te3**	**Te4**
0	0	0	0	2
1	10	8	7	19
2	25	25	16	34
3	29	29	24	20
4	17.5	19	23	10
5	10	10	19	7
6	6	7	7	5
7	2	2	3	2
8	0.5	0	1	1
9	0	0	0	0
10	0	0	0	0

(rounded up or down except when below 0.5)

Pupil Performance - the Practical Task

The overall performance of all pupils undertaking an assessment task in trials during 1990 is shown in the diagram over the page. These aggregate statistics include both special schools, and teaching groups where delivery was deemed to be unsatisfactory. The assessments more closely match the performance which TGAT expected at the end of key stage 2, rather than key stage 3.



The general characteristics of this histogram can be described as follows:

- girls were on average, in all aspects, performing at half a level or more in advance of boys;
- levels of performance in AT1, AT2, and AT3 were remarkably similar,
- the level of performance in AT4 is at least one whole level behind that of the other ATs.
- pupil performance was enhanced when they operated in a controlled situation and with reference to a single class teacher - this was not surprising as it is the situation with which they were familiar.

The results of these early trials illustrated clearly the demands the Statutory Order placed on schools. Most teachers were unfamiliar with the programme of study and had little experience in delivering design and technology in the manner which the Order prescribed. Schools where whole cohorts of pupils trialled assessment tasks also experienced significant problems in relation to organisation and management. Similarly pupils had not been taught to the programme of study for key stage 3 nor had they had the benefit of design and technology in key stages 1 and 2. For the vast majority this trial was the first time they had faced the challenge of an integrated design and technology activity.

Teaching group assessment

As this trial took place within the context outlined above, teaching groups were classified in the following way:

1 - serious deficiencies in delivery and approach prevented pupils from revealing their true levels of performance, (this also includes serious disruptions to the programme);

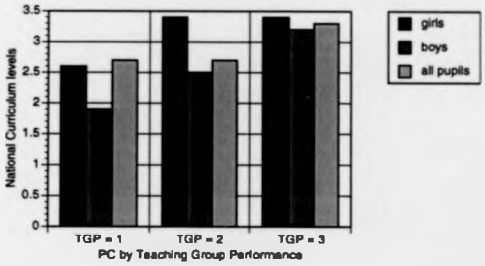
2 - satisfactory delivery and approach allowed pupils to demonstrate their capability;

3 - good delivery and approach provided the pupils with every opportunity to demonstrate their capability.

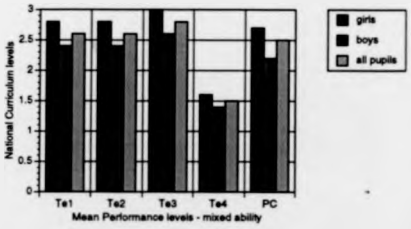
The distribution was as follows:

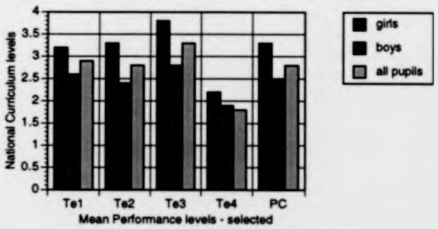
12% group 1, 29% group 2, 59% group 3

As the diagram below shows these judgements were reflected in the assessments which pupils in these groups achieved. A significant factor is the performance by gender.



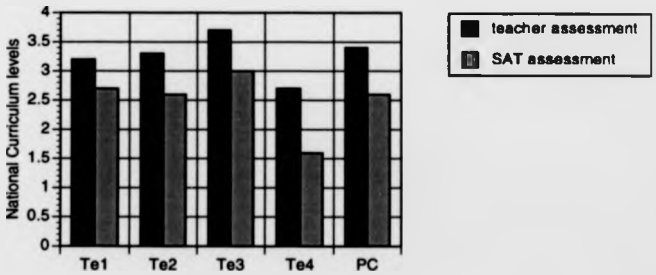
The disparity in achievement between girls and boys becomes less marked as the quality of the delivery of the assessment task improves. Girls appear less affected by delivery and environmental factors - mean PC inceases by 0.7 of a level; whilst boys performance increases by 1.4% of a level. Because this trial included a relatively high number of selective schools, pupil performance was also analysed from this perspective. Performance by mean PC shows that in a selective situation girls' achievement, throughout the activity,was enhanced by half a level but boys showed no such improvement.





Comparison of Teacher assessment with SAT assessments

Teacher assessments, by mean performance levels for ATs and PC, are shown alongside the task assessments in the diagram below.



Teacher assessments were consistently higher. It should also be noted that although the mean levels range from 3.68 for Te3 to 2.66 for Te4, teacher assessments ranged between 0 (for Te4 only) and 8 (for all ATs) and this was also the case with assessment tasks. Teacher assessments indicated that pupils were achieving their highest levels in Te3. The correlation between teacher assessment and task assessment was closest for Te 1 and 2. Te4 was acknowledged by teachers as where performance was weakest but their estimates were on average at least one level better than those achieved in the context of the assessment task. Pupils underachieved in the assessment task in both Te3 and Te4. In both cases this seemed to be due to organisational and timing issues. Sound evidence that within a time window it is difficult for pupils to do themselves justice in relation to Te3. It is also evidence that the majority of teachers perceived Te4 as the final task, instead of an activity which should be integrated into all aspects of the project, consequently the time allowed for it was minimal.

Achievement by level - Practical task

The table below shows the percentage of pupils achieving National Curriculum levels by AT and PC for all assessment tasks during the trial.

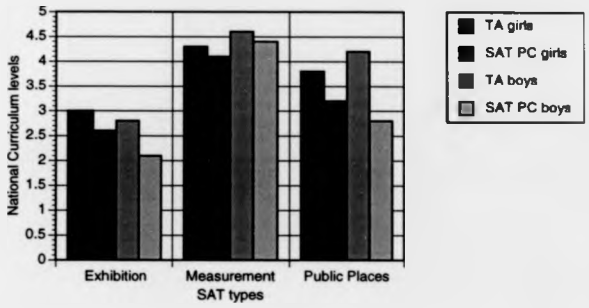
	% of pupils achieving levels				
	Te1	**Te2**	**Te3**	**Te4**	**PC**
0	2.5	1	8.5	21.5	1
1	18	18	17	34	17.5
2	29	31.5	20	22	27
3	27	27	23	13	30
4	13	11	16	5	14
5	6	6	12	2	3
6	2	3	5	2	3
7	2	2	1	0.5	1
8	0.5	0.5	0.5	0	0.5
9	0	0	0	0	0
10	0	0	0	0	0

(rounded up or down except when below 0.5)

This illustrates the degree to which the tasks differentiated and can be compared with the table which shows the percentage of pupils achieving National Curriculum levels by AT as determined by teacher assessment. As noted above the SAT scores are depressed but the pattern of differentiation is rather similar. There is a concordance between the spread of assessments for both task assessment and teacher assessment.

Assessment by type of task

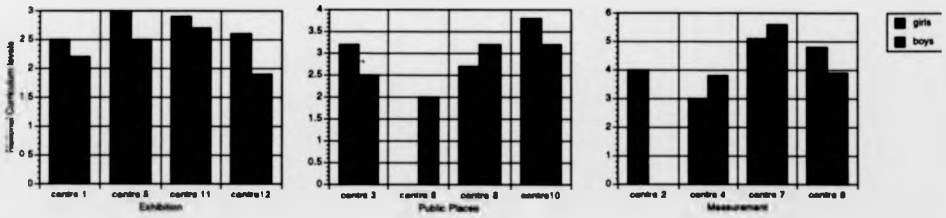
During the Summer three different SAT structures were trialled using three different themes. The histogram below shows the mean teacher estimate and the mean SAT performance by gender for each of these structures.



In examining this information the following points must be borne in mind:

- Exhibition was trialled in whole cohort schools (organisation issues, SEN pupils -30% in one school etc.),
- Measurement was trialled in groups of 40 in 4 schools. This included, by chance, two selective schools.
- One of the other schools undertook the SAT exclusively as a CDT activity.
- Public places was trialled in groups of 40 in 4 schools. This included one group, in an all boys school, identified as low achievers. Thus boys constituted 64% of those involved in the trial and the group in the all boys' school 40% of the boy population in the trial.

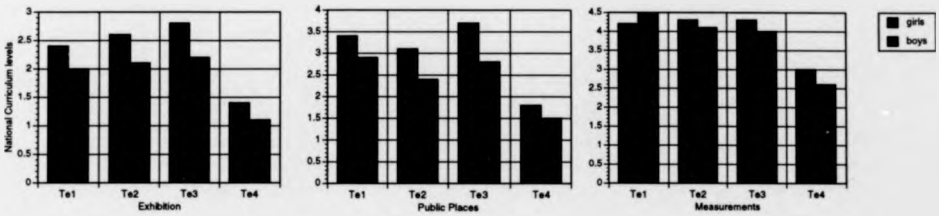
This diagram below shows mean PC performance by centre by SAT.



Measurement and Public Places both returned SAT results which for boys and girls taking Measurement, and girls taking Public Places were higher than the teacher assessments. The exception was boys taking Public Places, where the difference between TA and SAT assessment was one and half levels. This raised an important issue, did the context affect pupil performance and are some contexts gender sensitive? In this trial the all boys school, which took Public Places, contained identified low achievers. When this school was removed from the sample the mean SAT PC was also better than the teacher assessment and in two of the

other three schools boys bettered their teacher assessments. However, the issue of pupil motivation cannot be ignored.

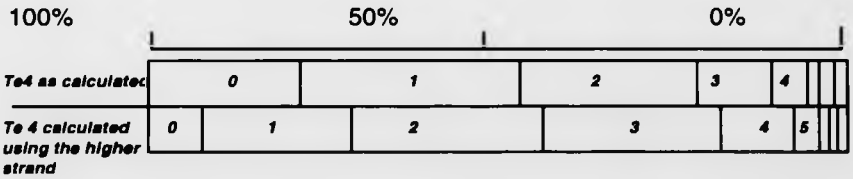
This evidence indicated that pupils performed better in a context and framework which was controlled by the SAT, as was the case with both Measurement and Public Places. However, it was possible that the focused or guided nature of these SATs constrained pupil performance in Te1. The levels achieved by attainment target for each SAT is shown below.



This indicates that performance levels are consistent between the ATs and the PC with the exception of Te4. The tighter framework of Measurement and Public Places did not, therefore, prevent pupils from producing evidence of achievement in Te1 as it appears consistent with overall performance.

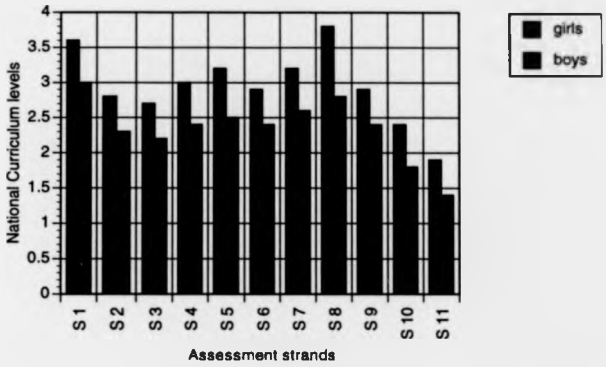
Underachievement in Te4 is consistent regardless of SAT type, with mean performances at least one level less than the PC score. This was the case even for Measurement which involved short tasks specifically targeted at Te4. Poor performance might also have been the result of the aggregation rule which was used to determine the Te4 level.

A 50% mastery rule was operated in Te4 by taking the lower of the two strands (there are three strands in the other ATs and the second highest strand score has been used). Invariably this has meant that the score for strand 11, the weakest strand, was the Te4 score. The Te4 level therefore indicated a 100% mastery. If instead of taking the lower of the two strand scores the highest was used there are significant changes in the level achieved, this is shown in table below.



Performance at strand level

The histogram below illustrates clearly that even at strand level the mean levels achieved show a consistent difference between the performance of girls and boys.



Of the three SATs trialled in the Summer, Measurement more closely characterised the kind of performance expected at the end of key stage 3. 50% of those trialling Measurement achieved PCs of 4 and 5 with 25 % achieving levels above 5. For those taking Public Places 50% achieved PCs of 2 and 3 with 30% achieving levels above 3. For Exhibition 65% achieved PCs of 2 and 3 with only 10% achieving levels above 3.

The 1991 Pilot

The Sample

All pupils in this pilot undertook a practical task. The objective of this task was that each pupil would be assessed in relation to each of the four attainment targets. The analysis was based on the 9450 pupils for whom a completed optical mark read (OMR) form was returned. However, of this total 368 were not included for the following reasons:

Teacher failed to assess pupils -	one group	- 28
Forms incorrectly completed		- 12
Pupils OMR not completed by the school		- 102
Absent for the whole activity		- 4
No evidence recorded in any AT (0000) profile		- 222
Total number of deletions		- 368 pupils

Of the total, 426 pupils were recorded as having special educational needs and 32 pupils as operating in their second language. A large number, 1938 of pupils, were assessed as having achieved no level of performance in relation to a particular

attainment target. The majority of these pupils had not been assessed as having special educational needs in relation to design and technology. It was therefore concluded that, due to teachers' inexperience in the context of National Curriculum design and technology, these pupils did not experience a genuine design and technology activity. Of this total of zero scores, 1,643 occurred in Te4 and 674 in Te3 (some pupils scored zero in both). Frequently teachers quoted lack of time as the reason for these high numbers, especially in relation to Te3. Whilst other teachers, in relation to Te4, may have failed to appreciate that evidence of achievement in this attainment target was being generated throughout the activity and not just at the end.

It might legitimately be asserted that achievement at level 1 requires no formal recording procedure and that pupils failing to achieve a level, in relation to a particular attainment target, were therefore not provided with an opportunity to discuss, describe or be observed engaging in their work. For example, one group of pupils achieved no attainment target levels higher than 1 and many scored 0. At moderation the teacher, when asked to explain this responded:
"I don't understand design and technology we've been doing Art as usual." 3
This was one example of how the context might have prevented pupils' true levels of achievement from being recognised.

Teacher assessment

Prior to the start or during the first few weeks of the assessment activity, teachers were asked to make an assessment against each attainment target for every pupil. Only one school had previously made assessments against NC criteria for Y9 pupils. To help them do this, teachers were provided with the level guide. These assessments might, in many cases, be regarded as estimates and may be based on a narrow perception of design and technology. In this subject there was, at the time of the trial, no reliable performance profile with which assessments can be compared. However the teachers undertaking these assessments were generally very experienced and would have been consistent in their judgements even if their standards varied. The mean teacher assessments for all pupils completing the pilot are shown below.

	All	***Boys***	***Girls***
Te1	3.24	3.13	3.34
Te2	3.25	3.16	3.34
Te3	3.58	3.48	3.67
Te4	2.91	2.77	3.04
PC	3.24	3.13	3.34

Achievement by level - teacher assessment

The table below details the percentage of pupils achieving each level by attainment target.

	% of pupils achieving levels			
	Te1	**Te2**	**Te3**	**Te4**
1	6	6	4	11
2	20	23	14	26
3	36	32	32	31
4	21	22	25	19
5	9	10	14	7
6	3	3	6	2
7	2	2	2	1
8	1	1	1	0.3
9	0.2	0.3	0.2	0.2
10	0.1	0.1	0.1	0

(rounded up or down except when below 0.5)

Pupil Performance - the Practical Task

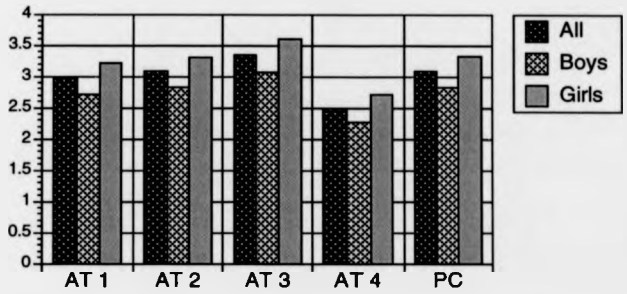
The analysis of pupil performance is based on attainment target scores which represent positive achievement, ie - only achievement in relation to attainment targets have been included. Below are the mean levels of performance for the various groups:

	mean achievement by attainment target			
	Te1	**Te2**	**Te3**	**Te4**
all pupils providing evidence in each attainment target	2.97	3.09	3.35	2.25
complete data set	2.63	2.8	2.94	2.01
pupils reported as SEN	1.56	1.62	1.79	1.10
pupils with at least one 0 in their profile	1.45	1.63	1.49	0.30

Mean pupil performance (for pupils with a complete assessment profile)

The overall mean performance of pupils completing the practical task successfully is shown in the diagram below. This does not include pupils who failed to complete the task due to absenteeism or leaving the school. It is possible that if statutory assessment takes place, regulations will enable teacher assessments to be substituted for missing information or the task would be completed at a new school. In this data, there is no substitution of teacher assessments; these performances

are based on evidence collected during the pilot practical task window. The profile component has been calculated by equally weighting the attainment targets and by rounding the mean up unless otherwise stated.



The histogram reveals the following general characteristics:

- both boys and girls achieve their highest level of performance in Te3;
- both boys and girls achieve their lowest level of performance in Te4;
- girls outperform boys consistently by approximately half a level in each attainment target;
- performance by attainment target is ranked 3 - 2 - 1 - 4 for both boys and girls.

Commentary on performance

The novelty of an assessment of this nature prevents comparisons being drawn with other data. However, the levels achieved are not consistent with the average anticipated by TGAT (level 5 to 6 at the end of key stage 3), but these norms were of course theoretical and were not substantiated. Deviations from the theoretical model may have been for the following reasons:

a. The task did not assess the curriculum delivered to the majority of pupils over the previous three years - KS3. Consequently the process was novel and pupils did not have anywhere near the full range of skills and knowledge on which to draw.

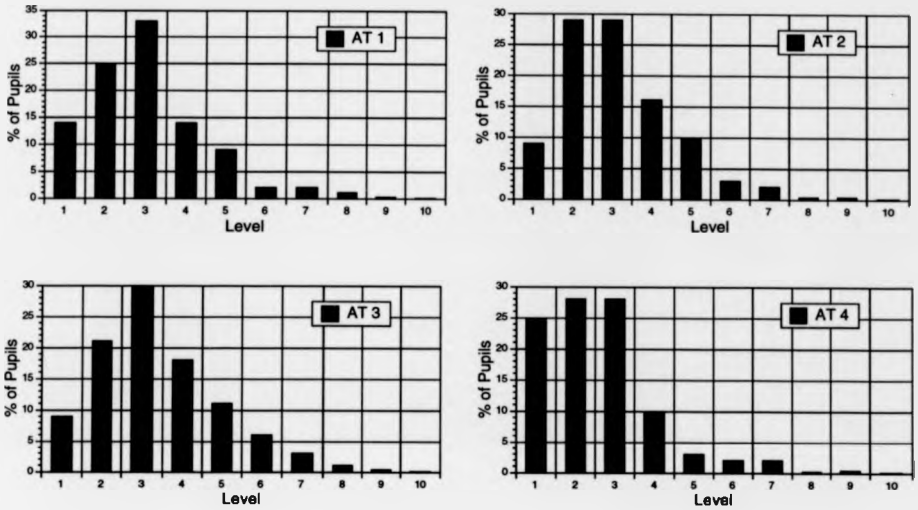
b. The philosophy of design and technology was not embedded in the teaching force. The majority of teachers were still operating within the confines of one of the subjects which constituted the federation, consequently many issues imperative to design and technology were not being addressed; the organisation and management issues arising from the bringing together of the subjects under the design and technology umbrella were creating friction.

c. Interpretation of the SoAs was creating a range of issues which was depressing achievement. Prime amongst these was teachers' inability to translate the SoAs into the context of what might reasonably be expected from a 14 year old in response to the criteria.

d. Teachers were finding it difficult to identify, recognise and record ephemeral evidence of achievement. Hence they were tending only to recognise what was

available in the form of permanent evidence in a project file. This resulted in overt pressure on pupils to produce a narrative of their activities, a description of their outcomes and an exposition of their reasoning.

Percentage of pupils achieving levels by attainment target



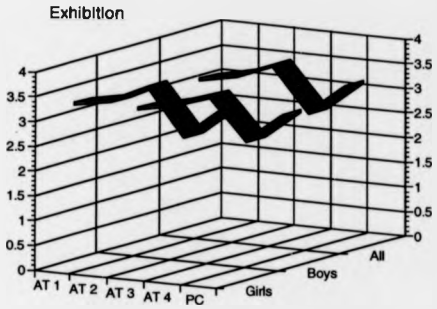
These histograms illustrate distribution by level in each of the attainment targets. Te4 was clearly the poorest performing attainment target. It was characterised by the significantly higher number of level 1 scores. There was undoubtedly a barrier which was preventing pupils achieving level 2. There was evidence that this was the reference to other times and cultures at level 2. Many teachers had not appreciated how this requirement could be addressed in a manner which was relevant to KS3 pupils. Level 4 also appeared to be an area where performance dropped significantly, possibly for the same reason.

Assessment by type of task

It is an essential requirement of statutory assessment that reliability is clearly established. A single task, it was assumed, would have a higher level of reliability. However the implications, in relation to resourcing and confidentiality, of a single task meant that other approaches required investigation. If standardisation could be demonstrated across a range and variety of tasks, a more flexible approach could have been justified. To achieve this, these three tasks were based on common elements which provided a standardising framework.

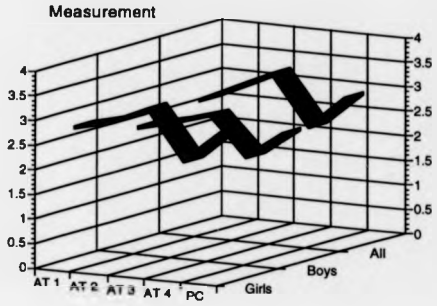
Exhibition - Broad

	All	Boys	Girls
Te1	3.05	2.81	3.27
Te2	3.2	2.97	3.41
Te3	3.49	3.21	3.76
Te4	2.61	2.39	2.83
PC	3.21	2.97	3.43
(mean up)			



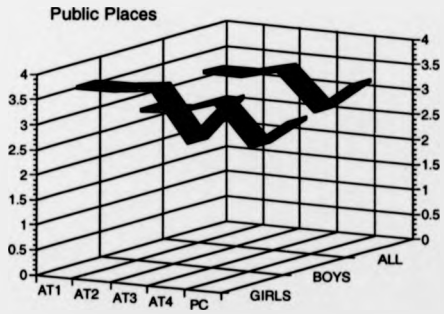
Measurement - guided

	All	Boys	Girls
Te1	2.61	2.40	2.77
Te2	2.95	2.58	2.95
Te3	3.29	2.80	3.29
Te4	2.30	2.03	2.30
PC	2.95	2.57	2.95
(mean up)			



Public Places - focused

	All	Boys	Girls
Te1	3.24	2.84	3.66
Te2	3.24	2.88	3.67
Te3	3.43	3.12	3.76
Te4	2.69	2.33	2.77
PC	3.26	2.88	3.65
(mean up)			



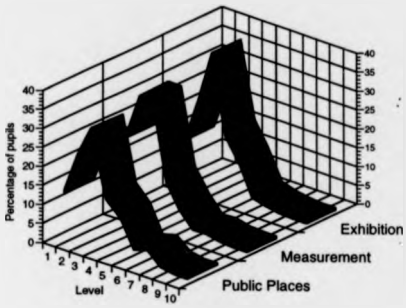
An analysis of these performance levels reveals the following:

- girls out performed boys in every attainment target in every practical task;
- Public Places produced the best mean performances;
- Public Places resulted in girls' best mean performances;
- Exhibition resulted in boys' best mean performances;
- Measurement resulted in both boys' and girls' lowest performance;
- in every task, for both girls and boys, performance by AT is ranked 3 -2 -1- 4;
- in PC terms, the theme affects girls more than boys (girls 0.7, boys 0.4 between highest and lowest mean PC);
- variation between attainment target scores is consistent by gender regardless of practical task (between highest and lowest AT score - girls 0.93, 0.99 and 0.99) (between highest and lowest AT score - boys 0.82, 0.77 and 0.79).

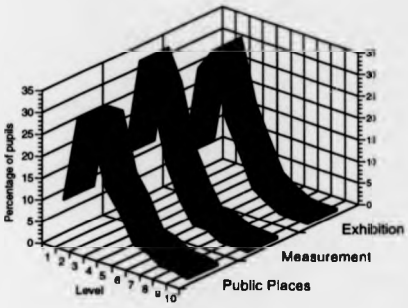
Comparison of attainment target levels by SAT type

% of pupils achieving levels

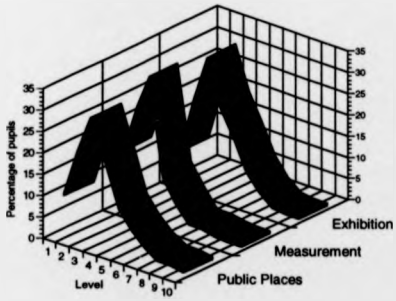
Te1	Exhibition	Measurement	Public Places
1	11	19	12
2	22	31	23
3	36	31	31
4	16	11	14
5	10	5	10
6	2	1	3
7	1	1	4
8	0.3	0.3	1
9	0.1	0.1	0.4
10	0	0	0.4



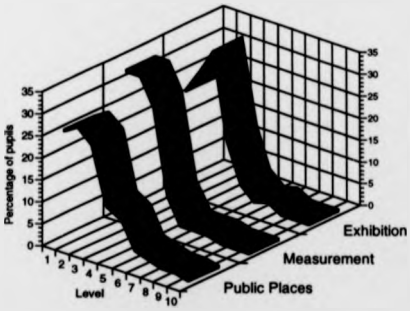
Te2	Exhibition	Measurement	Public Places
1	7	13	9
2	27	35	28
3	32	28	28
4	18	14	15
5	11	7	10
6	4	2	4
7	2	2	3
8	0.2	0.3	1
9	0.1	0.1	1
10	0	0	0.2



Te3	Exhibition	Measurement	Public Places
1	7	13	9
2	19	24	20
3	30	32	29
4	20	10	20
5	14	5	10
6	7	1	6
7	3	1	3
8	0.5	0.	1
9	0	0	1
10	0	0	0.4



Te4	Exhibition	Measurement	Public Places
1	21	32	25
2	26	32	27
3	32	26	24
4	13	7	11
5	4	1	9
6	1	1	3
7	2	1	2
8	0.1	0	1
9	0.1	0	1
10	0	0	0.3



(rounded up or down except when below 0.5)

In all attainment targets, Exhibition and Public Places produced very similar percentages of pupils at each level. Levels of achievement in Measurement were depressed by between 5% to 8% from level 5 downwards, in relation to the other tasks. This is most surprising in relation to Te4 as Measurement included a short tasked designed to produce evidence of achievement in relation to this attainment target.

Comparison between Teacher assessment and the pilot task

If the assessments produced as a result of the practical task are compared with the teacher assessments the following points emerge:

- the mean PC obtained by the girls was almost identical to the task score of 3.33;
- the mean levels obtained by girls for Te1, 2 and 3 were very similar;
- the mean levels obtained by girls for Te4 from the SAT were 0.4 of a level lower than the teacher assessment score;
- the mean PC obtained by the boys from the SAT was 0.3 of a level lower than the teacher assessment score;
- the difference in levels between attainment target scores and teacher assessment scores was approximately 0.3 of a level except in Te4 where it was half a level;
- performance by attainment target was ranked 3 - 2 - 1 - 4 for both boys and girls.

When comparing performance by level for the practical task with a similar break-down by teacher assessment, the following conclusions can be drawn:

- the percentage of pupils in each attainment target from levels 6 to 10 was identical,
  (approx. 6% of the cohort achieved level 6 and above for both task & teacher assessment);
- the percentage of pupils in Te1, Te2, and Te3 at level 5 was almost identical;
- approximately 7% fewer pupils than expected achieved level 4 in Te1, Te2, and Te3;
- expectations of achievement in Te4 were depressed from level 5 downwards;

Teacher assessments anticipated that girls would out perform boys but by a smaller margin. Girls' mean performance was in line with expected teacher assessments whilst boys under performed by 0.3 of a level. Performance at level 6 and above was as expected but fewer pupils achieved level 4 than anticipated and consequently higher percentages achieved levels 1, 2 and 3. Achievement across the attainment targets in both teacher assessment and attainment target assessment was ranked 3, 2, 1 and 4.

The 1992 Pilot - the tasks

The sample

The Pilot sample in 1992 for design and technology was drawn from a broad cross-section of schools, ranging from inner city schools to those in remote rural areas, and representing a small sample of independent and grant maintained schools, CTCs and several special schools, as well as mainstream schools. The sample of mainstream and special schools was selected at random by the SEAC's Schools' Coordinating Unit. Schools in all other categories were invited to participate by the development agency.

A total of twenty three English LEAs and six Welsh LEAs were involved in the design and technology Pilot. In England, seventy two mainstream schools and sixteen special schools completed the design and technology Pilot and in Wales twenty five mainstream schools and six special schools completed. Two independent schools, four CTCs and one grant maintained school also completed all aspects of the pilot. Of the schools which indicated a willingness to take part, 18 withdrew at some stage of the pilot; 13% of the sample of mainstream schools selected by the Schools' Coordinating Unit. Of these, 13 were mainstream schools in England, 3 English medium schools in Wales and 2 Welsh medium schools.

A total of 9,970 pupils and 800+ teachers were involved in the technology Pilot. 6,774 pupils were in English mainstream schools and 156 pupils in English special schools. 2,195 pupils in Welsh mainstream schools and 46 pupils in Welsh special schools took part in design and technology. 29% percent used Welsh medium and 71% percent used English materials. 239 pupils in independent schools, 475 pupils in CTCs, and 85 pupils in grant maintained school took part.

Teacher assessment

Mean performance by attainment target for all pupils and by gender is given below. The profile component, for all 1992 data, was calculated using the following weighted attainment target values Te 1 -15%, Te 2 - 25%, Te 3 - 40%, Te 4 - 20%. This weighting followed the laying of the assessment Order for KS3 National Curriculum assessment for the first statutory assessment in 1993.

	all pupils	boys	girls
Te 1	3.9	3.54	4.33
Te 2	3.97	3.62	4.39
Te 3	4.24	3.89	4.66
Te 4	3.74	3.41	4.12
PC	4.06	3.72	4.45

The main features of these assessments were as follows:

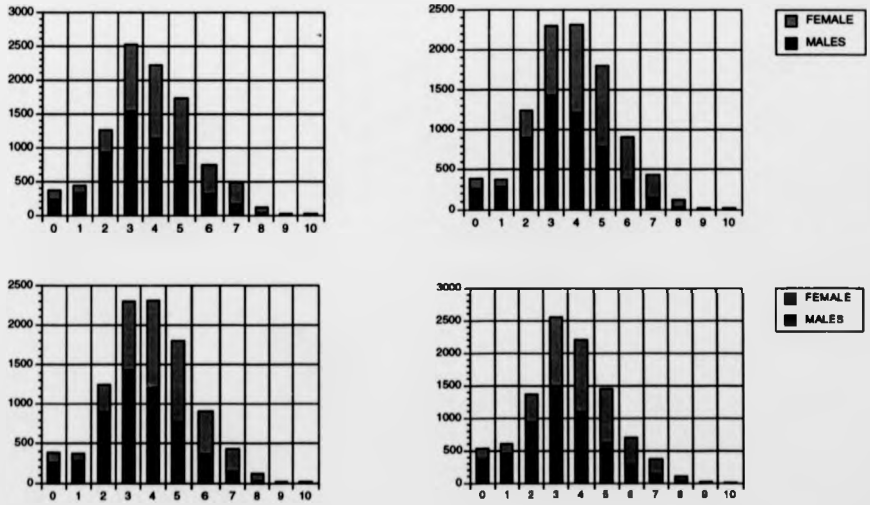
- performance by AT was ranked 3 - 2 - 1 - 4 for both boys and girls;
- girls outperformed boys in each attainment target by at least 0.6 of a level;
- the use of weighted attainment targets marginalises the assessments for Te1 and Te4 when producing a PC. The average of Te2 and Te3 produces profile components of - all-4.1, boys-3.76. girls-4.52. The effect of including Te1 and Te4 was to reduce the mean performance by an average of 0.05.

This repeated the pattern of the data collected from the major trials held in the previous two years. Teachers consistently assessed pupils at lower levels than anticipated by TGAT for pupils at the end of KS 3. This may have been because these pupils still fell outside the statutory requirements of the National Curriculum. Alternatively, it could be that the demands of the Order were not compatible with the curriculum provision for design and technology at KS 3.

It is noticeable that, even though these levels are low, over a three year period teacher assessment of performance has increased. Indeed between 1991 and 1992 the reported increase is as much as a whole level in, for example, Te4. Clearly teachers are becoming more accustomed to the demands of the Order and this is influencing the teaching of Y9 pupils. The improvement is reflected in each attainment target and for both boys and girls.

Achievement by level - Teacher assessment

The following four histograms illustrate the number of pupils achieving levels for teacher assessment by attainment target. Each bar shows the number of boys and number of girls at that level.



Pupil Performance - the practical task

The mean assessment for performance in the practical task is shown below, for all pupils and by gender.

	all pupils	*boys*	*girls*
Te 1	3.73	3.31	4.21
Te 2	3.76	3.40	4.18
Te 3	4.04	3.67	4.46
Te 4	3.50	3.14	3.88
PC	3.92	3.54	4.32

(PC calculated using weighted attainment TEs)

The main features of these assessments were as follows:

- performance by attainment target was ranked 3 -2 -1 -4 for all pupils and boys
- girls performed better on Te 1 than Te 2.
- girls outperform boys in each attainment target by at least 0.8 of a level;
- when using weighted attainment targets to arrive at a PC, the resulting score was in line with the observations relating to teacher assessment;
- the difference between practical task and teacher assessment scores was, by attainment target for all pupils, consistently about 0.25 of a level lower, with the exception of Te 4 where it was 0.5 of a level lower.

It would appear that within the more restricted opportunities which the practical task offered, pupils did - marginally - under perform in relation to teacher expectations. Clearly performance in Te 4 dropped by a more significant amount. Teachers clearly identified this as being due to two factors: lack of time and the task not offering opportunities to meet some of the statements in Te 4. Despite teachers' reservations that the task did not offer pupils opportunities to evidence Te 1 this is not reflected in the levels which pupils achieved. Despite these differences, when the PC was determined using weighted attainment targets the mean difference between teacher assessment and long task scores was only 0.14 of a level.

Analysis by practical task taken

The way in which schools allocated pupils to practical tasks varied considerably. However, the practical tasks actually taken did reveal some interesting trends.

Practical task	*% of all pupils*	*% of boys*	*% of girls*
1 - food	21	15	27
2 - constr.	37	42	34
3 - textiles	11	8	15
4 - graphics	22	23	18
5 - control	9	12	6

The aspects of the subject which are traditionally associated with boys accounted for 68% of the practical task entries for all pupils. 77% of boys took practical tasks in these aspects as did 58% of girls. The aspects of the subject which are traditionally associated with girls accounted for 32% of the long task entries for all pupils. 42% of girls took long tasks in these aspects as did 23% of boys. Clearly there has been a greater movement of girls into the typically male dominated areas of the subject than there has of boys in the opposite direction. Schools reported that they needed to use this full range of practical tasks, given the time window in which the long task has to take place. There were clear implications of this message for 1993, which in the event were ignored.

Performance by practical task taken

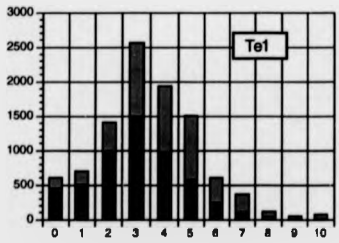
Using weighted attainment targets, the levels achieved by long task were as follows:

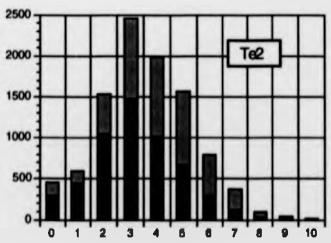
Practical task	*mean PC* **all pupils**	*mean PC* **boys**	*mean PC* **girls**
1 - food	4.1	3.5	4.4
2 - constr.	3.7	3.4	4.1
3 - textiles	4.1	3.6	4.4
4 - graphics	4.0	3.7	4.4
5 - control	3.7	3.5	4.2

Practical tasks taken in construction and control materials produced the lowest performance levels, 0.3 to 0.4 of a level below the other tasks. The pattern, in relation to gender performance, is very similar to the levels produced when all practical task assessments are combined - girls outperform boys by a minimum of 0.7 of a level.

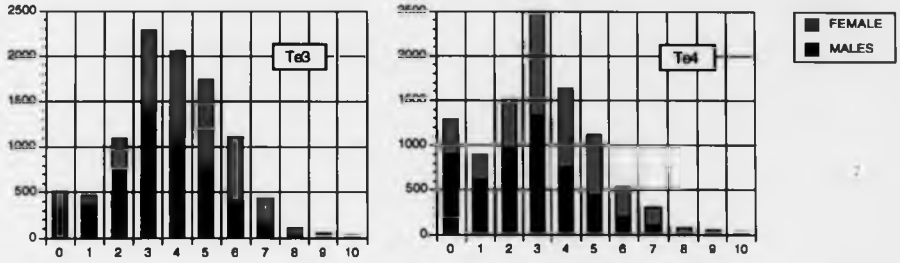
Levels achieved by pupils for the practical task

The following four histograms illustrate the number of pupils achieving levels for the long task by attainment target. Each bar shows the number of boys and number of girls at that level.





FEMALE
MALES



The 1992/3 Pilot - the practical task

The sample

The following data is based on the pupils taking part in a trial prior to the first statutory assessment of design and technology. For reasons of security, in relation to legal aspects, participating pupils were of Y10 age. The trial was conducted in four LEAs in England and Wales. For the majority of pupils, this was the first practical activity of a GCSE course in design and technology. The numbers involved were low due to the high level of security which surrounded the trial. In total, 367 pupil OMRs were completed, 74% of the number anticipated when the trial was arranged. Of this total, 328 completed the practical task and assessments were made of their work, 66% of the anticipated total. The work of thirty nine pupils was not assessed by teachers, for a variety of reasons; this represents about 10% of the pupils for whom OMRs were completed. It had been decided that tasks would only be offered in three materials in the first statutory assessment.

Task of entry

Pupils were entered for the three tasks in the following percentages:

Construction materials	51%
Control	13%
Food	36%

This entry pattern was likely to be very similar to the first statutory assessment. Construction materials certainly accounted for the largest number of entries and Control the least.

Tier of entry

Pupils were entered in the following tiers:

tier 1-4	10%
tier 3-6	55%
tier 5-8	30%
tier 7-10	5%

This entry pattern, in relation to tier, was similar to that anticipated for the first statutory assessment. At least three-quarters of pupils were entered in the middle two tiers and over 50% of the total were entered in tier 3-6. There is some evidence that a policy of tactical entering was employed by some schools, given the inability of a pupil to fall off a tier in the practical task. This resulted in a higher entry in tier 5-8. The entry level in the trial at this tier was quite high, possibly, as the triallists were Y10 pupils. Pupil performance indicated, however, that many pupils were over-entered i.e., placed in a tier above their ability level.

The entry by task and tier is shown below. The percentage is given in relation to pupils taking that task and in relation to all pupils.

	Construction		**Control**		**Food**	
	within task	***all pupils***	***within task***	***all pupils***	***within task***	***all pupils***
tier 1-4	11%	6%	2%	0.3%	11%	4%
tier 3-6	61%	31%	39%	4%	51%	19%
tier 5-8	27%	14%	46%	6%	30%	11%
tier 7-10	1%	0.5%	13%	2%	8%	2%

The entry pattern for Construction and Food was similar and accords with what was expected, with the exception of the relatively high entry in the top tier in relation to Food. The entry profile for Control, though, was different. Almost 50% of pupils were entered in tier 5-8 and 60% were entered for the top tier. There was some evidence that teachers perceived the Control task as more appropriate for able pupils and this, in turn, has affected the entry profile. Observation evidenced that pupils taking Control did so in smaller group sizes. This was often at the expense of those pupils taking tasks in the other two materials.

Time taken on the task

This was the first time that the task had time restrictions; teachers were asked to apply the 'up to 12 hours' ruling. The average time, across all tasks was 10.7 hours. There was no significant evidence that pupils completed any of the tasks in substantially less time, or that pupils failed to complete within this time. The average completion times are given below.

Construction materials	10.94 hours
Control	10.22 hours
Food	10.69 hours

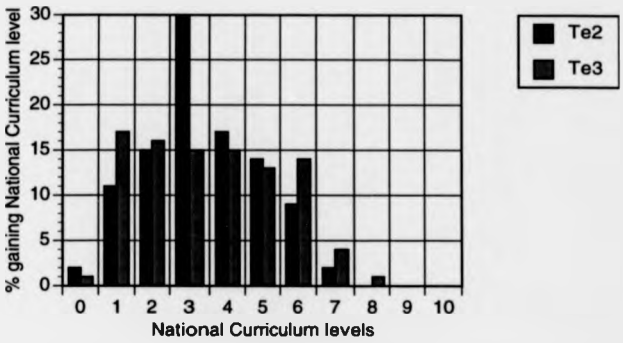
If pupils were over-entered, it was quite likely that they would fail to complete the

task and would produce insufficient evidence on which an assessment could be made.

Pupil Performance - the practical task

Construction materials

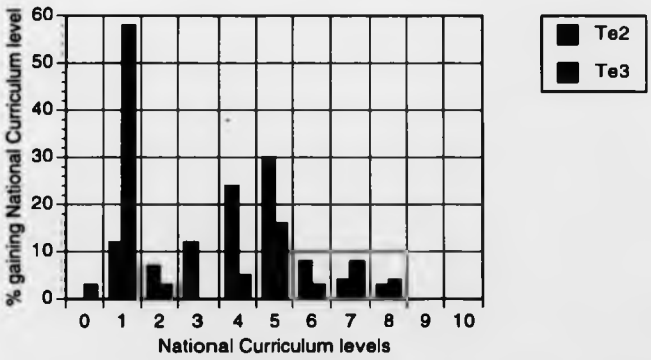
The histogram below shows pupil performance in relation to Te 2 and Te 3.



The mean level achieved for Te 2 was 3.4 and for Te 3 was 3.6; a slightly better performance in relation to Te 3 than for Te 2. Performance in both attainment targets indicates poor performance for Y10 pupils.

Control

The histogram below shows pupil performance in relation to Te 2 and Te 3.

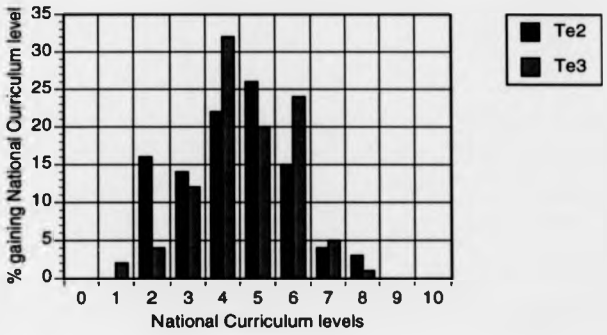


The mean level achieved for Te 2 was 4 and for Te 3 was 2.7. The graphs indicated that performance in relation to Te 2 was far better than for Te 3 but both indicate poor performance for Y10 pupils. The performance in relation to Te 3 is clear evidence of pupils being over-entered. Teachers also focused pupils' attention on constructing the circuit - the level 5-8 design requirement. This they may have achieved, but as a result they failed to satisfy the demands of the design

requirements at tiers 1-4 and 3-6 (a high percentage achieving level 1). As a result, many pupils could not be awarded high levels even though they may have achieved some of the requirements of these levels. It also indicated that teachers had not paid sufficient attention to the demands of the assessment requirements until after the task was completed.

Food

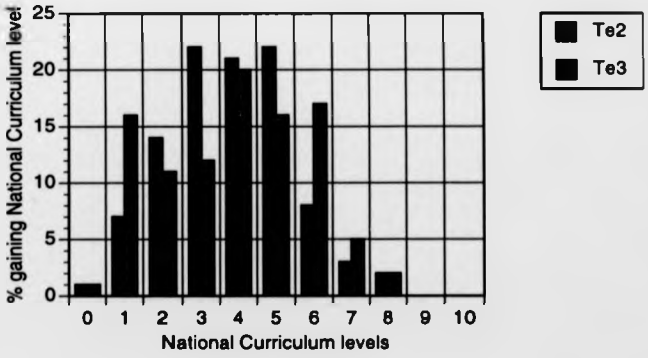
The histogram below shows pupil performance in relation to Te 2 and Te 3.



The mean level achieved for Te 2 was 3.7 and for Te 3 was 4.6. The histogram indicated that performance in relation to Te 3 was much better than for Te 2 but both indicate poor performance for Y10 pupils. There is evidence that pupils had difficulties with the mathematical aspects of the design brief, i.e. establishing the kcal. value of the snack bar, even though this represented mathematical achievement at a similar or lower level. (Clear evidence that the subject boundaries into which the curriculum divides knowledge are not easily broken down!)

Performance across the tasks

The histogram below shows pupil performance in relation to Te 2 and Te 3 for all three tasks.



The mean level achieved for Te 2 was 3.7 as was the mean level for Te 3. As noted in relation to each task, these performance levels are poor. Pupils have undoubtedly found the requirements of working to a detailed specification extremely demanding. Tasks of this nature revealed how little pupils really know about technical aspects of the subject on which these activities were focused.

Section B

The 1992 Pilot - the tests

This was the first year in which pencil and paper tests were taken as means of determining pupil performance in relation to two attainment targets - Te1 and Te 4.

Entry levels

The number of pupils entered for each test band is stated below:

	total	*as %*	*% boys*	*% girls*
test 1 *(levels 1 to 4)*	2271	22.7	16	7
test 2 *(levels 3 to 7)*	6775	68	35	33
test 3 *(levels 6 to 10)*	915	9.3	4	6

These entry statistics revealed two interesting aspects:

- at least twice the number of boys are entered for the lower test band compared to girls;
- 50% more girls are entered for the higher test band than boys.

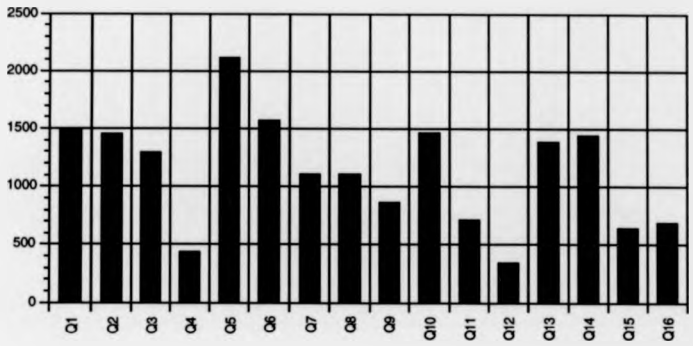
Clearly, on the basis of teacher expectations, boys were twice as likely as girls to be underachieving and girls were twice as likely as boys to be high achievers. The majority of the cohort, 68%, was entered within the key stage range. The large number of pupils being entered for test 1 was a cause for concern. In effect this means, that at least one in five pupils was entered for a test on the basis that their teachers did not expect them to achieve beyond level 3.

Criterion referenced tests created genuine problems in relation to the process used to determine overall performance. If a pupil's answer to a question was deemed to have satisfied the assessment criteria, a statement of attainment, then how were all these positive assessments to be aggregated? Are some, possibly those at a

higher level, seen to be of greater importance or might failure at a lower level result in higher levels being discounted? There were many issues of this nature which had to be addressed. Because of these factors performance is firstly presented in relation to pupils satisfying assessment criteria by question. It should be noted that each of these tests had three inclines of difficulty.

Test 1 - levels 1-4

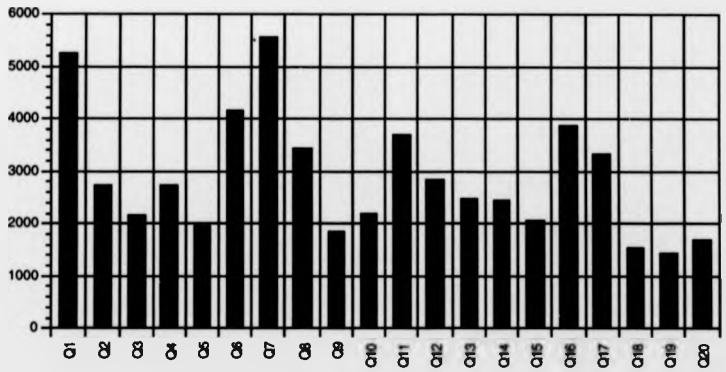
The histogram below shows the number of pupils answering each question on this test paper correctly.



The question which most pupils correctly answered was a level 1 question - statement Te 1.1a - 92% answered correctly. The question found most difficult by pupils was level 4 question 12 - statement Te 1.4b - 15 % answered correctly.

Test 2 - levels 3-7

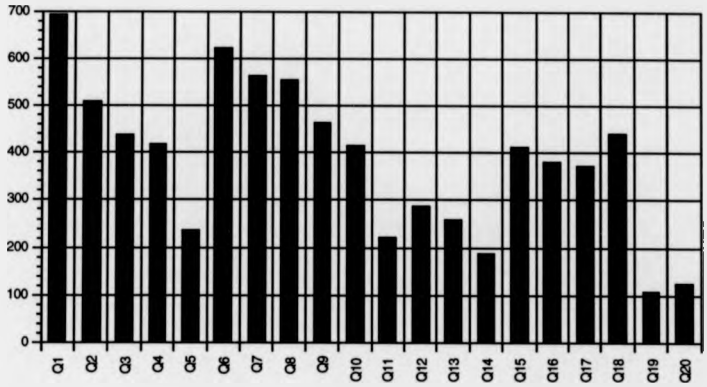
The histogram below shows the number of pupils answering each question on this test paper correctly.



The question which most pupils correctly answered was a level 3 question - statement Te 1.3a - 82% answered correctly. The question found most difficult by pupils was at level 6, question 19 - statement Te 4.6b - 21 % answered correctly.

Test 3 - levels 6 -10

The histogram below shows the number of pupils answering each question on this test paper correctly.



The question which most pupils correctly answered was a level 6 question - statement Te 4.6c - 82% answered correctly. The question found most difficult by pupils was at level 9, question 19 - statement Te 4.8b - 12 % answered correctly.

Correct response

The percentage of correct responses by test paper was as follows

test 1 - 49.93%
test 2 - 42.55%
test 3 - 42.2%

A further analysis was undertaken to examine the correct responses by level, regardless of test taken. To achieve this the number of correct responses at a level was calculated as a percentage of all answers provided at that level for pupils taking that test. The following tables shows these percentages by level for each test.

				National Curriculum level						
1	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	
73	56	47	24							**test 1**
		70	42	36	34	30				**test 2**
					62	52	41	29	26	**test 3**

percentage of correct answers at a level

This demonstrates that within each test there was a clear incline of difficulty. It also indicates in the overlap areas, that pupils entered for the higher test paper produced a higher percentage of correct answers to these questions. If the percentage of questions correctly answered at a level is calculated regardless of

test entered the following picture emerges.

National Curriculum level										
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	
73	56	64	46	36	38	32	41	29	26	**all tests**

percentage of correct answers at a level

This further analysis only effected level 3 to 7, the levels in which overlap occurred but with the exception of level 5. This table is dominated by the much larger cohort of pupils which took test 2, However, there is a decline from level 1 to level 10. It was unclear, however, if this was desirable and should form a model for future development.

'Floor and Ceiling' effects

'Floor and ceiling' effects are those which may result in pupils either failing to record a level or under-performing due to incorrect entry. Examining these pilot results in relation to pupils failing to achieve a level was difficult because it was dependant on the aggregation rules adopted. The table below shows the number/ percentage of pupils who failed to get a question correct, by test paper, by section.

	section A		**section B**		**section C**	
	number	***%***	***number***	***%***	***number***	***%***
test 1	294	13	49	2	376	16
test 2	654	9	219	3	1374	20
test 3	90	9	40	4	252	27

A pupil failing to get any question correct in Sections A and C would have failed to achieve a level for Te4 and failure in Section B would result in no recorded level for Te1. As in all other assessment procedures (including practical tasks) Te4 produced the poorest performance and consequently the highest percentage of failures. This was most pronounced in Section C. Pupil fatigue - this was the only one and a half hour test pupils experienced as part of the end of key stage testing procedure - may well have been a major factor and in test 3 pupils' inability to complete the test in the time available was another. The numbers failing to demonstrate any positive achievement in Te 1 was 3%. These figures also included pupils who might genuinely have been in the W (working towards) category having not yet reached level 1. The number of pupils failing to get any question correct was as follows:

	all pupils	***boys***	***girls***
test 1	23	19	4
test 2	79	44	35
test 3	13	6	7

This represents 1.18% of pupils taking a test. These pupils would not have recorded a level for either attainment target. If pupils were entered for a test which created a ceiling to their possible level of achievement, they should have answered all the questions correctly at the highest level of the test for which they were entered, i.e. all level 4 questions in test 1, all level 7 questions in test 2, given the aggregation rules being applied. The number of pupils in this category is shown below - these pupils represented 4% of those taking part in the pilot.

test 1	***all pupils***	***boys***	***girls***
achieved all level 4 questions	66	48	18
test 2			
achieved all level 7 questions	361	144	215
test 3			
achieved all level 10 questions	27	4	23

It is extremely unlikely that 27 pupils did get all level 10 questions correct. It was not possible however, due to the numbers involved, to remark pupils' papers.

Aggregation rules

Based on previous smaller scale trials and evaluations six aggregation rules were applied to the test results. Using these rules the following mean levels resulted for the two attainment targets assessed via the test.

Rule 1 - the two highest statements achieved in an attainment target are aggregated, divided by 2 and rounded down;

	all pupils	boys	girls
Te 1	4.78	4.42	5.2
Te 4	4.47	4.15	4.85

Rule 2 - the two highest statements achieved in an attainment target are aggregated, divided by 2 and rounded up;

	all pupils	boys	girls
Te 1	5.27	4.92	5.69
Te 4	5.02	4.7	5.39

Rule 3 - the highest level at which both statements are satisfied;

	all pupils	boys	girls
Te 1	4.49	4.1	4.93
Te 4	4.33	3.97	4.72

Rule 4 - the highest level at which two statements are satisfied, i.e. the second highest statement (roll back rule);

	all pupils	boys	girls
Te1	4.77	4.41	5.19
Te4	4.58	4.24	4.97

Rule 5 - the highest level at which all statements are satisfied;

	all pupils	boys	girls
Te1	3.24	2.9	3.59
Te4	3.38	3.14	3.63

Rule 6 - the highest level at which any statement is satisfied;

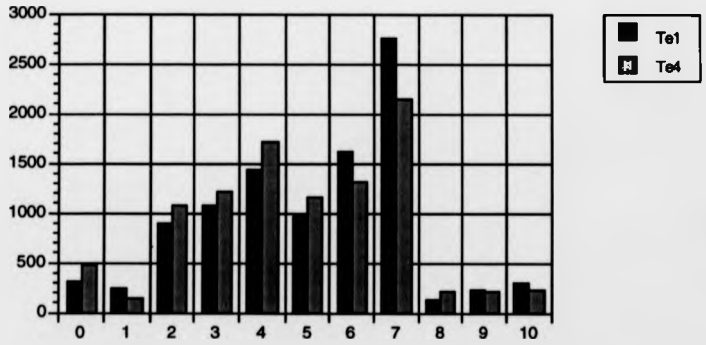
	all pupils	boys	girls
Te 1	5.59	5.2	6.02
Te 4	5.39	5.08	5.74

Regardless of the rule girls, produced higher test scores for both attainment targets. The difference in achievement between girls and boys is obviously dependent on the rule applied. It is most exaggerated by the easiest rule - 6, and least obvious with the most difficult rule - 5. As in the practical task Te1 produced higher scores than Te4.

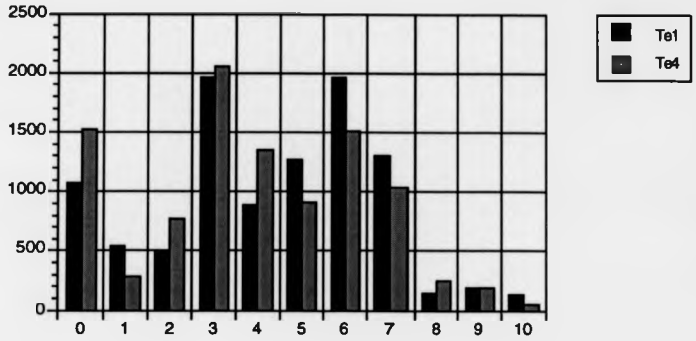
A number of these rules were not worth pursuing as they were either too demanding or too lenient. An analysis of the evidence suggests that rule 2 and rule 4 were the most sensible and, therefore worthy of further analysis. Rule 2 did produce higher assessments, however this rule rounded down (rule 1) produced scores so similar to rule 4, which is more straightforward to apply, that it was not worth pursuing.

Numbers achieving levels in the test

The histograms below illustrate the numbers of pupil achieving levels in the test using the two selected aggregation rules.



Rule 2



Rule 4

Combining test and practical task assessments

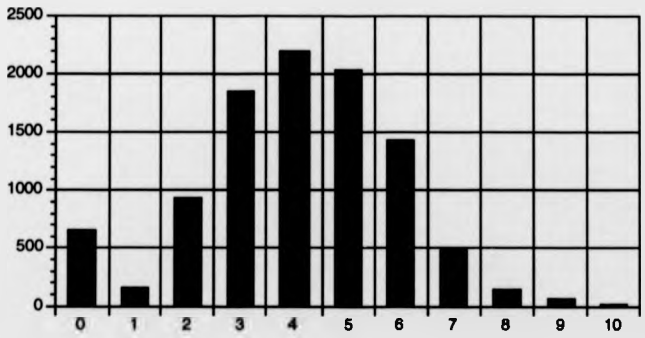
If rule 2 and rule 4 are used for test results and combined with the levels obtained by pupils for Te 2 and Te 3 from the practical task the levels for the design and technology PC could be examined. The four ATs were aggregated using the weighting published in the draft assessment order - Te 1 - 15%, Te 2- 25% , Te 3 - 40%, and Te 4 - 20%..

The levels obtained for the practical task and teacher assessment are provided to allow for easy comparison.

Rule 2

		practical task	teacher assess.
Mean PC - all pupils in the pilot	- **4.34**	3.92	4.06
Mean PC - boys in the pilot	- **3.99**	3.54	3.72
Mean PC - girls in the pilot	- **4.74**	4.32	4.45

The test scores increased mean performance levels by almost half a level - 0.42, in relation to the practical task. This aggregation procedure also resulted in higher mean scores than for teacher assessment. The histogram below shows the number of pupils at each level using this procedure.



If the results are examined by practical task the following picture emerges.

	practical task levels			task and test combined subject score		
	all pupils	*boys*	*girls*	*all pupils*	*boys*	*girls*
task 1	4.1	3.5	4.4	4.42	3.91	4.77
task 2	3.7	3.4	4.1	4.19	3.87	4.65
task 3	4.1	3.6	4.4	4.47	3.96	4.78
task 4	4.0	3.7	4.4	4.47	4.18	4.87
task 5	3.7	3.5	4.2	4.2	4.13	4.57

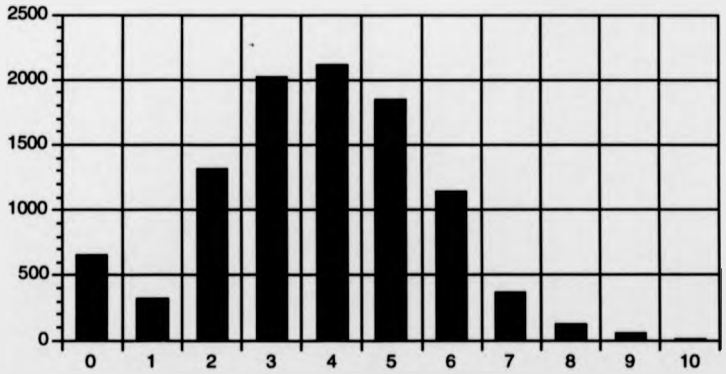
(n.b. Not all schools reported which practical task pupils took.)

This shows that, using this aggregation procedure, the test was fair, regardless of the long task the pupil took. The test enhanced performance for both boys and girls for all tasks - when schools reported the task taken.

Rule 4

		long task	teacher assess.
Mean PC - all pupils in the pilot	- **4.04**	3.92	4.06
Mean PC - boys in the pilot	- **3.69**	3.54	3.72
Mean PC - girls in the pilot	- **4.45**	4.32	4.45

The test scores marginally increased performance levels by 0.12 of a level, in relation to the practical task. This aggregation procedure results in very similar mean scores for teacher assessment. The graph below shows the number of pupils at each level using this procedure.



If the results are examined by practical task the picture over the page emerges.

	long task levels			task and test combined subject score		
	all pupils	boys	girls	all pupils	boys	girls
task 1	4.1	3.5	4.4	4.12	3.61	4.47
task 2	3.7	3.4	4.1	3.90	3.57	4.37
task 3	4.1	3.6	4.4	4.15	3.64	4.48
task 4	4.0	3.7	4.4	4.17	3.87	4.59
task 5	3.7	3.5	4.2	3.98	3.82	4.28

(n.b. Not all schools reported which long task pupils took)

This shows that, using this aggregation rule, the test was fair, regardless of the long task the pupil took. The test improved all mean levels of performance, for both boys and girls, for all cases where the long task taken was reported.

The 1992/3 Pilot - the tests

Entry levels

The total number of pupils who took the test and whose papers were marked was 382, 77% of the anticipated entry. The entry by tier and task is shown below.

	Construction		Control		Food	
	within task	*all pupils*	*within task*	*all pupils*	*within task*	*all pupils*
tier 1-4	15%	2%	1%	0.2%	11%	4%
tier 3-6	58%	29%	46%	6%	51%	19%
tier 5-8	26%	13%	39%	4%	30%	11%
tier 7-10	1%	0.5%	13%	2%	8%	2%

There was some evidence that teachers entered pupils for a lower level tier for the test than for the task. In total this affected 14 pupils, approximately 4% of the total. The most noticeable effect was the reversal of entries in Control in relation to tier 3-6 and tier 5-8. Although the numbers entered in the top two tiers in Control was higher than for the other two materials, the entry profile was similar. It should be noted that the entry to the test was principally based on performance in relation to Te 2 and Te 3 in a specific material context.

The following tables show the number of pupils achieving credits by level for each tier of entry. For example the first table shows tier 1-4, Construction materials. If the level 2 column is inspected, it can be seen that if the mastery level was set at 6, 14 pupils would have gained this level (9 at 6 credits, 3 at 7 credits and 2 at 8 credits). Two of these pupils would also have gained level 3 (2 at 7 credits) and 1 level 4 (1 at 6 credits). Each level column adds up to the total taking the test.

Te4 - Construction materials

tier 1-4

no. of credits	level 1	level 2	level 3	level 4
0	2	2	2	6
1	0	1	4	10
2	1	1	4	7
3	5	2	9	1
4	3	5	4	3
5	6	4	4	1
6	4	9	0	1
7	6	3	2	0
8	2	2	0	0

tier 3-6

no. of credits	level 3	level 4	level 5	level 6
0	3	8	7	28
1	5	7	27	26
2	14	17	29	27
3	16	21	18	16
4	19	20	23	10
5	24	18	8	8
6	16	11	5	2
7	13	12	1	1
8	8	4	0	0

tier 5-8

no. of credits	level 5	level 6	level 7	level 8
0	2	4	4	13
1	3	6	3	4
2	5	10	9	13
3	15	11	6	7
4	7	4	9	11
5	12	11	0	4
6	5	5	9	1
7	4	1	5	0
8	0	1	8	0

tier 7-10

no. of credits	level 7	level 8	level 9	level 10
0	0	0	0	1
1	0	0	0	0
2	0	1	0	1
3	1	0	2	0
4	0	1	0	0
5	0	0	0	0
6	1	0	0	0
7	0	0	0	0
8	0	0	0	0

Te4 - Control

tier 1-4

no. of credits	level 1	level 2	level 3	level 4
0	0	0	1	0
1	1	0	0	0
2	1	0	0	1
3	0	0	0	1
4	0	1	1	0
5	0	1	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0

tier 3-6

no. of credits	level 3	level 4	level 5	level 6
0	4	4	8	6
1	0	1	2	0
2	1	2	3	3
3	3	1	3	2
4	4	1	2	3
5	3	4	0	3
6	3	1	2	1
7	1	3	0	2
8	1	3	0	0

tier 5-8

no. of credits	level 5	level 6	level 7	level 8
0	1	0	4	3
1	1	0	1	1
2	3	0	2	6
3	3	1	2	2
4	2	2	5	1
5	0	2	2	2
6	4	4	0	3
7	3	4	0	0
8	1	5	2	0

tier 7-10

no. of credits	level 7	level 8	level 9	level 10
0	2	2	1	1
1	0	0	0	0
2	1	2	3	0
3	1	2	0	1
4	2	0	1	0
5	0	0	1	2
6	0	0	0	1
7	0	0	0	0
8	0	0	0	1

Te4 - Food

tier 1-4

no. of credits	level 1	level 2	level 3	level 4
0	0	0	0	1
1	0	0	1	3
2	2	2	1	4
3	0	0	4	1
4	3	5	4	3
5	2	3	3	3
6	6	5	0	2
7	3	2	1	1
8	2	0	1	0

tier 3-6

no. of credits	level 3	level 4	level 5	level 6
0	0	0	1	2
1	0	1	6	8
2	3	9	10	7
3	8	9	7	20
4	9	10	8	13
5	15	11	9	9
6	5	8	7	4
7	19	11	7	0
8	9	9	13	5

tier 5-8

no. of credits	level 5	level 6	level 7	level 8
0	0	1	0	15
1	0	1	9	12
2	5	5	8	3
3	8	9	5	2
4	8	9	4	7
5	2	9	6	0
6	7	6	7	1
7	6	0	0	0
8	4	0	1	0

tier 7-10

no. of credits	level 7	level 8	level 9	level 10
0	0	1	1	2
1	0	2	1	0
2	2	3	1	0
3	1	2	2	4
4	0	1	0	2
5	3	0	2	2
6	1	1	2	0
7	3	0	1	0
8	0	0	0	0

If these scores are aggregated across Te 4, the following scoring pattern emerges.

All materials

tier 1-4

no. of credits	level 1	level 2	level 3	level 4
0	2	2	3	7
1	1	1	5	13
2	4	3	5	12
3	5	2	13	3
4	5	9	8	6
5	7	9	9	2
6	10	14	0	3
7	9	5	3	1
8	4	2	1	0

tier 3-6

no. of credits	level 3	level 4	level 5	level 6
0	7	12	16	36
1	5	9	35	34
2	18	28	42	37
3	27	31	28	38
4	32	31	33	26
5	42	33	17	20
6	24	20	14	7
7	33	26	8	3
8	18	16	13	5

tier 5-8

no. of credits	level 5	level 6	level 7	level 8
0	3	5	8	31
1	4	7	13	17
2	13	15	19	22
3	26	21	13	11
4	17	15	18	19
5	14	22	8	6
6	16	15	16	5
7	13	5	5	0
8	5	6	11	0

tier 7-10

no. of credits	level 7	level 8	level 9	level 10
0	2	3	2	4
1	0	2	1	0
2	3	6	4	1
3	3	4	4	5
4	2	2	1	2
5	3	0	3	4
6	2	1	2	1
7	3	0	1	0
8	0	0	0	1

The key issue was at what score should the mastery level be set. The following table shows the number of pupils who would have fallen off the tier if the mastery level was set at each of the possible scores. This analysis was done by tier.

numbers failing to gain this number of credits at each level									
	8	*7*	*6*	*5*	*4*	*3*	*2*	*1*	*0*
level 1	43	34	24	17	12	7	3	2	0
level 2	45	40	26	17	8	6	3	2	0
level 3	46	43	43	34	26	13	8	3	0
level 4	47	46	43	41	35	32	20	7	0
level 3	190	157	133	91	59	32	14	9	0
level 4	190	164	144	111	80	49	21	12	0
level 5	193	185	171	154	121	93	51	16	0
level 6	201	198	191	171	145	107	70	36	0
level 5	106	93	77	63	46	20	7	3	0
level 6	105	100	85	63	48	27	12	5	0
level 7	100	95	79	71	53	40	21	6	0
level 8	111	111	106	100	81	70	48	31	0
level 7	18	15	13	10	7	4	3	3	0
level 8	18	18	17	17	15	11	5	3	0
level 9	18	17	15	12	11	7	3	2	0
level 10	17	17	16	12	10	5	4	4	0

The important row is the first in each tier, as the majority of pupils who failed to achieve this level would probably also have failed at higher levels. If the mastery level was 6 credits, 75%, it can be seen that very large numbers of pupils would not achieve a level - 64% in tier 2 for example. Even at 5 credits, those not achieving a level is 44%. Performance is weakest on the Construction material test and strongest in Food, with Control closer in performance scores to Food.

Applying a mastery rule

The previous information includes each pupil's performance, at each level in the tier of the test paper which they took. In determining the level achieved only each pupil's best performance should be considered. A minimum achievement of 50%, 4 credits, has been considered up to a 100%, 8 credits. Less than 50% success would clearly be unacceptable as a mastery level. This analysis has been carried out irrespective of the tier which a pupil took. Therefore percentages in the first row, 0 level, are those pupils who failed to achieve a level on all of the four tests at that mastery level - see over the page.

Construction materials - Te 4

% achieving highest level at this number of credits

level	4	5	6	7	8
0	18	33	53	72	89
1	1	1	1	2	1
2	7	6	5	2	1
3	10	12	11	7	2
4	17	17	11	7	2
5	17	8	4	1	0
6	12	9	2	1	0
7	9	10	11	6	4
8	8	2	0	0	0
9	0	0	0	0	0
10	0	0	0	0	0

Control - Te 4

% achieving highest level at this number of credits

level	4	5	6	7	8
0	13	22	39	61	80
1	0	0	0	0	0
2	2	2	0	0	0
3	7	4	2	0	0
4	11	15	11	11	7
5	0	2	7	2	0
6	35	33	28	20	7
7	11	2	2	4	4
8	13	11	7	0	0
9	0	0	0	0	0
10	9	9	4	2	2

Food materials - Te 4

% achieving highest level at *this number of credits*

level	4	5	6	7	8
0	8	18	37	57	78
1	1	1	4	1	1
2	2	2	3	1	0
3	7	10	5	9	4
4	13	10	10	6	3
5	19	21	22	18	9
6	31	20	10	4	4
7	8	11	7	3	1
8	7	1	1	0	0
9	2	2	2	1	0
10	3	1	0	0	0

All materials - Te 4

% achieving highest level at *this number of credits*

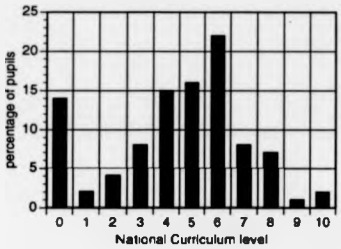
level	4	5	6	7	8
0	14	26	46	65	84
1	1	1	2	2	1
2	5	4	4	2	1
3	9	10	8	7	3
4	15	14	11	7	3
5	15	12	10	7	3
6	21	16	8	4	2
7	9	10	8	5	3
8	8	3	2	0	0
9	1	1	1	0	0
10	2	2	1	0	0

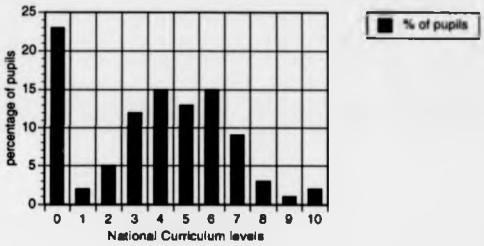
Performance patterns seem reasonable with the exception of levels 5 and 6 in the Control test. Level 6 appears to have been more easily achieved than level 5 which might, in fact, have been the case. Level 5 asked pupils to report on their own task, whilst level 6 asked them to evaluate some theoretical solutions. There was other evidence which suggested that pupils were better at responding to new material, rather than recalling their own work, which might not have been very successful.

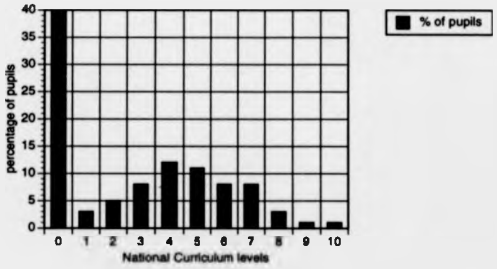
Mastery at credits of 7 and 8 is clearly out of the question as representing a far too difficult challenge. The original credit mastery model on which the tests were designed proposed a mastery level of 6. If this was applied 46% of pupils would not achieve a level for this attainment target; this would be unacceptable. At a mastery level of 5, 26% would not achieve a level. Although this also appears high, a lower mastery level would not justify the award of a level. It must be remembered that the level awarded will indicate competence in relation to the statements of attainment

at that level. The levels achieved at a mastery of 4, 5, and 6 credits are shown graphically below.

Final recommendations for the first statutory assessment could only be reached after reviewing performance in relation to Te 1. It would have been possible, however, to have had different mastery levels for each attainment target, although this would have been difficult to justify. Pupils needed to achieve a level on both attainment targets, consequently the number falling off a tier could only increase when both attainment targets were taken into account. In the trial, it was possible that pupils were not well prepared, were not correctly entered and that teachers applied the mark scheme too rigorously; all of these factors may have depressed achievement levels. These issues, however, were likely to occur in the first statutory assessment. If all these factors are taken into account a mastery level of 5 would seem appropriate in relation to Te 4.







Te 1 - Identifying needs and opportunities

All pupils did the same test in relation to Te 1 regardless of the practical task which they had taken. The number of credits gained by pupils at each level is shown over the page.

tier 1-4

no. of credits	level 1	level 2	level 3	level 4
0	0	0	0	1
1	0	0	0	4
2	0	1	5	7
3	0	1	12	4
4	0	1	6	6
5	0	2	9	2
6	0	8	5	14
7	0	11	5	8
8	47	23	1	1

tier 3-6

no. of credits	level 3	level 4	level 5	level 6
0	6	7	9	16
1	3	4	19	41
2	3	10	9	41
3	15	15	30	38
4	18	19	50	24
5	38	31	33	25
6	43	53	21	10
7	43	46	8	5
8	32	21	7	6

tier 5-8

no. of credits	level 5	level 6	level 7	level 8
0	2	4	12	46
1	4	10	8	15
2	9	16	14	23
3	12	15	18	14
4	12	22	15	9
5	20	13	16	7
6	24	16	15	3
7	15	8	9	2
8	13	7	4	2

tier 7-10

no. of credits	level 7	level 8	level 9	level 10
0	1	3	1	1
1	3	3	0	1
2	1	5	4	4
3	4	3	1	3
4	6	2	6	1
5	1	0	2	1
6	1	0	2	1
7	0	0	0	2
8	0	0	0	0

The key issue was at what score should the mastery level be set? The following table shows the number of pupils who would fall off the tier if the mastery level was set at each of the possible scores. This analysis is done by tier.

	number of pupils who would fall off the tier if the mastery level was set at this number of credits								
	8	7	6	5	4	3	2	1	0
level 1	0	0	0	0	0	0	0	0	0
level 2	24	13	5	3	2	1	0	0	0
level 3	46	41	36	27	21	9	4	0	0
level 4	46	38	24	22	16	12	5	1	0
level 3	174	131	88	50	32	17	9	6	0
level 4	185	139	86	55	36	21	11	7	0
level 5	199	191	170	137	87	57	28	9	0
level 6	200	195	185	160	136	98	57	16	0
level 5	98	83	59	39	27	15	6	2	0
level 6	104	96	80	67	45	30	14	4	0
level 7	107	98	83	67	52	34	20	12	0
level 8	109	107	104	97	88	74	51	36	0
level 7	18	18	17	15	9	5	4	1	0
level 8	18	18	18	16	14	11	6	3	0
level 9	18	18	16	12	6	5	1	1	0
level 10	18	16	15	10	9	6	2	1	0

The important row is the first in each tier, as the majority who failed to achieve this level would probably also have failed at higher levels. In tier 1-4, for example, every pupil achieved a level 1 even at a mastery of 8 credits. At tier 3-6, a mastery level of 5 credits would result in 24% of those taking this test failing to record a level.

Applying a mastery rule

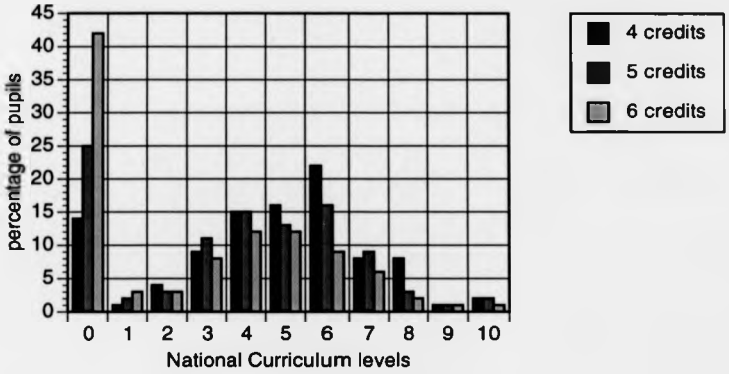
The previous information included each pupil's performance, at each level in the tier of the test paper which they took. In determining the level achieved, only each pupil's best performance should be considered. As in the analysis of Te4 mastery levels of 50%, 4 credits, up to 100%, 8 credits, have been considered. This analysis has been carried out irrespective of the tier which a pupil took. Therefore, percentages in the first row, 0 level, represents those pupils who failed to achieve a level on any of the four test papers at that mastery level.

All materials - Te 1

	numbers achieving highest level at this number of credits					*% achieving highest level at this number of credits*				
level	4	5	6	7	8	4	5	6	7	8
0	34	60	100	183	264	9	16	26	48	69
1	2	3	4	11	24	1	1	1	3	6
2	11	13	16	23	21	3	3	4	6	5
3	11	24	36	44	24	3	6	9	12	6
4	81	97	107	60	18	21	25	28	16	5
5	79	69	48	25	14	21	18	13	7	4
6	88	58	34	18	11	23	15	9	5	3
7	39	34	25	12	4	10	9	7	3	1
8	23	14	7	4	2	6	4	2	1	1
9	5	2	2	0	0	1	1	1	0	0
10	9	8	3	2	0	2	2	1	1	0

As with Te4, mastery at credit levels of 7 and 8 is clearly out of the question. At a mastery level of 6, 26% of pupils would fail to achieve a level; whilst at 5 credits, the number failing to record a level would drop to 15% and at 4 credits, 9%. Pupils performed better on Te 1 than on Te 4 (at 6 credits 46% fail to record a level;at 5 credits, 26% failed and at 4 credits, 14%). In all previous trials and pilots, this has always been the case, Te 4 has always resulted in pupils' poorest performance. The profile of performance, discounting those who failed to record a level, is consistent. However, the difference between achievement at level 3 and level 4 is significant. The majority of pupils took tier 3-6 and the majority of these achieved level 4, as would be expected. The entry for tier 1-4 was only 25% of the tier 3-6 entry and the majority of these achieved level 2, as would be expected;

consequently, the numbers achieving level 3 are significantly low. Below, the levels achieved at a mastery of 4, 5, and 6 are shown graphically.



Performance - Te 1 in relation to the practical task

If the pupils taking the test are subdivided into the practical task which they took, and the various mastery levels are applied, the percentage failing to record a level is as follows.

	Construction	Control	Food
mastery level			
4	10%	2%	12%
5	16%	8%	19%
6	26%	16%	34%
7	50%	41%	53%
8	69%	76%	75%

It can be seen that pupils taking Control, by this indicator, achieved the best performance in relation to Te 1. Pupils taking Food, who achieved the highest levels in relation to Te 4, demonstrated the lowest performance in relation to Te 1. Another comparison was made by examining the average level of achievement for Te 1 and Te 4 if a mastery level of 5 was applied.

	Construction	Control	Food	All
Te1	4.2	5.3	4.1	4.3
Te4	3.3	4.7	4.3	3.7

This confirms that the more able pupils were entered for Control. The performance of pupils taking Food was fairly consistent. The significantly lower performance of pupils taking Construction materials in relation to Te4 was very marked.

Determining a profile component

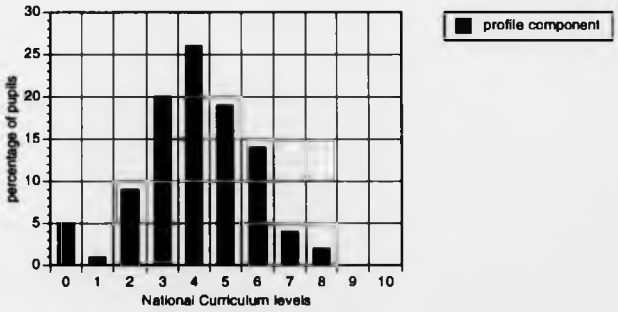
From the analysis, the fairest mastery level at which a sensible degree of success can be recorded is 5 credits, 62.5%. This mastery level was used in determining each pupil's profile component. In this analysis only pupils who had assessed levels in Te 2 and Te 3 were included. If a pupil failed to record a level in one of the attainment targets in the test, the level below the lowest in the tier was substituted. If a pupil failed to record a level in both attainment targets, they are deemed not to have achieved a profile component. The mean PC scores are shown below.

	Construction	**Control**	**Food**	**All**
mean PC	3.9	4.1	4.7	4.2

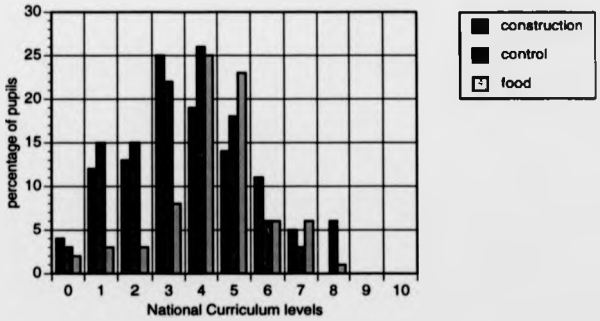
The numbers who failed to record a score because they fell off the tier in both attainment targets are shown below.

	Construction	**Control**	**Food**	**All**
% not recording a level	17%	2%	16%	15%

The graph below shows the profile component for all pupils taking part in the pilot.



The graph below illustrate the distribution of pupils in relation to the practical task which they took.



Time taken on the test

The table below indicates the time taken by pupils to complete the test in relation to task and tier.

Tier	Construction	Control	Food	All
1-4	52	50	52	51
3-6	70	56	61	62
5-8	68	68	70	69
7-10	80	80	71	77

With the exception of tier 3-6 the time taken is very consistent across the tier, with one exception. Construction tier 3-6, had the highest entry, resulted in the lowest scores but pupils spent relatively more time on the tests. Examining the test papers, it is difficult to see why. Answers are generally superficial and there is no evidence that these pupils found the test more difficult.

Conclusion

The extensive trialling and piloting which took place, even though a wide variety of instruments was used, resulted in some fairly consistent conclusions. Most notably, the Task Group on Assessment and Testing had indicated that the expected average level of performance at the end of key stage 3 would be between level 5 and level 6. There was no evidence in this research that pupils' achievement, when measured in the context of a genuine design and technology task, would have approached these levels of performance by the time of the first statutory assessment. Up until the 1992 Pilot pupils' performance, as would be expected, was gradually improving. The mean PCs in the three major trials were 2.7, 3.2 and 3.92. However, with the introduction of a differentiated practical task, focused on two attainment targets in the pre-1992/93 trial, performance slightly dipped. The mean levels for both attainment targets being 3.7. This cannot be compared directly with the PCs previously quoted as these resulted from the aggregation of all the attainment targets. If, for example, only Te 2 and Te3 were aggregated for the 1990, 1991 and 1992, the mean PC would have increased as these two attainment targets always produced the highest levels. This dip in performance is consequently more marked than the statistics suggest. Specified tasks, rather than open-ended ones were thought novel. It is likely that as teachers became accustomed to such specified, differentiated tasks that pupil performance would have increased accordingly.

Given the rate of increase evidenced between 1990 and 1992, 0.5 of a level in the

first year and 0.9 of a level in the second year, it is conceivable that the TGAT's norms (because in reality that is what they were) may have been achieved within three or four years. It would have been fascinating to have monitored these performance levels to see if they reached the predicted plateau and remained there. Without the use of calibration procedures, as used by GCSE examination groups, standardisation may have proved more problematic. Without statutory assessment, there is no way of monitoring mean performance in the context of the revised National Curriculum Order. This may create considerable issues when Teacher Assessment is the subject of a statutory order in 1997. Reported non-moderated teacher assessment may create a climate in which standards consistently rise in relation to GCSE performance which is relatively stable.

Other significant general conclusions can be drawn from all phases of this research in relation to pupils' performance as measured against the attainment targets. The traditional roots of the subject in making, still have a significant effect on pupil performance. Te3 produced the highest level of performance on every occasion. This is probably a direct result of both pupil and teacher motivation in this aspect of the subject. It is the aspect which the majority of pupils enjoy most and, at this key stage, the one on which teachers focus their teaching. The order of the attainment targets in relation to pupil performance was also highly consistent. Following Te 3, performance in the other attainment targets was always Te2, succeeded by Te1 and final Te4. It is not surprising that the two attainment targets with the highest practical content, Te 2 and Te3 resulted in the highest levels of performance. The inability of teachers to understand the importance of Te1 and their lack of conviction of its relevance was evidently transmitted to pupils. Evaluation is a vital component for success yet teachers found it difficult to embed this procedure into the process of designing and making and gain purposeful value from it. Instead, it was largely seen as an end of activity 'bolt-on' . This was reflected in the resulting assessments which could give little value to this tokenistic approach. Even when practical tasks were designed to focus on this attainment target there was no detectable improvement. Te1 and Te4 are undoubtedly key skills in the commercial world of product innovation, possibly more so than Te2 and Te3, but education has as yet not perceived their value and accorded them the esteem they warrant. The down-grading of these two attainment targets has been formally recognised in the revised Orders. This is seen by many as a retrograde step, encouraging a practical 'hands-on' approach to the subject rather than one which is thoughtful and informed.

Another highly significant feature was the manner in which girls outperformed boys in every type of task and test and in every attainment target. This pattern was established in 1990 and conformed to in all later trials and pilots. One must

conclude, that if these attainment targets represented the definition of design and technology between 1989 and 1993, then girls demonstrated a higher level of capability. However, the demands of statutory assessment for concrete evidence may have affected the way in which assessments were reached. This is supported by evidence that suggests that girls performed closer to their teachers' predicted assessments than boys, who typically did not match their teachers' assessments of their capability. Perhaps the demands of standardised tests were better understood by girls and they conformed more closely to these requirements. Both girls and boys demonstrated better levels of performance in tasks with a less technical content. This is perhaps a symptomatic indicator of design and technology teachers' strengths and weaknesses, within the breadth of the subject. It also illustrates the subject's weakness in relation to the use of science and mathematics in assisting pupils in making informed objective decisions about design and manufacturing issues rather than ones based merely on opinion.

The introduction of tests and tasks in which pupils were required to design and make to a specification was to ensure that pupils demonstrated their capability within pre-defined constraints. There was obviously no guarantee that evidence from an open-ended activity would necessarily conform to the restrictions of the statutory framework, which by its nature lacked flexibility. Similarly the tests sought to ask thought provoking and relevant questions. Questions of a general nature might have only provided partial evidence of capability; whilst those about the what, how and why of pupils' designing and making presented a more complete picture. As with any assessment procedure it is essential that the person being assessed understands fully what is expected of them so that they have a real opportunity to demonstrate their capability.

The results from these trials and pilots measured what pupils achieved, but indirectly they also measured teacher competence, expertise and commitment to the National Curriculum Order. Consequently, teacher evaluation of this process is highly pertinent and it is described and discussed in the following chapter.

References

1. Thorndike, R.L. & Hagen, E. (1951, 2nd edition 1972). *Measurement. and Evaluation in Psychology and Education*, New York: John Wiley & Sons, page 193
2. Hagedon, R. & Labovitz, S. (1971). *Introduction to Social Reserch.* New York: McGraw - Hill, page 79

Chapter 6

Teachers' evaluation of the assessment process

Synopsis

At every stage of the evolution of a statutory assessment process, the teachers delivering and assessing the procedure were consulted and their views sought in both qualitative and quantitative forms. Teacher evaluation followed every phase of the development and was used to inform subsequent phases. However, teachers' views had to be balanced against the specification and the continual changes which were being made to it. The requirement to respond to these imposed factors may have created the impression that teachers' views were largely ignored and in some respects, this was true. The teachers' message in relation to workload was loud and clear and eventually, backed by the force of law, it could not be ignored. This chapter, however, focuses on teachers' responses in relation to all aspects of the assessment procedure. It would not be feasible to report the evaluation of each developmental phase so two, particularly significant ones, have been chosen.

The first half of the chapter looks at the evaluation of the 1991 national pilot. This was the last pilot conducted under the original specification. This assessment procedure was based on an integrated approach to the attainment targets undertaken via a single practical/long task. The second half of the chapter presents the evaluation outcomes from the 1992 pilot. This was the final national pilot before the first statutory assessment. The intention was to pilot the model which would be replicated in the coming statutory assessment as closely as possible. Conducted in response to a revised specification, it required pencil and paper tests as well as practical tasks to be used in determining pupils' capability. The evaluations addressed many similar issues but the approach was not identical because of the different specifications under which each was carried out. The 1991 evaluation was more comprehensive and addressed a broader range of issues.

The chapter is divided into the following sections and sub-sections

The approach to teacher evaluation

The 1991 National Pilot

Teacher evaluation - the context and the respondents
The administration of assessment tasks in school
Manageability
Assessment and moderation
Other issues

The 1992 National Pilot

Teacher evaluation - the long task
The manageability of the long task
Teacher evaluation - the tests

The approach to teacher evaluation

Throughout the fifty two months in which schools were involved in this research fourteen significant trials and pilots took place. These involved a large number of teachers both in administering the assessment process and marking and moderating pupil outcomes. Cumulatively these teachers provided an invaluable resource in relation to all the key issues concerning statutory assessment. At the outset the majority was committed to the implementation of National Curriculum design and technology because, with its introduction, the subject appeared to have gained the status which previously they felt it had been denied, Teachers were less enthusiastic about the assessment requirements that this status conferred, but they were prepared to cooperate because of the associated benefits. The changes brought about by political intervention, mainly in response to teacher evaluation and protests at key stage 1, created confusion, dismay and frustration. This was gradually reflected in teachers' evaluations and willingness to cooperate. The work described in this thesis was greatly informed by participating teachers and their comments were given high status in relation to changes and modifications. Frequently the rapid changes to the specification prevented teacher's common sense views from being implemented. The developmental approach was always intended to be one of partnership and collaboration and this was maintained until the details of the first statutory assessment were released.

Every teacher taking part in any trial or pilot was given the opportunity to complete an evaluation form. In addition to the quantitative information, collected via OMR forms (optical mark read), teachers were given every opportunity to supply qualitative information. This was collected via meetings and face to face discussions as well as more formal response forms. This chapter, because of the extensive nature of the qualitative information, is primarily based on the quantitative data. In 1991 teachers were asked a question to which there were three graded responses. A fourth box allowed them to provide their own answer. If possible, these responses were coded to the a, b and c responses. In addition, a separate coded sheet was provided for fuller comments. In 1992 teachers responded on a four point ranking scale. For the purposes of reporting outcomes to SEAC these were sometimes collapsed to produce an approval/non-approval response. These teachers were also provided with a structured response sheet for detailed comments. Not all teachers answered every question. When percentages are quoted they are in relation to the responses reported, the percentage of missing data is not stated.

Following the 1991 pilot all participating schools were provided with a complete report of the trial but because of the confidential nature of the last pilot before the first statutory assessment, this was not allowed by SEAC in 1992.

The 1991 Pilot

Teacher Evaluation - the context and the respondents

INSET took place during January 1991 and the first schools embarked on the SAT by the beginning of February. All schools finished by the designated date and all returned OMR forms in relation to pupil assessments by the 7th July. Very few pupils, however, experienced an identical SAT - a reflection of the nature of design and technology. If the structural variables such as teacher expertise, environment and curriculum design are coupled to the requirement placed on pupils to identify their own task it was apparent that this would inevitably be the case.

The subject breakdown of teachers completing the evaluation form was as follows:

Art & Design 10%
Business Studies 1%
Craft, Design & Technology 44%
Information Technology 3%
Home Economics 24%
Textiles 14%
Others 4%

The Administration of assessment tasks in school

INSET

The introduction of the National Curriculum in design and technology posed the majority of schools with a new curriculum challenge. Consequently INSET provision assumed a much wider significance than preparation just for the pilot of an assessment procedure, involving delivery, assessment and standardisation. The piloting proved to be a potent means of exemplifying the subject to the majority of schools taking part, especially as the materials which had been developed fully embraced the philosophy which underpins the statutory Order. Evaluation of the INSET undertaken is provided in appendix 6.1, p. 329, along with other information pertaining to teacher evaluation, appendix 6.2, p. 333.

Post SAT Evaluation

Following the SAT activity teachers were asked to reflect on the value of the pre-SAT INSET. The 36% receiving INSET from trained personnel, gave a mean rating of 2.91 on a 1 to 4 rating scale. Whilst the 56% who received INSET from a colleague gave a mean rating of 2.56. Both means represent a good approval rating for the quality of the INSET provided. Coordinators were provided with a school kit containing all the materials required for INSET. This kit received a mean rating of 3.2. It was categorised as excellent by 36% of the respondents and good by 46%. The kit was thought to provide sufficient information by 82% of the coordinators and a similar high approval rating was given to the, "Delivering SATs"

video in terms of helping teachers understand what a SAT involved. Of the 75% of teachers receiving INSET at a school level, the length provided varied as shown below:

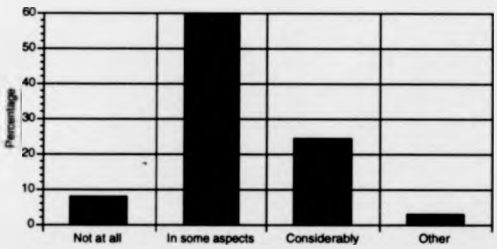
12% less than 2hrs, 38% half a day, 50% one day.

Manageability

Teaching style

During the period of this research, design and technology INSET was concerned with more than just assessment issues. It was interesting to observe if involvement in the pilot helped teachers in their understanding of NC objectives and influenced their approach to the subject. Coordinators were asked to reflect on the effect on their colleagues and 89% reported that colleagues' understanding and delivery had been enhanced in some aspects, with 35% of this total believing many aspects had been enhanced. At a faculty level, 88% of coordinators felt that taking part in the pilot had been of some help in implementing NC design and technology. 50% of this total felt that involvement had been very beneficial.

Individual teachers were asked if delivering the SAT had involved changes in how they taught the subject. The response, shown below, clearly illustrates the inexperience which teachers have in approaching design and technology from the perspective of the statutory Order.



Organisation

This was a crucial aspect of a SAT. The responses indicate that teachers did find it difficult to manage the activity undertaken during the pilot. Teachers were asked if their approaches to organisation changed during the course of the pilot. Three-quarters of those delivering the SAT responded positively with 25% acknowledging that the change was considerable. This was a clear indication of the novelty of the SAT to the majority of teachers and it undoubtedly affected their ability to manage their delivery and the situations which developed as a result.

The analysis by subject expertise of this question identified art and design as being the subject which included the highest percentage of teachers who found the

changes considerable. This was in contrast to textile teachers who found the approach relatively similar to their normal mode of operation.

Did your normal approaches to organisation change during the pilot?

	(% of responses)				
	not at all	***in some aspects***	***considerably***	***own resp.***	***no. resp.***
Art & Des.	8.8	47.1	32.4	8.8	2.9
Bus. Stud.	——	100.0	——	——	——
CDT	9.3	61.6	23.2	2.2	3.7
Home Econ.	6.0	49.0	22.0	10	13
Info. Tech.	18.0	72.0	9.0	——	1.0
Textiles	18.0	72.0	9.0	——	1.0

It is not surprising then that teachers responded to overall manageability in a less than positive fashion. On the four point rating scale the mean rating was 2.1. Around 45% found the manageability difficult and 15% very difficult, with the remaining 40% describing it as good or better. Below are shown responses in relation to the three SATs.

	excellent			***poor***	***mean rating***	***no resp.***
	4	3	2	1		
Exhibition	--	38%	38%	16%	2.24	8%
Measurement	2%	28%	48%	12%	2.22	10%
Public Places	--	25%	54%	19%	2.1	2%

There are some differences: Public Places appeared to have provided teachers with more management challenges and Exhibition with less -the SAT with the tightest structure, Measurement, was rated somewhere between the two but significantly had the lowest number of poor responses.

If a similar analysis is undertaken by subject expertise, the results support the previous analysis, that textiles and the CDT teachers had least problems in managing the activity. Of the principal contributors, art and design teachers had most difficulties. A third fewer art and design teachers rated the manageability as good than in the three other key subjects. The novelty of the activity for both business studies and information technology teachers is clearly apparent.

	Excellent			***poor***	***mean rating***	***no resp.***
	4	3	2	1		
Art & Design	—	24%	47%	20%	2.03	9%
Business Studies	—	25%	25%	50%	1.75	
CDT	2%	33%	44%	13%	2.24	12%
Home Economics	—	33%	43%	19%	2.15	5%
Info. Tech.	—	10%	63%	27%	1.82	
Textiles	—	38%	42%	14%	2.26	6%

Length of the SAT

For the design and technology SAT, pupils were given up to fourteen hours to complete the task. Up to four hours was taken in completing the short task and the quest. Teachers were asked to report if the time allowed was sufficient. This is analysed firstly by SAT type and then by subject expertise.

The time the SAT allowed pupils to provide evidence of their capability was?

	% of responses (no response - not shown)			
	insufficient	***reasonable***	***excessive***	***own response***
Exhibition	46	32	7	13
Measurement	58	25	6	7
Public Places	43	34	13	8

The time the SAT allowed pupils to provide evidence of their capability was?

	% of responses (no response - not shown)			
	insufficient	***reasonable***	***excessive***	***own response***
Art & Design	41	35	9	14
Bus. Studies	75	25	—	—
Craft, Des. & Tech.	51	26	7	13
Home economics	50	36	7	4
Info. tech.	63	18	10	—
Textiles	42	44	10	2

SAT type analysis reveals that Measurement was seen to require more time, 12% to 15%, than the other two SATs. Although teachers found Measurement easier to manage because of the tighter structure, this appears to have some drawbacks in relation to the time available.

Disregarding information technology and business studies, because of the small number of respondents, the surprising result by subject expertise was the high percentage of home economics teachers who found the time allowed insufficient. It was predictable that craft, design and technology teachers would be the most dissatisfied with the time allowed, the nature of the resistant materials used in this aspect of the curriculum being directly responsible. In earlier trials home economics teachers have weighted responses towards the time allowance being excessive. Yet again, the style of SAT seemed to fit textiles most closely.

The amount of additional time required, by SAT, shows little difference between the SAT models. The one significant feature was the much higher percentage of pupils who required more time to complete Measurement.

What % of pupils required more time ?

up to	10%	25%	50%	75%	100%
Exhibition	21	11	16	20	32
Measurement	22	12	12	12	42
Public Places	20	10	23	12	35

When categorised by subject expertise, craft, design and technology teachers stand out clearly as those whose pupils required a lot more time - 70% at least needing 50% more time. By comparison, home economics teachers report many fewer pupils needing such an amount of extra time and textiles seems to be closest to fitting the SAT model.

What % of pupils required more time ?

up to	10%	25%	50%	75%	100%
Art &Design	17	30	12	9	32
Business Studies	25	25	—	—	50
Craft, Des. & Tech.	12	8	20	18	42
Home Economics	28	12	30	5	25
Info. Tech.	28	12	—	25	35
Textiles	40	5	12	13	30

The teachers' perception that pupils required more time was not supported by the pupils, they felt the time available was adequate. It is central to design and technology that pupils gain experience and consequently learn how to manage the available resources. When deciding on a task, foremost in their mind should be the question: can it be completed in the time available? Evidence shows that high levels are only evidenced when pupils exceed the time available.

Preparation time

All teachers were asked to comment on the amount of preparation time required. Overall about 60% felt no more preparation was required than normal, 14% of these stating the preparation was less than usual. Around a third felt the SAT took longer to prepare, craft, design and technology teachers being slightly above this average. Art and design teachers commented strongly that less time was required which is surprising as they indicated greater difficulty with managing the activity.

	% of teachers(no response - not shown)			
Was preparation time	***less than***	***about the same***	***more than usual?***	***own resp.***
Art & Design	40	14	18	12
Business Studies	—	100	—	—
Craft, Des. & Tech.	12	40	38	4
Home Economics	20	33	32	6
Info. Tech.	10	38	52	—
Textiles	4	46	32	8

Teachers were also asked how much extra time was required to prepare for the SAT on average per week.

	% of teachers				
	no extra time	*up to 1 hour*	*2 hours*	*3 hours*	*4 hours*
Art & Design	70	3	23	—	4
Business Studies	100	—	—	—	—
Craft, Des. & Tech.	65	10	16	4	5
Home Economics	71	10	12	3	4
Info. Tech.	56	28	8	8	—
Textiles	70	12	14	2	2

This analysis produced a similar picture with 5% more craft, design and technology teachers than in the other three main deliverers having to do additional preparation. By SAT type, Measurement was found to require more preparation which was surprising as it had the most structured framework.

Resources

There was a strong indication from teachers that, in relation to the normal level, more resources (consumable materials) were used by pupils during the task.

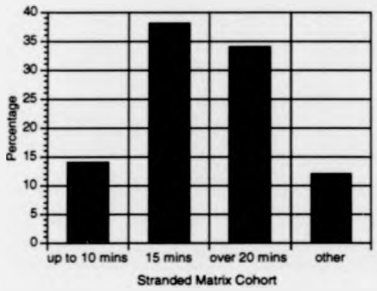
	% of teachers (no response - not shown)			
Were the consumables	*less than*	*about the same as those normally required?*	*more than*	*own response*
Art & Design	6	15	65	10
Business Studies	—	—	100	—
Craft, Des. & Tech.	3	32	48	12
Home Economics	7	10	70	4
Info. Tech.	3	18	54	18
Textiles	2	32	60	4

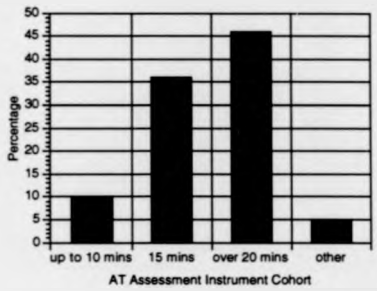
Craft, design and technology teachers had less difficulties in supplying what pupils required, whilst Home Economics found the resourcing issue significantly more demanding. This was also a comment on the way in which departments resourced projects. Pupil centred activities require a greater variety of resources than the traditional teacher initiated project. It was also a reflection of the requirement for basic design orientated materials to be available in every area of the faculty, which was previously not the case. The quality and quantity of available resources varied greatly from school to school. This raises the question, if the test is intended to be standard should not the resources also be similar?

School facilities were investigated via the coordinators' evaluation form. Thirty two percent felt the facilities required for the SAT were more than usually required. The remaining 60% commented that the facilities required were about the same or less than normal. Pupil performance was felt by 18% of the respondents to have been seriously affected by the lack of facilities whilst 33% felt performance had been slightly affected in this respect.

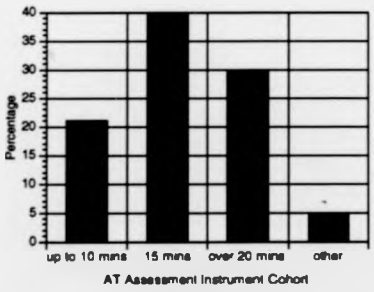
Assessment and moderation

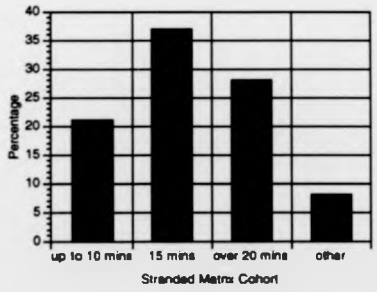
All aspects of evaluation in respect of assessment were analysed in relation to each of the assessment devices. The objective of both was to establish a pupil's operating level in each attainment target. Although the assessment devices were allocated to clusters of schools, the AT instrument was used by a third more teachers than the stranded matrix. A crucial aspect is how long the assessment procedure took. To put this in context teachers were asked to comment on how much time they would normally devote to assessing a Y9 pupil's achievement during a term. The analysis indicated that the cohort using the AT instrument spent more time carrying out assessments - 12% more in the 20 minute plus category.





At the end of the assessment procedure, teachers were asked how long the assessment had taken.





The assessment procedure for the SAT appears to have taken less time than teachers would normally spend on assessment during the term. Both devices show 7% more teachers carried out the assessment in under 10 minutes. In addition, the number of teachers normally spending over 20 minutes decreased by 6% for the stranded matrix and 16% for the AT instrument. Even though 60% of teachers managed individual assessments in under 15 minutes, there were 30% of teachers who took in excess of 20 minutes.

Record keeping

In the infancy of the National Curriculum, teachers of design and technology were still coming to terms with the task of recording performance. The outcome, either the product or the project file or both, has traditionally been the sole evidence for determining achievement. Aspects of the statements of attainment were concerned with operational performance and teachers needed to record what was taking place if credit was to be given. Only 7% of teachers managed any recording during a lesson, 30% after a lesson and 36% not at all. This question provoked 18% of teachers to make their own response. The general tenor of these comments centred on the impossibility of what was being asked. If coded these replies would boost the, "not at all," category to 54%.

The difficulty of recording achievement was supported by the response that only 13% of SAT assessments were based jointly on the SAT outcomes and teacher records. Teacher recall of performance was the basis of 46% of SAT assessments and 28% were based on the traditional outcome alone. This emphasises the very difficult task teachers operating in practical areas had of helping children progress individual work, maintain a safe working environment for all the pupils and in addition record pupil achievement. Most teachers regarded the latter as being of the least importance. On a subject analysis only 31% of CDT teachers carried out recording during or after a lesson and 60% carried out no recording at all. Of art and design teachers, 47% made records on which assessments were based.

	% of teachers (no response - not shown)			
pupils' achievement recorded	***during the lesson***	***after the lesson***	***not at all***	***own response***
Art & Design	3	44	27	12
Business Studies	—	50	50	—
Craft, Des. & Tech.	5	26	41	19
Home Economics	12	26	29	20
Info. Tech.	—	36	27	27
Textiles	6	36	30	20

Coordinators were asked for their view of record keeping during the SAT. Some schools tried to keep to an agreed format - 46% - but, of these, in only 13 % of schools did all teachers use the same procedure. The remaining responding - 45% - let teachers use their own methods. In commenting on the quality of record keeping only 3% were rated as detailed and informative whilst 40% were felt to be rather spasmodic, probably non-existent. Another 47% felt teachers did as much as could be expected which may have included both after lesson notes and teacher recall. The record keeping folder helped 60% by providing a model.

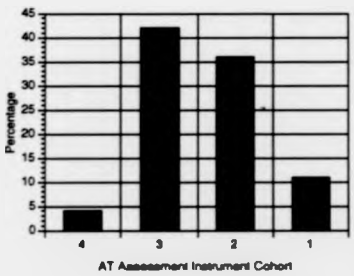
Teacher assessment

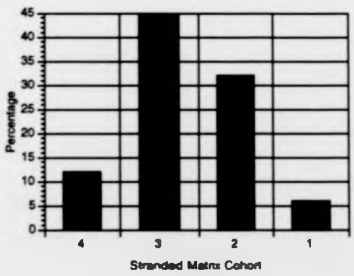
Teachers found the Statutory Order for design and technology difficult to interpret. This was one of the reasons why teachers were slow to put it into operation. All the schools taking part in the pilot had indicated that teacher assessments would be made for Y9 pupils and that these pupils would have followed the programme of study since the beginning of Y8. However, by the time of the pilot only one school had made teacher assessments. The majority had begun to adopt National Curriculum assessment procedures for Y7 pupils but some schools still had not started these procedures three years after the Order became law.

Consequently, teachers needed help to find a quick yet reasonably accurate method of producing a teacher assessment. To do this the level guide was developed (it consisted of a single cogent statement that attempted to pick out the most distinctive feature of each level for each attainment target). The level guide met with approval from 75% of teachers. Only 22% did not find it helpful and preferred to use their own system. There was no difference in approval rating from any of the evaluation subsets.

Assessment devices - the stranded matrix

Of the teachers using this system 74% found it satisfactory to use and 13% found it too difficult to use. Even though it organised the SoAs and offers unambiguous procedures 40% of teachers found it too complex. In contrast 16% found it rather superficial. A further 40% found it fair and objective. On the rating scale from 4 (excellent) to 1 (poor), this system produced a mean of 2.65





The AT Instrument

This received an 80% approval rating from teachers with 12% finding it too difficult to use. Some teachers found the structure difficult, 35% reporting that it was over complex. However 50% rated it fair and objective, 13% found it superficial. On the rating scale this system produced a mean of 2.41.

The labels - in relation to assessment

The labels were a key aspect of the assessment procedure. They were designed to be useful to both pupils, during the activity, and teachers when assessing the pupils' outcomes. Teachers found them very useful, 80% rating the labels as helpful in assessing work. Only 13% felt they were of no use. This response was similar regardless of the assessment device used. Significantly a higher percentage of teachers using the AT instrument rated the labels as very useful compared with those using the stranded matrix.

Moderation

There were two stages to the moderation procedure. Coordinators were encouraged to organise in-school moderation which was followed by a formal moderation carried out by a trained moderator. Of the schools taking part, 60% had a plan for moderation, 18% working to a set procedure. The remaining 42% insisted on teachers carrying out assessments in groups and coming to a consensus view. Those not working to a plan believed that informally teachers worked together but this may or may not have been the case.

Coordinators were asked for an overall view of the quality of assessments in their school. Around 20% stated assessments were inconsistent, 62% reasonably consistent and 9% standardised and consistent. This compared with the evidence from the external moderation procedure shown on the graph below. External moderators placed schools in one of four classes:

1. Assessments consistent and standard
2. Majority of assessments consistent and standard
3. Assessments consistent but not standard
4. Assessments inconsistent

It was not anticipated that assessments in the group 1 schools would produce higher AT levels. Indeed, the possibility was that this group of schools would produce lower but more secure assessments.

Each school was supplied with sets of SAT sample assessments. These received

an approval rating of 52% whilst 40% of teachers commented that they found them of little help. This figure was high but it was discovered during moderation that a significant number of teachers had not seen these sample assessments. Coordinators asked a similar question stated that 70% of these sample assessments had been helpful to colleagues. Only 25% felt they had been of little assistance.

Other issues

Disruption

Coordinators were asked to comment on whether the SAT created any additional disruption. Teacher absence was seen by 50% to be no more problematic than usual but 38% did feel additional problems were created. These coordinators commented on the constraint of operating within a time window, continuity and fairness of opportunity for pupils being the key issues. Normally pupils might be supervised by a non-specialist teacher in a non-practical environment in the event of teacher absence. This obviously caused problems in the SAT context. Schools adopting a team teaching approach recorded less problems in this respect than those with delivery based on class groups.

The SAT was felt to have caused major timetable disruptions in 13% of the schools in the pilot, 45% registering no disruption and 32% some minor disruption. These results were very similar when coordinators were asked about the likely situation in two years' time. This was surprising as most schools were planning to move to blocked timetables, which would allow more flexibility in the curriculum, by 1993.

Schools provided information of events which disrupted the SAT, they listed the following:

Bank holiday	Easter holiday
Building works	INSET
School visits	Medicals
GCSE Moderation.	

Two thirds of the schools taking part noted some disruption during the SAT period. Undoubtedly the most common complaint by teachers in this respect was that the SAT coincided with GCSE exams. Many felt it unfair that both assessment activities should take place concurrently.

It was felt by the majority of teachers that the Easter holiday created a serious disruption to the activity. Pupils of this age found it very difficult to remotivate themselves following the vacation and consequently time was lost. Teachers would prefer the activity to be contained in a term and ideally half a term. Many

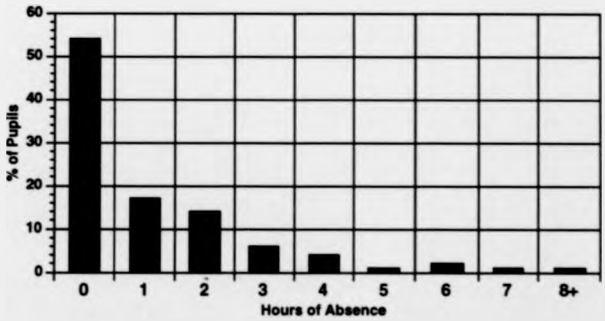
recognised, however, that with the diminishing time in the curriculum for design and technology it is becoming impossible to undertake an appropriate task in less than a term. Additionally teachers reported in many schools that group sizes are being increased. Over twenty pupils in a class was common and some schools have adopted a whole class, thirty plus, approach to this subject. This had a serious impact on the ability of schools to deliver NC design and technology.

Schools were also asked if the SAT had disrupted school routines outside the design and technology department. Eleven schools, 9% of the sample, felt this had occurred. The examples cited were:

'additional workload on teachers'; 'disrupted scheme of work'; 'affected pupils as they wished to do extra work during breaks'; and 'had to book the hall to show the video'.

None of these reasons gave cause for concern, one was indeed a cause for optimism.

If a SAT involved a long task, then pupil absence was an important factor. The graph below shows pupil absence in hours during the SAT window.



It is difficult to decide at what point absence seriously affected performance. Comparisons with teacher assessments are only valid if there is confidence in these assessments. Approximately 10% of pupils missed more than one fifth of the allowed time. An analysis of these pupils' performance was required to determine at what point teacher assessments could be substituted due to absence.

Style and fairness

Teachers and coordinators were asked a range of questions about the style and fairness of the SAT. Teachers were asked if the model of SAT allowed pupils to demonstrate their capability. Overall, 55% felt the task was reasonable, 10% too demanding and 24% felt the task provided insufficient opportunity for pupils to

demonstrate capability. As there were three models, this question is analysed by SAT type below.

In the context of NC design and technology does this model of a SAT offer pupils....

	% of teachers (no response - not shown)			
	little	***reasonable***	***too demanding***	***own response***
	opportunity to demonstrate their capability?			
Exhibition	14	66	9	10
Measurement	31	44	10	12
Public Places	29	51	12	6

There was clearly greater satisfaction with Exhibition, the SAT model which provides the widest range of opportunities. The structured SAT, Measurement, received the lowest acceptance rating. Measurement and Public Places, both of which in different ways either focused or guided the task, were seen by a significant percentage to offer little opportunity for pupils to demonstrate capability. If the same question was asked by subject expertise the results were as follows.

	% of teachers (no response - not shown)			
	little	***reasonable***	***too demanding***	***own response***
	opportunity to demonstrate their capability			
Art & Design	—	44	14	6
Business Studies	50	25	—	25
Craft Des. & Tech.	22	57	10	10
Home Economics	23	60	8	8
Info. Technology	27	36	18	18
Textiles	26	58	8	4

Of the prime deliverers (CDT, HE and textiles), at least 57% felt the SAT offered a reasonable opportunity to demonstrate capability. One in four teachers believed there was little opportunity to demonstrate capability. In the main these teachers believed the time window was too short and teacher assessment provided a more reliable measure of pupil capability.

The themes were seen as offering a fair challenge by 58% of the respondents; 17% found them narrow and restrictive and 6% too difficult. An analysis by SAT theme revealed the following.

If your pupils had followed the KS3 programme of studies would they have found the theme:

	% of teachers (no response - not shown)			
	restrictive	***fair***	***too difficult***	***own response***
Exhibition	4	67	8	10
Measurement	35	36	7	12
Public Places	14	64	7	7

Quite clearly Measurement was seen as being far more restrictive than the two other themes. Teachers commented on the, "technical nature," of the theme which they thought was, "unfriendly and irrelevant."

analysis by subject of teachers delivering Measurement (no response - not shown)

	restrictive	*fair*	*too difficult*	*own response*
Art & Design	45%	33%	5%	20%
Business Stud.	—	—	—	100%
Craft, Des. & Tech.	30%	47%	8%	5%
Home Economics	40%	33%	5%	14%
Info. Tech.	75%	10%	15%	—
Textiles	54%	30%	—	2%

Clearly craft, design and technology teachers did not experience the same difficulty with this theme as teachers in other subject areas. This has implications if themes are nominated which are technical in nature.

When all the themes were analysed by subject expertise, there appeared to be very little difference between the way in which the themes were viewed across the subject areas.

% of teachers (no response - not shown)

	restrictive	*fair*	*too difficult*	*own response*
Art & Design	15	62	3	9
Business Stud.	—	50	—	50
Craft, Des. & Tech	17	57	8	8
Home Economics	15	61	2	8
Info. Tech.	28	46	9	9
Textiles	14	62	4	8

On the evaluation form, teachers were asked to suggest themes which they felt would be appropriate for a SAT. Despite the large number of titles there was virtually no agreement. Leisure and Travel appeared a few times as did Educational Toys but no title reached double figures in over 300 separate responses. There was evidence of subject bias in the suggested themes, some examples are given below:

Art and Design	- Organic structures, pattern and evolution
Business Stud.	- no suggestions
Craft, Des. & Tech.	- Mechanisms, Packaging and Movement
Home Economics	- Shopping, Community Health and Family Holidays
Info. Technology	- Communication
Textiles	- Castaways, Festivals and Childrens' Clothing

Many teachers obviously found it difficult to perceive the wider issues relating to design and technology. Several teachers commented that, "it is not our job.....only for a fee.....no theme is suitable." The number of these responses were more significant than any others!

Teachers were asked to comment on the structure of the SAT rather than the theme. The overall response was 33% too loose, 41% about right and 16% too tight. Analysis by SAT type is more revealing.

Do you think the structure of the SAT was:					
	too loose	***about right***	***too loose?***	***own resp.***	***no resp.***
Exhibition	33%	43%	9%	7%	8%
Measurement	18%	43%	28%	5%	6%
Public Places	18%	43%	27%	6%	6%

The tighter structured SATs received almost identical ratings and the percentage of teachers who believed these structures to be too tight is very close to those who feel that Exhibition was too loose. It is clearly a challenge to create the right balance.

Analysis by subject expertise revealed one clear piece of evidence. Art and design teachers are more uncomfortable with a structured activity. In contrast textiles teachers, who had less problems with managing the activity, felt the structure was too tight. It is clear that within subject areas there are sub-categories, for example the traditional needlework teacher as opposed to the genuine textiles teacher. It was difficult in any subject analysis to take these individual subject interpretations into account. There was however a significant group of around 32% who would prefer a tighter structure.

Do you think the structure of the SAT was:					
	too loose	***about right***	***too tight?***	***own resp.***	***no resp.***
Art & Design	18%	24%	40%	12%	6%
Business Stud.	25%	75%	—	—	—
Craft, Des. & Tech.	30%	48%	12%	5%	5%
Home Economics	38%	35%	12%	6%	9%
Info. Tech.	36%	45%	18%	—	1%
Textiles	44%	38%	10%	2%	6%

Teachers were asked if they felt this style of SAT was appropriate for pupils at the end of key stage 3. Very few, only 2%, felt it was too simple, 52% believed it to be appropriate and 28 % too complex. Analysis by subject generally reflected these responses except in relation to business studies and information technology. These teachers would however have found the approach extremely novel given their background.

For pupils at the end of key stage 3 is this style of SAT:				(no response - not shown)
	too simple	***appropriate***	***too complex?***	***own response***
Art & Design	—	50%	32%	15%
Business Stud.	—	25%	75%	—
Craft, Des. & Tech.	4%	54%	25%	12%
Home Economics	—	54%	24%	14%
Info. Tech.	—	18%	45%	18%
Textiles	—	60%	28%	6%

For pupils at the end of key stage 3 is this style of SAT:				(no response - not shown)
	too simple	***appropriate***	***too complex?***	***own response***
Exhibition	2%	53%	20%	13%
Measurement	—	45%	32%	17%
Public Places	4%	54%	31%	7%

An analysis by SAT type showed the style of Exhibition and Public Places received the greatest approval. However, fewer respondents found Exhibition too complex when compared to the other SATs. Measurement again received the lowest appropriate rating despite its tighter structure which many were requesting.

Pupils' views

The prime objective of a SAT activity was to determine each pupil's true level of capability. This will only be achieved if pupils are motivated by and committed to the task in hand. An additional form of motivation may occur when the majority of pupils understand the status and importance of a SAT. During the pilot, schools were encouraged to treat the SAT as a normal activity and not put pupils under undue pressure. Some schools attempted to make pupils adopt a serious approach to the task by stressing the importance of the assessment in summative terms. There was little, if any, evidence of parental pressure affecting either performance or outcomes other than that which might normally be expected. This will undoubtedly change as parents become tuned into the process of national curriculum assessments.

To discover pupils' views of the SAT, an in-depth questionnaire was undertaken with a representative cross section. The sample was selected across subject areas and performance to be consistent with the pilot sample. This evidence was gathered using interviews. This technique was used in an attempt to overcome the deficiencies in data from previous pupil questionnaires. All interviews took place within a week of pupils' completing the SAT. Consequently, the sample consisted of approximately 1% of the pupils in the pilot in 5 schools - 8% of the sample. Of these pupils, 78% did the Exhibition SAT, 19% Measurement and 16% Public

Places. The subject was certainly enjoyed by this group of pupils, as 90% rated design and technology as good or better; however just less than half chose to continue the subject at key stage 4.

These pupils confirmed that for the majority of Y9 pupils, design and technology was a novel experience. 100% thought the SAT was different from what they normally did and 75% found it quite a bit different. More importantly 89% felt they had learnt something from the task (8% learnt very little, 23% learnt a lot). The experience from the perspective of these pupils must be described as very positive.

Responses in relation to the materials were mixed and did not always reflect the views of their teachers. For example, two thirds of the pupils found the Quest helpful when starting to think about the project, whilst less than 50% of teachers rated this effective.

There was some evidence that pupils taking different SATs responded differently to the pupil materials. The video associated with Public Places received the highest response rating with 67% reporting that it helped them find a topic for their project. Exhibition received a 42% rating and Measurement 30% in relation to this criterion. The Quest for Public Places was rated as very useful by 78%, Exhibition 69% and Measurement 48%. Teachers consistently reported that pupils found it difficult to decide what to do in response to a context and pupils supported this, 62% indicating that they found it quite difficult. When this point was pursued pupils typically responded,

" we're usually told what to do."

This supported the early response that it was different from what they did normally. There was no marked difference between SAT types in relation to this issue.
The labels proved very useful to 62% of the pupils and only 8% found them of no use - this included 5% who were told not to use them!

One element of clear disagreement with teacher evaluation was in relation to the time allowed for the SAT. Two thirds of the pupils felt they had enough time and one third would have liked more time. In comparison, over half the teachers felt their pupils needed more time. Another 10% of teachers felt the SAT was too long but no pupil felt this was the case. Exhibition and Public Places conformed to this pattern whilst Measurement was distinctly different. Only 8% of pupils piloting Measurement felt that they did not have sufficient time. This may well be a reflection of the greater structure in this SAT which was developed from the guided model. If pupils supplemented lesson time to complete the SAT, this was generally undertaken at home. Only 9% did much extra work during break or lunch whilst 31% did a lot of additional work at home. Two thirds did little if any extra work at school whilst 32% did little, if any, extra work at home.

In summary, the majority of pupils found the activity novel and beneficial. Some felt this novelty prevented them from doing as well as they could but most were reasonably satisfied with their performance. Some differences emerged between teacher and pupil perceptions and between the style of the three SATs piloted.

Evaluation of teachers previously taking part in a pilot (warm teachers)

Although the number of warm teachers participating in the pilot was small, an analysis of their evaluation forms was undertaken. Responses were generally more positive when compared with the cohort evaluation, as illustrated below.

These teachers appeared to be more familiar with the NC Order as 70% regarded theassessment procedure as fair and objective. This familiarity was also reflected in assessment time - 82% reporting assessments taking less than 14 minutes per pupil. Mean ratings of 3.45 for the level guide, 91% rating 3 or above on a 4 point scale, 3.0 for the assessment device and 2.8 for the assessment procedure, illustrate the higher levels of approval in relation to all aspects of assessment. 84% felt the SAT was a fair challenge for pupils having followed the programme of study and 74% felt the style was appropriate for pupils at the end of this key stage. 45% felt that their teaching style had not changed during the SAT and another 45% reported changes in some aspects. In the whole evaluation cohort only 8% reported NC changes, 60 % in some aspects and 23% considerable changes. This was clear evidence that involvement in an earlier trial had had a lasting effect on these teachers. These teachers also rated the pupil materials more highly. The video received a mean rating of 2.64 and the quest of 2.30. Perhaps these teachers appreciated more clearly how to use these materials for the benefit of their pupils.

The previous experience gave these teachers greater confidence in managing the activity; a manageability rating of 2.5 compared with 2.1 for all teachers. Exposure to the issues relating to the Order made teachers more responsive to its requirements and appreciative of how the SAT reflected all these aspects.

Overview

In this analysis of SAT administration, there were many areas of conflicting evidence. This was a direct response to the then current state of flux in this area of the curriculum. Teachers and schools were still coming to terms with and adjusting at different rates to the implications of the national curriculum. It would have been surprising if a clear consensus in relation to these issues had emerged. A majority supported the developments and the strategies employed. With experience of the procedures and processes teachers would be able throughout the key stage, by exposing pupils to the PoS, to furnish them with the ability to tackle this type of task with confidence.

The National Pilot 1992

Teacher evaluation - the long task

The manageability of the long task

The nature of design and technology brought together teachers from the six contributing subjects. In this pilot the percentage of teachers, by subject background, managing the long task was as follows:

Subject	*percentage*
Art and design	- 10%
Craft, design and technology	- 44%
Home economics	- 26%
Information technology	- 3%
Business studies	- 3%
Textiles	- 12%
Others	- 3%

For pupils to be able to provide a personal response to the long task it was important that it was described in a manner which could be understood, yet was not too prescriptive. If pupils found the task could be interpreted in a wide variety of ways, teachers would probably find the management of the task difficult. Consequently, teachers were asked:

Was the way in which the long task was described, too specific or too loose?

loose	about right	tight	too tight
16%	61%	17%	5%

The majority were content with this aspect as the 'about right' response accounted for almost two thirds of all responses. As these were the first tasks in which the material in which they should be undertaken was specified, it was important to discover teachers' reaction to this approach. One reason for the introduction of this approach was to ease management issues.

Should the material in which a task must be undertaken be specified?

Yes - 33%	No - 67%

This response indicates that teachers had, after thirty months, fully accepted the requirements of the Order even though those responsible were keen to modify the approach. Another aspect of interest was teachers' view of artefacts, systems and environments.

Which of these: artefacts, systems and environments, is the most suitable for end of KS 3 assessment?

artefacts - 60%;	environments - 17%;	systems; - 24%

This was expected, especially as some of the materials specialists were restricted to outcomes which could only be products. Not surprisingly, the majority of teachers supervising control material tasks felt systems was an essential category. The number of tasks was also a key issue in relation to management. In this pilot there had been five tasks, one in each of the important material bases within design and technology. To ascertain if teachers agreed, they were asked:

How many tasks should there be?

number of tasks	1	2	3	4	5	more
	21%	0%	0%	0%	70%	9%

The greatest majority was clearly in favour of the regime which they had just experienced. This, for example, allowed textiles teachers to supervise textile tasks. It accepted the traditional divisions of the subject, enabling specialist teachers to operate in specialist environments. One fifth of respondents indicated that this approach distorted the intentions of the statutory Order because key decisions were made for the pupil. These teachers would have preferred a single investigation task which required pupils to make decisions about what they were going to do before determining the most appropriate medium in which it might be realised. Those requesting more tasks were mainly business studies teachers who believed the place of their discipline had not been fully recognised.

Tasks of this nature may have created resource problems in school but 54 % of teachers needed no more resources than normal; 35% thought more resources were required and 12 % used far more resources than normal. Organisation and teaching style may also have required modification to accommodate a task of this nature and indeed 46% felt that this was the case; 23% reported no change and 31% only minor changes. Some of the 46% reporting changes commented that the long task had forced them to return to traditional practices which had been abandoned in favour of a more integrated approach to the subject; this was fair comment!

Appropriateness of the tasks

Teachers were asked to rate the appropriateness of the task for which they had been responsible. This meant that a specialist food teacher was required to rate the food task in relation to the KS3 programme of study and not necessarily what the pupils had been taught in relation to food.

	food	***construction***	***textiles***	***graphics***	***control***
poor	9%	11%	15%	13%	25%
satisfactory	23%	21%	23%	22%	31%
good	58%	53%	51%	48%	37%
excellent	10%	15%	11%	17%	8%

The control material task was the only one seen by the majority as being inappropriate. Many teachers felt the task was ambiguous and did not place enough emphasis on components. They indicated that in their view control was synonymous with electronics.

Assessment materials

Teachers were also asked to comment on the information which was provided to help them manage the activity and the actual structure of the task.

Did the long task guide (teacher material) provide you with sufficient information?

poor	***satisfactory***	***good***	***excellent***
11%	33%	51%	5%

Was the long task framework (teacher/pupil material) easy to follow?

poor	***satisfactory***	***good***	***excellent***
10%	27%	53%	10%

Teachers views were also sought on the activity card which set out the theme and context for the task; this card also provided pupils, by tier, with a differentiated structure.

Did the activity card describe the task and how to tackle it clearly for the pupils?

tier	***poor***	***satisfactory***	***good***	***excellent***
levels 1-4	16%	31%	40%	14%
levels 3-7	10%	28%	50%	12%
levels 6-10	12%	27%	46%	15%

The structure was provided for two key reasons: to help pupils plan their projects and to make them aware of what criteria would be used to asses their work. Teachers were asked to comment on both these aspects. These structures were drawn entirely from the programmes of study so it was important to discover if pupils understood the language and terminology used.

Did the differentiated structures help pupils plan their projects?

tier	***poor***	***satisfactory***	***good***	***excellent***
levels 1-4	30%	39%	26%	5%
levels 3-7	21%	41%	35%	3%
levels 6-10	28%	33%	35%	5%

Did the differentiated structure help pupils understand what was appropriate to achieve required levels?

tier	***poor***	***satisfactory***	***good***	***excellent***
levels 1-4	39%	41%	18%	3%
levels 3-7	33%	41%	24%	3%
levels 6-10	36%	36%	23%	5%

Did the language and terminology of the structures pose particular problems?

tier	poor	satisfactory	good	excellent
levels 1-4	35%	42%	18%	5%
levels 3-7	15%	48%	33%	4%
levels 6-10	19%	45%	31%	5%

Although the majority rated these aspects at least satisfactory, the high percentage in the poor category was indicative of the formal nature of these materials. It is not surprising that pupils entered in the 1-4 band, found formal examination style materials unhelpful and difficult to understand. Such pupils are not used to this style of working and indeed it is inappropriate to their educational needs. This was one of the problems posed by an entitlement curriculum which included standardising testing as an essential component. It is evident that a style of working which requires systematic responses to detailed specifications and structures was one to which pupils of this age were not accustomed.

The prime aim of the task was to determine pupils' capability in relation to the two attainment targets being assessed, so teachers were asked:

In the context of NC design and technology, did this model of a long task offer pupils a reasonable opportunity to demonstrate their capability?

attainment target	poor	satisfactory	good	excellent
Te2	9%	28%	47%	16%
Te3	15%	26%	42%	17%

If your pupils had followed the whole KS3 programme of study, would they have found the tasks a fair challenge?

poor	satisfactory	good	excellent
5%	16%	51%	29%

These two responses indicate that at least two-thirds of teachers thought the tasks were reasonable. The fact that this increased to 80% if pupils had followed the complete programme of study, was a good indication that this model might be acceptable as the basis for the first statutory assessment in a year's time.

Teacher evaluation - the tests

The manageability of the tests - marking

The marking of National Curriculum tests was becoming a highly contentious issue for all teachers. Perhaps in design and technology there were even more reasons for concern as pupils of this age typically would not have been set pencil and paper tests. Consequently teachers were being asked to undertake what was very definitely an addition to their normal workload. Attempting to make this task

unambiguous and straightforward was essential. The subject background of teachers marking the test was as follows: art and design - 9%; CDT - 53%; HE - 26%; business studies - 3%; textiles - 5% and others 3%. This is consistent with the distribution of those supervising the long task, except for the significant decrease in textiles teachers marking the test. There is no valid reason why their participation dropped by 50%. Marking statistics are provided below.

		level 1-4	level 3-7	level 6-10
average number	***all tests***	***test 1***	***test 2***	***test 3***
of papers marked	25	6	18	5
average reported marking time	390 mins.	13mins.	17mins.	22mins.

Did you find the layout of the marking scheme straightforward?

	poor	***satisfactory***	***good***	***excellent***
% of teachers	13%	23%	50%	14%

How would you rate the ease of interpretation of the mark scheme?

	poor	***satisfactory***	***good***	***excellent***
% of teachers	31%	41%	26%	2%

How closely did pupils' answers correspond to the 'what to look' for column?

	poor	***satisfactory***	***good***	***excellent***
% of teachers	15%	67%	18%	0%

Teachers found the mark scheme difficult to interpret and identified very little correspondence between pupils' answers and the examples in the mark scheme. When writing questions and mark schemes based on process driven statements it is difficult to write questions which have right or wrong answers and therefore unambiguous marking criteria. Whilst the statements which constituted Te 1 and Te 4 are used as the basis of assessment criteria this will always present a difficulty. It was apparent from a review of test papers that some teachers interpreted the examples as the only right answer. Constructing an unambiguous mark scheme for the aspect of the test in which pupils reflect on a practical activity presents serious difficulties. Tests are a device which are more efficient at assessing knowledge rather than experience(s). As pupils had also not been taught many of the aspects of the programme of study which were being tested, teachers were unfamiliar with the content of the marking criteria and this hindered interpretation.

An assessment procedure which itemises aspects of capability in such a discrete way will encourage a similar approach in those carrying out the assessment, At an INSET session a group of teachers discussed for 15 minutes the correct material

for making a vacuum forming mould and still did not reach agreement. If teachers found it difficult to agree on an aspect which is relatively unambiguous and straightforward it is obvious that the questions which dealt with decision making and the reasons for taking decisions are certain to be more controversial.

Although teachers experienced difficulty with the mark scheme two thirds felt the questions closely interrogated the statements of attainment and only 5% could see no connection. 51% thought the questions presented a sufficient incline of difficulty. Perhaps the high percentage who felt the incline was either too sharp, insufficient or confused were also identifying a basic problem with the statements of attainment.

Relationship of the test to the long task

The first part of the task asked pupils, via a series of questions linked to statements of attainment to reflect on the first practical task which they had already completed. The second half asked them to develop the theme into a new context. This restricted outcomes to ones in textiles and plastic; teachers were asked:

Were the tests fair regardless of the long task taken by the pupil?

unfair			***fair***
28%	27%	36%	10%

55% of teachers felt the tests favoured pupils who took certain of the long tasks. An analysis by subject background showed that this included all the Home Economics teaches. They felt the failure to ask questions about food issues in the last part of the test would disadvantage pupils who took this long task. An analysis of the levels obtained by these pupils proved that this was not the case. Teachers may have failed to recognise that pupils would draw on a much broader experience of design and technology than just the long task when answering the test. It was, of cause, a procedure designed to assess pupils' capability at the end of a key stage. Teachers' lack of familiarity with the programme of study was also apparent, in their response to these two questions:

Would a paper of this nature present a fair challenge to pupils at the end of key stage 3, if the programmes of study had been followed?

unfair			***fair***
18%	26%	40%	16%

How well did the tests reflect the aspect of the programme of study identified for the long task?

unfair			***fair***
25%	39%	39%	7%

Pupils' response to the questions

On the basis of marking the test paper, teachers were asked to identify the questions which they felt pupils had problems understanding. Their response is shown below by test and question level

que. level	*1*	*1*	*1*	*1*	*2*	*2*	*2*	*2*
test 1	4%	1%	6%	6%	6%	6%	6%	6%
que. level	*3*	*3*	*3*	*3*	*4*	*4*	*4*	*4*
test 1	9%	19%	5%	16%	22%	24%	49%	13%
test 2	9%	10%	3%	10%	13%	17%	56%	5%
que. level	*5*	*5*	*5*	*5*	*6*	*6*	*6*	*6*
test 2	32%	28%	15%	21%	11%	15%	16%	18%
test 3					2%	4%	3%	6%
que. level	*7*	*7*	*7*	*7*	*8*	*8*	*8*	*8*
test 2	40%	15%	21%	39%				
test 3	16%	6%	8%	13%	10%	10%	24%	9%
que. level	*9*	*9*	*9*	*9*	*10*	*10*	*10*	*10*
test 3	9%	14%	18%	19%	10%	14%	12%	17%

Seven questions can be identified as having caused pupils particular problems.

level 4	Te 1.4a	Te 1.4b
level 5	Te 4.5a	Te 1.5b
level 7	Te 4.7a	Te 4.7a
level 8	Te 1.8b	

In most cases questions which appeared on two tests were rated quite differently. For example the two questions which related to Te 4.7a were rated on test 2 at 40% and 39% but the ratings for the same question on test 3 were 16% and 13%. The exception is the question relating to Te 1.4b which teachers believe created genuine problems for pupils regardless of the test. This was confirmed by pupils performance, only 15% correctly answered the question on test 1 and 27% on test 2 (% of correctness expected for a level 7 question).

Finally, the average reported time spent by teachers marking these test papers was 6hrs. 30mins and the maximum reported time was 21hrs 45mins. However if the average marking times and the average number marked are used to calculate the time spent by each teacher, a time of 8hrs. 14mins. is arrived at. It is unlikely that teachers reported the total time incorrectly so possibly the average times given to mark the test papers were exaggerated by about one-third.

Conclusion

When introducing a novel form of school assessment, two extremes of approach could be adopted in relation to teachers. At one extreme, teachers could be regarded merely as operatives, invigilators charged only with the responsibility of administering the test in a standard and prescribed fashion. Diametrically opposed to this, would be a system in which teachers were fully involved in interpreting the context of the assessment task, were responsible for resourcing all aspects of the task and where their objective observations formed a significant element of the assessment process. Statutory National Curriculum assessment in design and technology gradually regressed from the latter approach, in which the teacher was central to the process, to the former one in which they were becoming little more than bystanders. Indeed, if the subject had not been deregulated as part of the Dearing review[1], it would have been subjected to the regulations now applied to the core subjects. It is difficult to envisage how these could have been imposed on a subject in which practical activity is such an integral component. The two evaluations reported in this chapter represent the final pilot in which teachers were seen as central to the process (1991) and the first in which external constraints moved the teacher from the role of facilitator towards that of an invigilator (1992).

These evaluations present a picture of how teachers coped with the assessment process. It reveals the very real practical problems which teachers had to resolve in trying to provide pupils with a fair and standard opportunity to evidence their best level of performance. It is also indicative of the difficulties which teachers were having in delivering National Curriculum design and technology. Possibly the trialling or piloting of the assessment task was, for many teachers, their first genuine experience of what this subject involved. Therefore, their increasing frustration with the assessment process was symptomatic of their views of the Order. A crucial change which National Curriculum assessment introduced was the attempt to measure the effectiveness of the teaching to which a pupil had been exposed during a key stage. The final scores would say something about the teaching as well as the pupil. It was therefore essential that the teachers should trust the process and have confidence in it. The more involvement teachers could have in the development process the greater would be the chance of success; the more they felt their views and opinions were being listened to and valued, the more comfortable they would feel with the end product, the reported assessments.

The effect of teachers' evaluations was, however, not always as teachers might have intended. An example is the evidence from the 1991 pilot, that each teacher's subject background and expertise had a significant bearing on their capacity, and possibly willingness, to implement the Order and its statutory assessment

procedures. The management and organisation of teachers were probably therefore influences on pupil performance. Teacher quality is a variant which is unavoidable and a factor which statutory assessment was intended to identify. But if teachers were being required, for example, to operate in areas outside their expertise, as in the 1991 pilot, and as a consequence pupil performance was suffering there could have been significant repercussions. Performance issues which were clearly related to such unequal resourcing could well have been politically embarassing, had they occurred during a statutory assessment. Evidence that this was the case in the pilot, enhanced the argument for standard restrictive tasks undertaken in specialist areas with specialist teachers, as in the 1992 pilot.

The 1991 evaluations had resulted in the assessment procedure reinterpreting the Order in a reactionary fashion, in opposition to the views of many teachers. Teachers' judgements and opinions on the piloting of a statutory assessment procedure should have been invaluable to the development process. However, because of the way in which some of these findings were interpreted and the intentions behind some of the teachers' comments this was rarely the case. Difficulties in managing the task and its assessment could have been seen as a need for additional INSET but the opposite view was adopted and the procedure was, incrementally, made less dependent on teachers. Whatever a teacher's standpoint, the confusion which the continual changes created further exacerbated the climate of conflict which became ever more stormy as the first statutory assessment approached.

Reference

1. Dearing, R. (1993a). *The National Curriculum and its assessment: Interim Report*. London: National Curriculum Council/Schools Examination and Assessment Council

Chapter 7

Evaluation and conclusion

Synopsis

In this final chapter the four issues and questions established in chapter one are reviewed in the context of what has occurred since this research commenced. A number of key documents have been published since the Spring of 1993 when the first statutory assessment of design and technology started: The Dearing review of the National Curriculum, the HMI review followed by the joint NCC and SEAC review of the design and technology component of technology, and finally the SCAA consultation document and the revised Order. These have significantly changed the context and immediate relevance of this research but it is argued that the research still has considerable value and relevance to those concerned with statutory assessment and design and technology in particular.

Following a discursive appraisal of the four key issues eight issues are isolated and conclusions drawn concerning these particular aspects.

The chapter is divided into the following sections and sub-sections

Reaching conclusions

- The climate in which the research was undertaken
- Issues of validity and reliability
- Issues relating to criterion-referencing
- The impact of assessment activities on learning experiences
- The relationship between different modes of assessment

Final conclusions

- The role of assessment in changing the curriculum
- The style of assessment device
- The nature of assessment criteria
- The interpretation of assessment criteria
- The nature of practical tasks
- The standardisation of pupil material
- The nature of tests
- The assessment process

Reaching conclusions

The climate in which the research was undertaken

The period covered by this thesis, September 1989 to April 1993, represents forty four of the most turbulent months in education in this country. The Education Reform Act, the most substantial piece of education legislation for half-a-century set out radically to change education for five to sixteen year olds. This was to be achieved by imposing a statutory curriculum on state schools. This approach to reform had been emerging for the previous two decades, as detailed in chapter 1, and a consensus had developed that this was the right way forward. That this imposed curriculum might be enforced by a statutory system of assessment procedures had, however, not been made apparent until the final legislation began to take shape. The blueprint for the assessment procedure was provided by the Task Group on Assessment and Testing. Initially these proposals found favour with educationalists but politicians were less enthusiastic. Politicians and civil servants allowed TGAT's credibility to be undermined by teachers, principally those conducting statutory assessment tasks at key stages 1 and 3, before attempting to impose a more rigid assessment framework.

Throughout the eighties politicians of all parties became increasingly convinced that the state education system was failing the pupils, their parents and the nation at large. Even educationalists agreed that a reappraisal of aspects of the system was long overdue. However, they would not have agreed on this solution: in effect imposing compliance to a statutory curriculum via a regime of testing. Testing is often the policy makers' solution to real educational problems as it appears to measure performance accurately, identifying strengths and weaknesses. As an approach to quality assurance, testing lacks subtlety and all too frequently it does not result in the anticipated effective and permanent change. As has been shown in this thesis (see chapter 3) the adoption of a testing strategy created an additional range of problems, ones which did nothing to improve the quality of education and which are still being resolved. The political requirement was, and still is, to force the system to be accountable. Testing, it is argued, would produce scores which, following aggregation, would result in simple numbers which could be subjected to statistical analysis, allowing goals to be set and performance monitored. Clearly, there was a possibility that the assessment process might have become so powerful that success in the tests could have become the sole objective of teaching and learning.

During these four years, assessment was at the very heart of the turbulence. Others, such as Daugherty (1994)[1], provide general accounts of the way in which policy developed, but this thesis provides detailed evidence of this process from a research and development perspective. It does so through the dimension of one

subject, design and technology, and at one key stage, key stage 3. Thus it provides a unique insight into the effect of policy on practice. The research has generated a rich resource of information, especially concerning the development of appropriate research tools, the results of the application of those tools and the views of teachers and others on their effectiveness. The research was, by necessity, both proactive and reactive. The development of these research tools led the subject's development during the period in different ways. During the first phase, 1989 -1991, it provided teachers with a model of what National Curriculum design and technology meant in curriculum terms. Whilst in the second phase, 1991 -1993, the development had to react to the more traditional view of assessment which was by then setting the agenda.

The data gathered during this period are unique because the National Curriculum has now been revised both in format and content. Design and technology, of all the subjects in the curriculum, was in Graham's opinion (1993):
'The one true revolutionary subject to enter the National Curriculum.'[2]
The development of novel assessment procedures for what many claimed to be a 'new' subject was worthy of close scrutiny. Much can be learnt for the future concerning implementation strategies, subject evolution, and assessment procedures. This thesis set out, in chapter 1, to examine four issues and the questions which they raised. The research undertaken must now be evaluated to determine what conclusions can be reached in relation to these issues. This thesis did not attempt to determine whether there was a need for statutory assessment and what value these assessments would have. It is concerned not with why but how. Currently, at key stage 3 only the core subjects, English, maths and science, are the subject of statutory tests but at sometime in the future it is probable that some or all of the foundation subjects may be exposed, in a statutory way, to assessment scrutiny. This thesis provides evidence which could be used to ensure that if and when this occurs, it is accomplished in a way which is more beneficial to all concerned.

Issues of validity and reliability

The first issue was:
Which aspects of the 1989 Statutory Subject Order in Technology, relating to design and technology, could be assessed in a valid and reliable way by a national statutory assessment procedure?

Central to this question is the term 'assessed' and any response will be determined by the character of the assessment employed. Assessment is not synonymous with testing. If it were this question could be dealt with in a relatively simple way. But much has changed since this question was first framed and, in the context of a statutory framework, assessment has for the present been replaced by testing.

The nature of assessment

The 1990 Statutory Order defined design and technology as a process; as such it became a multidimensional discipline in which all aspects were interrelated. The subject profile component was termed design and technology capability. The term capability was used to emphasise the active nature of the subject. Capability was described, in the Interim Report of the National Curriculum Working group and stated as the overall objective of the subject in the final proposals, as:
'to operate effectively and creatively in the made world.'[3]
This is what the interim report concluded pupils gained from design and technology activity which could be learnt in no other way. The report also emphasised that:
'design and technology is always purposeful (i.e. developed in response to perceived needs or opportunities, as opposed to being undertaken for its own sake), takes place within a context of specific constraints (e.g. deadlines, cash limits, ergonomic and environmental requirements as opposed to unconstrained, blue-sky research) and depends upon value judgements at almost every stage.'[4]
These definitions were rarely referred to and, like the two documents from which they were taken, quickly disappeared following the publication of the statutory Order. The statutory Order contained only the legal framework of the subject, consequently the supporting arguments for the inclusion of many aspects in the Order were consigned to history.

Capability was used because, in a succinct and elegant fashion, it exemplified the objectives of the subject. If this was the accepted objective of the subject then it can be powerfully argued that any assessment procedure should concord with this approach. As described in chapter 3, National Curriculum assessment was driven by the developments in mathematics and science. These subjects were constrained to testing only the knowledge component, an approach which was from the outset at odds with the notion of design and technology capability. Assessment models more sensitive to 'purposeful activity' were being developed elsewhere during this period, in particular in relation to vocational qualifications. This was extremely relevant to design and technology, as it had been identified by ministers as being one of the key subjects which might provide a vehicle for the introduction of vocational qualifications into the secondary sector. From 1986 onwards a new national system for the assessment of occupational competence had been devised and introduced. National Vocational Qualifications (NVQs) are concerned with what an individual can do in the workplace. Competence is defined in specific terms:
'It is a description of an action, behaviour or outcome which the person should be able to demonstrate and it must be assessable.'[5]

Assessment for an NVQ involves the collection and evaluation of evidence against

performance criteria. An individual is required to demonstrate that his or her performance meets the prespecified standard. The notion of 'performance' is crucial to competence. It is not discussing or describing how something might be done or the merits of doing it one way rather than another, it is actually doing it! GNVQs (General National Vocational Qualifications) aim to bridge the academic/ vocational divide. These qualifications do not confirm occupational or professional competence. They are based, like the National Curriculum, on statements of attainment rather than statements of competence. NVQ assessment tasks must be undertaken in the workplace but this is not the case for GNVQs. In both the assessment of learning outcomes must be appropriate and reflect the purpose of the course or programme of study.

This approach provides support for the notion that assessment should be in accord with the character of the subject, in the case of design and technology - capability. Performance is as crucial to competence as it is to capability. The term 'authentic assessment' has been used by amongst others, Goldstein (1992)[6] to describe an interactive model of assessment which sees it as part of the learning procedure. Gipps (1994) confirms that the intention behind the TGAT report was that Standard Assessment tasks should be:

'... good examples of performance assessment. Performance assessments demand that the assessment tasks themselves are real examples of the skill or learning goals, rather than proxies.'[7]

As Resnick and Resnick (1992) observed,

'We cannot teach a skill component in one setting and expect it to be applied automatically in a context very different from the context from which it is practised or used'[8]

What did 'authentic assessment' mean in the context of the profile component design and technology capability? Clearly any assessment task needed to be a purposeful activity undertaken in response to perceived needs or opportunities, within a context of specific constraints. If validity is simply explained as the extent to which an assessment measures what it purports to measure then the 'authentic assessment' described above would have undoubtedly provided a valid assessment of the design and technology Order.

The assessment tasks, described in detail in chapter 4, covered a range of approaches. Some would fit the notion of 'authentic', others clearly would not. It will be evident to the reader of this thesis that the change in specification in 1991 can be seen as the point in time when the development of statutory assessment departed from the objectives of the TGAT report. The assessment instruments developed prior to that date all complied with the requirement of the previous paragraph. Following 1991, the assessment process was divided into two

components, practical tasks and writtten tests. The validity of such an approach is questionable. If the view is taken that the four attainment targets are interactive and interdependent, then any device which attempted to assess them individually or in pairs clearly is not. The design and make task used as the vehicle to assess Te2 and Te3, in 1992 and 1993, had undoubtedly a higher level of validity than the tests used to assess Te1 and Te4 during the same period. The various practical tasks were purposeful and undertaken within a context of specific constraints, they were not though in response to perceived needs or opportunities, an aspect dealt with by Te1. It is difficult to justify the validity of tests in relation to design and technology, given the rationale on which the Order was based. The model adopted for the tests viewed validity at the level of each independent statement, a concept clearly at odds with the notion of capability.

Reliability is most commonly associated with marking and tests can be designed which are extremely reliable in this regard. Such tests frequently offer a valid approach to examining independent items, such as facts or concepts. Ones which typically, in the context of a question, require a prescribed correct response. For example, multiple choice tests are highly reliable measurement instruments. It is also unlikely that a test, in design and technology, would be undertaken in an 'authentic' context. For example, normally tests take place in a neutral environment, one in which no external resource is allowed when responding to the test questions. Such conditions are hostile to the notion of designerly activity. Tests are generally controlled by regulations in relation to time and resources. Applying such restrictions to attainment target 1 - identifying needs and opportunities, could only produce an artificial and spurious assessment. This approach could not respect the integrity of design and technology capability. Design and technology was not defined in relation to factual knowledge or conceptual grasp in isolation, as these are passive. It is an individual's capacity to use their understanding of knowledge and command of skills to tackle and solve a task which is fundamental to the subject.
'The designer does not need to know all about everything so much as to know what to find out, what form the knowledge should take, and what depth of knowledge is required for a particular purpose.' [9]
An assessment task should provide opportunities for these key skills to be evidenced. What is seen as a valid approach to certain subjects should not be imposed on another for reasons of bureaucratic consistency and neatness.

Assessment is however, not synonymous with testing. In deciding on the characteristics of any assessment to be employed, one must understand fully its purpose. TGAT proposed a model in which SAT assessments would satisfy a variety of intentions. For example, be both formative and summative; formative to

support learning and summative to enable effective monitoring. Much has been made of this distinction in trying to establish a model of National Curriculum assessment for the future. Sir Ron Dearing in his Interim Report (1993) of the National Curriculum review concluded:

'There should only be a diagnostic element in such tests (national statutory tests) if this can be shown to be a cost-effective supplement to the primary summative purposes of the test.'[10]

The summative purpose of end of key stage assessment is now accepted as the primary purpose. Tests in the core subjects now have this principal aim, unlike those developed in the context of this research which had five objectives (see Chapter 4, page 100). These objectives attempted to make these assessments serve a number of purposes and consequently increase their usefulness.

Assessment devices

The first four years of the National Curriculum can, with the benefit of hindsight, be viewed as an experiment on a national scale. This experiment clearly established a better way forward for the future. A great deal was learnt about the limitations of what could be achieved by a statutory curriculum and how it should be formulated and structured. In design and technology the lessons were more fundamental than in other subjects. The uses of the subject's Order, in relation to assessment, had not been explicitly stated to those who first formulated it. In particular the subsequent legal requirement to adhere slavishly to the statements of attainment as assessment criteria was probably not fully understood. Possibly, the structure that had been established was key to this failing. The notion of two organisations, one responsible for delivering the curriculum and the other for assessing it seemed perfectly logical. However, the definition of the curriculum was the responsibility of the deliverer, not the assessor. The speed of implementation did not allow for small scale research to be undertaken on the efficiency of the statements of attainment as assessment criteria. If this had occurred perhaps some simple, yet key, changes might have rectified the situation. For example, reducing the assessment items in each statement. Within a few months of the commencement of this research an analysis of the statements had revealed significant issues which would impair their use as assessment criteria - see page 100 -102. This research adressed these issues by investigating a number of approaches which would diminish their effect. They involved:

- stranding statements within attainment targets;
- collapsing statements to produce a level statement;
- simplifying the statements to identify a simple progression statement.

The parallel development team at Goldsmiths' College investigated the hierarchy of the statements at each level. For each approach justifications had to be articulated which demonstrated its legality, see appendix 4.6, page 297.

The commonsense view underpinning these approaches was that reliable assessments would only be achieved if the procedure was straightforward and robust. In the context of National Curriculum assessment this meant up to 40,000 teachers would have been involved in delivery, recording, assessment and moderating in a standard way. The reliability of an assessment is the extent to which an assessment would produce similar scores on more than one occasion when given by more than one assessor. This is obviously dependent on both the assessment instrument and the assessment criteria. The premise adopted in this research was that if the final required result was a single subject score then the closer the assessment decisions were to the reported score the more accurate they would be. Fewer points of aggregation would result in a simpler process which would decrease the wastage of information and effort. Attainment targets were designed to be the pillars of each subject and as such, reliable scores at this level were crucial to statutory assessment. The 'levelness' approach to assessment is described in detail on pages 106 -109 and appendix 4.6, page 247. This holistic approach to assessment is similar to that advocated by the Assessment of Performance Unit's report on design and technology, yet it is not identical. The attainment targets define process domains within the subject. This is seen as the level at which assessments are most reliable rather than at the level of the whole subject. Initial, overall impressions can assist assessment but they can also subvert accuracy in relation to the component parts. Which is of key importance in a national statutory context? Undoubtedly the robustness of the single subject score. The argument made by this research is that this score will be most reliable if arrived at from valid attainment target scores rather than from single assessment items, statements of attainment, or from a single judgement of capability.

By the time of the first statutory assessment this argument had been won in relation to the practical task, the task which would assess Te 2 and 3. For the first time in statutory assessment the statement of attainment was not the judgemental focus point. Teachers were asked to determine which level statement best described the pupils' performance or capability - a 'best-fit' approach. Evidence, such as that in appendices 4.5, page 291 and 4.6, page 297, indicates that making level judgements was a manageable task and had the potential to become reliable and valid. To achieve reliability and validity is a lengthy process as it requires the evolution of a professional consensus of what represents performance at each level. Such an evolution is rarely the result of legislation. Exemplification can assist, but interaction, discussion and debate are essential if teachers are to have a shared understanding of levelness, one which they have confidence in applying in a statutory context. The need to simplify the process was identified by all those concerned in revising and reforming the National Curriculum. In a rare example, assessment in design and technology indicated the way ahead and the revision of all subjects in 1994 adopted this approach based on level descriptors.

The answer to this first question concerning validity and reliability (page 219), depends wholly on definitions and circumstances. Teacher hostility, union militancy, media exposure and political bigotry created a context in which failure was probably inevitable. Teachers' realisation that the assessment was clearly focused on them and their schools inevitably led to noncompliance and a great deal of rancour. Union concerns to protect the working conditions of their members took the conflict out of the classroom and into the lawcourts. The media's legitimate duty to interpret the debate by simplifying the issues, hardened the resolve of the protagonists and polarised the arguments. In the face of this onslaught a beleaguered government became dogmatic and authoritarian in trying to impose its will on a profession which, unusually, appeared to have the weight of public opinion on its side.

What system could have been assessed successfully in this context? What changes to policy could have resulted in the successful introduction of an assessment system based on the 1990 Statutory Order? The first key change would have been the adoption of a subject focused approach rather than a system based approach. Each subject should have had the opportunity to devise a system which was in harmony with and sympathetic to existing practice. In this context, solutions could have been developed more quickly and more openly. Changes of practice are extremely difficult to achieve. The use of new terms does not necessarily bring with them changed understandings and procedures. The term 'programme of study' was in teachers' minds synonymous with a syllabus. Key stage 3 teachers are familiar with GCSE syllabuses which are always published with sample papers. It is, for many teachers, the specimen papers which define and interpret the syllabus. Teachers, in coming to terms with the National Curriculum, did not benefit from such an interpretation until weeks before the assessment took place. GCSE examination groups have to develop syllabuses and specimen papers over relatively short periods of time and publish them for schools as a coherent document. The assessment of the National Curriculum would have stood a better chance of success if this approach had been adopted. For many teachers the first statutory assessment was their first view of the papers. If samples had been available when the pupils being assessed had embarked on the key stage a wholly different series of events might have occurred.

Much of the design and technology teachers' antagonism to the assessment tools which emerged resulted from the reinterpretation of the Order in the light of political decisions. If specimen material had been available from the outset this might not have been the case. The political imperative to introduce the Order and operationalise it within the shortest possible period of time eventually led to its failure. The Dearing revision of the National Curriculum stands a better chance of success because it was based, in part, on the evidence of trying to assess the first

Order (evidence resulting from this research and of those concerned with other subjects). The curriculum, its content, assessment objectives and procedures must always be developed and disseminated as a coherent package. Once this package has been agreed, it can be implemented without change before the first assessment procedure. The introduction of the first National Curriculum was for most teachers a tragedy with an incomprehensible plot as even the authors had not determined the major elements until just before the final scene.

The subtext of this first question is really - was the Order to blame? Along with the test developers it certainly became the scapegoat (see chapter 3, page 68 onward). As the basis for a research programme it presented many challenges and, in the form in which it was presented, would have proved extremely problematic to assess nationally. The main impediment was the lack of precision of the language used in the statements of attainment. This could only be identified from trialling these statements of attainment and observing how teachers coped with them. This was recognised by members of the SEAC steering committee:
'the agencies exposed the difficulty of devising standard tasks which would assess adequately some of the general, loosely-worded statements of attainment in those Orders.'[11]
However, later in the development process, when approval was given for the assessment criteria to be interpreted via the marking procedure and the task being undertaken by the pupil, there is some evidence that valid and reliable assessments could have been achieved over a period of time. This allowed the theoretical statements of attainment to be given substance and meaning which could be related directly and unambiguously to pupils' outcomes.

The central objective of the Education Reform Act was to raise the standards of pupil achievement. Assessment was required to monitor these changes. But during 1992 and 93 a widely held view developed that the Order was an obstacle to this prime objective. The NCC, following an internal review, concluded,
'the language of the Order is difficult for non-specialists to understand and the programmes of study lack focus and rigour.'[12]
(Non-specialists is a reference to the key stage 1 and 2 teachers who had to teach design and technology with little subject experience.)
It was for these reasons NCC recommended to the Secretary of State that the Order should be revised. If the language can be simplified so that there is less requirement for 'examiner interpretation' then statutory assessment becomes more likely to be successful. This research has highlighted the need for assessment criteria which have simple constructions, use plain language and describe authentic achievement which, as discussed earlier, can only be evidenced in an authentic context. The authentic achievement of design and technology capability could only be assessed in a valid manner by undertaking design and technology

tasks which involved pupils in evidencing aspects of all four original attainment targets at an appropriate level of achievement.

Issues relating to criterion-referencing

The second question was:
Is it possible to devise and implement assessment procedures, based on criterion-referencing, which will differentiate pupils' achievement across the range of design and technology in a fair yet consistent fashion?

The National Curriculum was the first formal assessment procedure to adopt a truly criterion-referenced approach to the assessment of design and technology. TGAT was explicit in recommending that:
'assessment results should give direct information about pupils' achievement in relation to objectives.' [13]
The commitment to this approach came 25 years after Glaser (1963) had published his seminal paper on this approach to assessment. He defined criterion-referenced assessment as:
'Measures which assess student achievement in terms of a criterion standard thus provide information as to the degree of competence attained by a particular student which is independent of reference to the performance of others.'[14]
The consequence of this approach is that the measurement of learning is described by what the learner can do rather than how well they have performed in relation to others or as a description of the learning input. A bank of criterion referenced statements arranged in levels according to difficulty is in effect equivalent to a desired set of learning outcomes. From the learning perspective, a teacher faced with the task of planning the delivery of a key stage might use selected statements of attainment as the objectives of certain elements of his or her teaching. In addition, statements of attainment are better motivators than syllabuses as they set pupils clear targets; the pupil knows what is being asked of them. This was the aim of the National Curriculum to unite teaching objectives and assessment criteria so that assessment would support learning and provide greater clarity of curriculum definition.

As has been discussed in chapter 2 - pages 45 to 51, traditionally assessment in design and technology has been largely on the basis of outcome alone. Every pupil would be set the same task and marks would be awarded on the basis of the solution produced. More often than not this was the result of the teacher norm-referencing within the group. Teacher judgements were probably highly reliable in the relationship of one individual to another. This approach was not viable and could not be sustained given the introduction of level related statements of attainment. The adoption of the ten level scale of achievement was an ambitious

enterprise. This hierarchical scale established a set of explicit criteria which would define progress for pupils from 5 to 16. It implied that a pupil would, in each attainment target, sequentially progress from one level to the next, as they were systematically being exposed to and taught the demands of each level. However, learning is not that straightforward and teaching is rarely that systematic. It also presupposes that the content of the attainment targets was hierarchical. As design and technology capability defined a process, the assumption was that each level described a more complex and sophisticated activity which required the employment of more demanding skills and greater depth of subject knowledge. The complexity of defining these criteria so that they were applicable to pupils over their eleven years of schooling and could be interpreted consistently by teachers was underestimated. In design and technology the third attempt at this task will become the legal requirement in September 1995. A series of hierarchical levels, as a basis for national calibration, has survived the Dearing Review despite severe opposition from a number of quarters. For example, the Social Market Foundation published two papers which argued the case for removing the scale as it had a detrimental effect on teaching and learning.[15]

The research on which this thesis is based was challenged with the responsibility of creating tasks which would be assessed by these criteria. The subject's generic practical nature, coupled to its multi-disciplined structure made the task uniquely different to those in other subjects. Three types of assessment activities were produced:

- contextual practical tasks assessed by outcome;
- prescribed practical tasks differentiated by task;
- tests.

The construct underpinning each of these and a detailed description of their evolution is provided in chapter 4, the purpose here is to determine how well each achieved the requirement to differentiate pupil achievement in a fair yet consistent fashion.

Contextual practical tasks assessed by outcome

This style of task was used between 1989 and 1991. They were typified by a highly resourced activity based on a context. Each had similar resources, for both teacher and pupil, which provided a standardising framework. In particular every pupil constructed their response around a set of identical labels which operationalised the assessment criteria in a practical pupil-oriented manner. As has been discussed earlier in this chapter, in the context of design and technology capability these tasks were the most valid. The most refined version of this type of task was piloted in the Summer of 1991. The three tasks used in that pilot were the fourth version and they had gradually evolved over a 24 month period with input from a

wide variety of sources (teachers, LEA Advisers, Academics and HMI). They represented a sound curriculum interpretation of the Order offering access to all pupils regardless of the resources and specialist facilities at their disposal. Nearly 10,000 pupils took part in this trial and the mean profile component level was 3.2. This was alarmingly low for pupils in Y9, but there were sound reasons for these performance levels (see page 152 - Commentary on performance). This type of task required pupils to perceive a need or opportunity from one of three specified contexts, consequently, the range of activities undertaken was numerous. So did the context affect performance, would it be fair in statutory assessment to have a range of contexts, or alternatively would it be unfair to only have one? If the mean profile components are considered, the performance on the three tasks was; Exhibition - 3.21, Measurement - 2.95, Public Places - 3.26. The difference between the highest and the lowest is 0.31 of a level, approximately 10%. There is no way of determining if this is an acceptable degree of variation and perhaps it is rather irrelevant. However, it was important that every pupil had the opportunity to achieve their best level of performance, regardless of the context. Correlation with teacher assessment is the best way to assess this aspect, however, teacher assessments were not wholly reliable, as they had been made only for research purposes, and were frequently based on how the pupil had performed in the assessment task.

Were the contexts fair in relation to gender? If the mean profile components are examined by gender the following picture emerges:

	Exhibition	*Measurement*	*Public Places*
boys	2.97	2.57	2.88
girls	3.43	2.95	3.65

Exhibition provides the best mean performance for boys and Public Places for girls. It is worth exploring this aspect as the design and technology Working Group envisaged a bank of tasks which could be used for statutory assessment purposes. If this had occurred, the evidence argues for the provision of more than one context, but perhaps, only if the pupil is allowed to choose. This approach would have raised serious management issues which teachers, during the infancy of the National Curriculum, did not manage to resolve. This prompts another question which this thesis has not sought to answer, what is the effect of the management structure employed by a faculty or department on pupil performance? With time the most efficient would have emerged and become models of good practice or means of achieving good end of key stage assessment performance.

What is remarkable, is the similarity of performance in relation to the attainment targets, regardless of the task or of gender. (This is illustrated graphically on page

156.) This offers further support for the notion of a bank of tasks. Both boys and girls achieved the lowest levels for Measurement, the task with the most demanding technical content. If choice were allowed pupils and schools might opt for contexts which are technically less demanding. This could result in a bias developing away from these aspects of the programme of study. To avoid this these results could be regarded as scores and performance means could then be standardised regardless of the task taken. For example, each task might have an associated difficulty factor or tariff. This procedure might be acceptable for comparing, say school performance, but would make nonsense of reporting an individual's performance. It would however allow a school the opportunity to select the task which fitted their programme rather than selecting the one which 'on the face of it' seemed the easiest. This would also necessitate the disassociation of school accountability from pupil performance at the end of key stage 3. Further research is required into the pupils' perception of the programme of study and how this affects choice in relation to projects and key stage 4 choices.

Prescribed practical tasks differentiated by task;

The introduction of differentiated tasks followed the new specification of September 1991. In the Summer pilot of 1992 the tasks were differentiated in relation to resource context, typically material base, e.g. food, in which the pupil would tackle the task. Structured process diagrams, at three bands of difficulty, also provided pupils with different objectives. This model was further refined for the first statutory assessment of 1993. In each material the task was specified at four levels of complexity. In this model the behavioural criteria, the statements of attainment, were combined with outcome goals derived from the programme of study to produce assessment criteria based both on process and product. These tasks only sought to assess Te 2 and Te3. This approach was not deemed fair by the nation's design and technology teachers as it was only offered in three materials: construction, control and food. This was a bureaucratic decision which, rather unfortunately, provided a focus for criticism and greatly affected teachers' attitude to the whole process. It should not be seen as a commentary on the structure developed for this task (see page 124). Although not the subject of this thesis this model, with some modifications, was employed for the non-statutory tasks of 1994. The introduction of two other materials, graphics media and textiles, overcame this criticism and the tasks were generally well received.

In 1991 the provision of five tasks based on specialist facilities produced mean profile components which ranged from 4.1 (food and textiles) to 3.7 (construction and control). The aspects of the subject traditionally associated with girls produced the highest levels of attainment and those associated with boys the lowest. The mean profile component for all pupils for all tasks was 3.92. This demonstrated an

overall increase of 0.7 of a level from 1991. The variation between highest and lowest is 0.4 of a level, approximately 10% as in the previous year. This performance does clearly illustrate that girls outperformed boys; this occurred in all trials and pilots of practical tasks undertaken during the period of this research. This is confirmed by noting that of all pupils taking food and textiles 65% were girls, whilst of all those taking construction materials and control 58% were boys. There was further evidence which identified a possible area of unfairness. Teachers reported that pupils required significantly more time to complete tasks in construction materials than in the other areas. This also affected control, as this required pupils to use similar materials in making a solution. If tasks are set in relation to different materials then there is a sound case for establishing a different rubric for each task. There is also considerable data on the relative performance of the genders. From the evidence presented throughout this thesis it would be sensible to conclude that girls outperform boys in all aspects of design and technology capability. However, it is possible that the emphasis placed on pupils to record their progress discriminates unfairly as it rewards those who are conscientious in this respect. There was substantial evidence that girls placed much greater value on this aspect of their work and were consequently rewarded. Further research is required into this issue,

The task differentiated tests produced for use in the first statutory assessment in 1993 were trialled, due to the need for confidentiality, in a limited way in the Autumn of 1992. The trial required teachers to enter pupils by tier, four different tiers, and task, three different tasks. The mean level for the attainment targets being assessed was Te2 - 3.7 and Te3 - 3.7. This was the first time that Te3 had not produced the highest performance figures, but it was the first time pupils had to undertake tasks which had to meet specified performance criteria. Performance varied between the tasks, more so than in previous trials. Teacher interpretations of the Control task caused significant underachievement in this respect (see page 165), whilst performance in the food task in respect of Te3, was far higher than in the other two tasks. Teachers did not relish the job of entering pupils in particular tiers, they believed that their decision was placing a ceiling on each pupil's possible level of achievement. Whilst acknowledging that at the end of key stage 3 the levels of pupil performance could vary considerably, teachers did not wish to bear the responsibility of deciding which of the differentiated tasks would provide the appropriate challenge for each pupil.

This highlights a key issue which this research was not designed to address and which should be the focus of research in the future: there was no reliable recorded level for the pupils being assessed as teachers neither had the confidence, nor were the systems in place, to produce consistent and accurate teacher

assessments. If teachers had this information entries could have been made with more certainty. Differentiated tasks, by their nature, will discriminate pupils' achievement but they rely on a fair and reliable entry procedure - a procedure which lies in the province of the school not of those devising the assessment process.

Tests

Tests were introduced in 1992 to assess attainment in relation to Te1 and Te4. The statements in these two attainment targets had not been written with the intention that they would form the basis of a test. Strategies had to be devised which would enable this to be achieved. These evolved from questions which required the assessor to judge whether a pupil's response satisfied a statement to questions in the 1992 pilot, to ones which were marked for the pre-statutory trial and the statutory assessment of 1993. In the first instance various aggregation rules could be applied to decide on the basis of the statements satisfied what level had been achieved. In the second, the marks obtained at each level were totalled, and the highest level at which these equalled or surpassed a pre-established mastery level was the recorded level of achievement. In 1992 all pupils entered at the same tier took the same test but in 1993 pupils took tests which were linked to the practical task taken.

The shift from asking the examiner to make a single judgement in relation to a criterion, towards a marking system was evidence that criterion referencing was not easily applied to the construction of traditional paper and pencil tests and the way in which they were marked. Even if questions could be set against the criterion, teachers did not feel a marking system which only allowed a question to be marked right or wrong was fair. Marking allowed some reward for answers which were not wholly accurate. This though indicated, that in the context of the test, the pupil had not fully evidenced the statement being assessed. Once the notion of marking and levels of mastery was introduced the system was also open to calibration. As Angoff (1974) noted quite correctly:
'if you scratch a criterion-referenced interpretation, you will very likely find a norm-referenced set of assumptions underneath.' [16]
Simply by making comparisons to any criterion-referenced assessment one is making a norm-referenced interpretation. For example, TGAT indicated that pupils should typically be achieving between level 5 and 6 by the end of key stage 3. If tests had mastery levels, the level could easily be adjusted to ensure that nationally this norm was achieved.

Did the tests differentiate pupil achievement and if so was this achievement related to design and technology or was it dependent, for example, on skills

relating to comprehension and expression? As argued in chapter 4, possibly the best answer is found by examining the correct response rate to questions. The performance graphs (see pages 168 -169) show clearly that response rate declined as questions became more difficult; 73% successfully answered level 1 questions whilst only 26 % were successful at level 10. Surprisingly, the test scores increased pupils' performance by 0.42 of a level, even though these two attainment targets only accounted for 35% of the weighted subject score. An analysis of performance by practical task taken also revealed that the test was fair regardless of the task taken.

The tests trialled in 1993 adopted a marking procedure, eight marks being available at each level. A key task in the trial was to determine at what level the mastery should be set for each attainment target. There was no reason why mastery levels should have been the same for both attainment targets or even for the section related to each practical task. It would though have been exceptionally difficult to have convinced teachers that different levels were fair. It can be seen from the analysis of this trial, see pages 175 - 183, that response rate to the questions differentiated achievement quite perceptibly. Is it possible to determine fairness when pupils were taking completely different tasks and tests? Was anyone in a position to decide if the assessment procedure relating to food was as difficult or easy as the construction or control material tasks? If fairness is equated with similar performance distribution possibly the only option is to calibrate, by adjusting the mastery levels, following marking. But these levels had to be decided long in advance of even the tests being taken. Such an approach would also have resulted in a truly examination context for what started out as classroom tasks. Deciding on a mastery level prior to the tests being taken illustrates the dilemmas this type of assessment posed. Compared to GCSE, which some claim is a criterion referenced examination, where calibration and grades decisions are taken post testing and marking, the situation is very different. At GCSE grade boundaries are set at different points for different examinations depending on pupil performance.

Were the tests fair to pupils? Initially, for the first cohort, it is unlikely that they could have been. For many the first test of this nature which they would have taken would have been the statutory test; any argument which claimed that inexperience created a fair context would be dismissing each pupil's right to demonstrate their best achievement. Once again, as in the response to the first question posed by this thesis, timescales proved to be the greatest obstacle to fairness. If pupils and teachers had been aware of the assessment process when they embarked on the key stage then tests of this nature would have had a greater degree of legitimacy. Each of these three assessment modes has the same objective, to produce a subject score. From a political perspective, all that was required was a number for each pupil which could be aggregated to produce a school, local education

authority or national average. This performance score could then be monitored on an annual basis and improvement detected - comparative, norm-referenced judgements. What was required was a summative score but naturally many might, mistakenly, interpret the score in relation to the assessment criteria. Just because the initial assessments were criterion-referenced, it cannot be assumed that post-hoc generalisations about the skills and knowledge mastered by a pupil achieving a certain level would be reliable. For example, does a pupil who has obtained a profile component of level 5 understand all aspects of the programme of study at that level and has he or she achieved all the statements of attainment in each of the attainment targets up to this level? Such conclusions should be treated extremely cautiously, as the subject score is far removed from the assessment decision, especially if the criteria have been subsumed into a scheme for a test which has been marked.

It has been demonstrated that each of these procedures does discriminate achievement. Whether these procedures were or would have become fair is more difficult to determine. If pupils had secure and reliable teacher assessments, performance correlations would have provided definitive conclusions. Where comparisons can be made, for example, 1990 trials, 1991 pilot and 1992 pilot, mean aggregated performances are very similar, but analysis has not been undertaken at the level of the pupil. For pupils to demonstrate their best achievement the evidence indicates that there should be choice. This would need to be in relation to the context or the material base depending on the nature of the task.

Although hypothetical, because both the National Curriculum has been revised and design and technology is no longer subject to statutory assessment, which of these procedures might have provided a model for the future? Tasks initiated via a context were deemed to be the most appropriate way of assessing Te1 but as this imposed significant demands on management it is unlikely that even with time and experience it would have been acceptable within a statutory regime. The style of task trialled and piloted in 1990 and 1991 does though provide a good model for teacher assessment. Level differentiated practical tasks available in an appropriate range of materials do have the potential to assess Te2 and Te3 in a statutory framework. This style of task could also have included Te4, which would have provided a fairer and more relevant assessment of this attainment target. A short paper and pencil test of Te1 which was common to all pupils would have completed the testing process. With weighted attainment targets 85% of the assessment would come from the practical task and 15% from the test. This would reflect accurately the very practical approach adopted by the majority of key stage 3 teachers to the teaching of the subject. Any increase of the weighting towards the

test would distort the subject for pupils of this age. Post the Dearing revision, the differentiated task model is ideal for the two attainment target model which has now become statutory.

The revision, despite opposition, has maintained the criterion-referenced approach but with the emphasis moving from the attainment targets to the programme of study. In time it is possible that teachers will appreciate and recognise the value of criterion referencing. The setting of performance goals is of value to the teacher in motivating pupils and recognising and rewarding achievement when it occurs. But there must also be recognition that this has implications for a formal assessment context in which reward is not always possible for partial achievement.

The impact of assessment activities on learning experiences

The third question was:

- can an assessment procedure be devised which will meet the political aims of National Curriculum assessment whilst providing pupils with valid learning experiences and pupils, teachers and parents with accurate and useful measurements of performance?

The Education Reform Act explicitly confirmed that the Government's approach to state education would, in the future, be assessment driven. Such a strategy, it argued, would enforce a market-driven approach to education, in common with government policy in many other aspects of the public sector. Assessment would provide information for the client who could, as a result, reach 'informed' decisions. Neither the Government nor the Department of Education and Science could have foreseen the consequences of this strategy and the degree of intervention which would be required in trying to make it a reality.

This objective, in uncomplicated terms, consisted of establishing a precise definition of the curriculum, so that pupils' performance could be tested, performance could be aggregated at a variety of scales and national standards could be established. This would then give parents norm-related information about the achievements of their children and the schools which they attended. However, also central to the Education Reform Act was the objective that the National Curriculum would allow every child to develop his or her potential, according to his or her ability, hence the adoption of criterion referencing. But here lay the ambiguity which, with hindsight, could only be resolved by compromising either one or both of these objectives, to foster the education of the individual or to institute and monitor national performance standards. The first objective, focused on the individual, was the one which the teaching profession could endorse whilst the second, essentially a quality assurance system, was the Government's priority. At

the outset the Government wished to submit every subject, whether it was core or foundation, to a testing and reporting regime. As the problems mounted their ambition was thwarted and priorities started to emerge. These became focused on the core subjects of English, maths and science.

So what was the level of the Government's commitment to design and technology at key stage 3? As has been described in chapter 2, this subject had with the introduction of the National Curriculum, reached the summit of its evolution. The Education Reform Act had formally recognised design and technology as an experience essential to every child's education between the ages of 5 and 16. Technology, of which design and technology was one of the two profile components, had been introduced in parallel with English; an implementation schedule only preceded by maths and science. Consequently, technology was often described as part of the extended-core, but this term was never used in any official publication. It is unlikely that technology had quite the same political importance as the other three subjects which were defined as the core. This was borne out by the deregulation of the subject during the period of the first statutory assessment;[17] unlike the core subjects, whose statutory assessment has remained a legal requirement from 1992, for maths and science, and 1993, for English. This, deregulation, could be explained in a number of ways but perhaps two have the most credibility. The first is that NCC's recommendation that the Order should be revised was an admission that it was faulty. If this were the case, it was argued, it would be unwise to use the Order as the basis for statutory assessment. This was a convenient, but flawed, argument as it was only applied to key stage 3. At key stage 4 new GCSE syllabuses, which assessed the 1990 Order, were introduced in September 1993 and were assessed in the Summer of 1995 and will be again in 1996, three years after the Order was deregulated and statutory assessment aborted at key stage 3.

The second explanation is perhaps more convincing. The Government, by the Summer of 1993, had finally appreciated that the scale of the bureaucracy necessary to assess all the National Curriculum subjects was unacceptable. It would have been unthinkable to capitulate on the core subjects but technology, with all its associated problems, was a permissible sacrifice. It is unlikely that at the outset the Government appreciated the need for in-service education and quality control mechanisms, if the system was going to be credible. The technical feasibility of a system which tested all pupils in 10 or 11 subjects, four times during their school career had been questioned from the outset, but the Government had pressed on regardless. Design and technology teachers, given the radical changes which the statutory Order imposed, required subject focused INSET to help them effect these changes. This was a priority, but in addition, as with all other subjects, INSET was necessary to help these teachers implement the assessment procedures. Despite the benefits which might have accrued for design and

technology from its statutory assessment , it cannot be viewed in isolation. It was part of a system which had already become grossly bureaucratic and totally unmanageable, with less than half the subjects being assessed. The Government had to take steps to recover the situation.

Despite the increasing criticism the Government attempted to defend its policy by maintaining the value of the tests. For example the schools minister, Baroness Blatch speaking late in 1992 maintained:

'I have no doubts that next Summer's tests will be of great value in raising standards still further and in providing objective and reliable information for teachers and parents about the achievements of their children'[18]

By Spring of 1993 the Government was becoming ever more isolated. Criticism mounted from across the political spectrum. It was to be expected that unions would question the system:

'The question remains whether the complexity of National Curriculum assessment, recording and reporting as currently conceived is in fact manageable by all schools.'[19]

but when a SEAC Council member, appointed because of his right wing views, commented:

'I agree with teachers; the new testing system is over complicated, bureaucratic and time consuming. Moreover, as often as not it doesn't even produce good tests.'[20]

it was apparent that a retreat was soon to commence. Conservative councils, notably Wandsworth, attempted to support the Government by challenging the union boycott in the courts. This did not succeed, as the High Court ruled that the action proposed by the NAS/UWT constituted a legitimate trade dispute. In response to this continuous and mounting climate of confrontation the Government started to back track by simplifying the system. Design and technology was only a pawn in this battle and, initially, few regretted its removal from the statutory regime.

Statutory testing at key stage 3 had failed largely because of the difficulties encountered with English. In 1992, 75% of schools had voluntarily taken part in national pilots of maths and science and these had been accomplished with relatively few problems. English though was more difficult to constrain within the new rigorous, short sharp testing regime demanded by the Secretary of State for Education, Kenneth Clarke. SEAC had not found an agency which could successfully devise such tests on the basis of the English National Curriculum Order. Contracts with two agencies had already been terminated, the second agency only having been appointed a few months earlier. With only a matter of months to go the English development was in a state of chaos. This, coupled to a long running debate about the nature of the English Curriculum, resulted in its becoming the focus of the public debate that was now taking place in the media in relation to testing.

The interest of the media was not confined to English. The publication of the design and technology practical task in November 1992 created a great deal of interest in the media (see chapter 3 pages 87 to 89) and a wide range of responses from the educational press. But there was to be no back down on English which had been at the centre of the dispute; it was design and technology, along with all other foundation subjects, which was deregulated in the context of testing. As a consequence the status of the subject was diminished and schools and headteachers across the country started reversing decisions which they had been forced to take by the statutory requirements of the first Order. For example, pupils were to revise key stage 4 subject choices and drop design and technology.

It was generally recognised that the nature of the technology Order made the development of standard assessment tasks in this subject perhaps more difficult than in other more traditional subjects:
'The Order in technology, for example, contained many loose, generalised statements of attainment, broad programmes of study with much discretion for the individual teacher..... How could the agencies charged with designing SATs produce valid, reliable tasks capable of being administered over a short span of time?'[21]
This thesis is evidence that this was indeed a complex task and it is difficult to measure success in relation to the impact of these assessment activities on learning experiences. Substantive evidence might only have emerged following several years of statutory assessment. Even the insight which might have emerged from the experience of one statutory assessment failed to materialise. The teachers' boycott of the first statutory assessments prevented the tests, tasks and procedures from being fully exposed to the 'consumers' - pupils, parents and teachers. From the outset this research had to lead the implementation of the new technology Order in order that there were genuine examples of design and technology which could be assessed. Trialling and piloting with schools throughout the country put these schools at the forefront of the implementation of the Order and these schools frequently became the purveyors and centres for local INSET. INSET delivered in the context of this research was highly regarded and over 40 one day sessions were held throughout the country. These were fully attended by HMI and LEA Inspectors and Advisers in addition to the teachers for whom they were intended. Clearly the type of experiences which were being designed as assessment tasks were also regarded as valid learning experiences. A confidential HMI report was produced on the implementation of the practical task in the Summer of 1993. The conclusions of this report were described by HMI Ives in a keynote address to an inservice conference.
'The SAT activity was for many pupils the most purposeful and coherent design activity undertaken throughout the key stage. HMI observed that pupils had

enjoyed the challenge because they knew what was expected of them, consequently they produced work of a quality not previously attained.'[22]

An independent report was commissioned by SEAC to evaluate the statutory process. It reported fully on the quality of the educational experience of the tasks and tests. The following extracts are taken from the report:
'Instructions for the Statutory Practical Task and Sample Test Questions: Design and Technology', was praised by deputy heads and heads of design and technology for being well presented, direct and detailed.'
(In a survey carried out for this report 99% of respondents, heads of design and technology departments, had seen and read this document and 88% - 'having read the document agreed that the document was useful')[23]
This report also commented favourably on the level descriptors used to assess the tasks:
'the assessment guidance proved to be very informative and helpful, particularly with key points being shown in bold..... Without this support on assessment, most teachers felt that their confidence in assessing the tasks would have decreased - the statements of attainment in the Order were considered by almost all teachers, to need clarification and amplification if the tasks were to be assessed with a high level of reliability.'[24]
Whilst being positive about the tasks, this report reached a similar conclusion to this research in relation to the tests:
'.... the design and technology tests were not able to assess what the reviewers judged the most important aspects of these subjects..... written tests, however modified, could not assess the current Orders with a high validity.'[25]

Although this research approached the central task from a number of directions it would be dubious to claim that it found a valid solution to this third key question. A variety of assessment tasks was devised which provided valid learning experiences, and consequently useful measures of performance resulted from them. As these tasks became more constrained, to meet the ever more tightly specified constraints, some of these objectives were compromised. The Practical tasks continued to provide valid learning experiences but no such claim could be made for the tests which produced assessments of Te1 and Te4. These were contrived devices which had little bearing on the notion of design and technology capability. If, as in other subjects, an attainment target had specified the knowledge component of the subject it would have been relevant to have tested pupils' understanding. A similar problem exists with the revised Order, although the linkage between the programme of study and the attainment targets is sufficiently robust to make the task achievable in a sensible way.

The relationship between different modes of assessment

The fourth key issue was:

- to investigate the relationship between end of key stage assessment, terminal and summative in nature, and ongoing teacher assessment, continuous and formative in nature.

The process by which teacher assessments were arrived at was not the priority of this research, but teachers were asked to provide teacher assessments for all pupils in advance of their taking an assessment task. If teacher assessments had been provided in which real confidence could have been placed, they would have provided a valid reference against which SAT scores could have been compared. The majority of schools which took part in this research, however, did not have a system in place for producing teacher assessments for the pupils taking part in these pilots. This should not be seen as a criticism of the teachers concerned, as there was no legal requirement to produce teacher assessments except for those pupils completing Y9 in the Summer of 1993 and for all pupils in subsequent years. It was apparent though that, even in 1993, many schools had not produced teacher assessments for those pupils who were to be the focus of the first statutory assessment. The evidence from reports commissioned in other subject areas was undoubtedly also true for design and technology:
'schools are adopting a 'wait and see' policy whereby practices were not being developed until absolutely necessary, i.e. until 'someone tells us we have to do it' as one head of science remarked.'[26]
In the light of events this pragmatic approach was fully justified.

The relationship between these two components of the assessment regime changed during this period of research. The shift in importance from statutory assessment to teacher assessment was very gradual but, in the context of the revised Order, it is only teacher assessment which schools will be legally required to report in 1997. During the early years of the National Curriculum teacher assessment was almost totally neglected as it was certainly not at the top of the Government's agenda. The Task Group on Assessment and Testing had accorded teacher assessment equal importance with end of key stage assessment, see chapter 1, page 19. But it was undoubtedly the Government's reluctance to acknowledge the importance of ongoing assessments carried out by teachers as part of their normal teaching duties which led to the spotlight being turned on to statutory end of key stage testing. It was not until mid-1991 that SEAC published any information on teacher assessment at key stage 3. A short policy statement, in the form of a pamphlet,[27] was followed by more substantial documents relating to general issues and more specifically maths and science. The author of this thesis was commissioned and wrote a similar document for design and technology but

before publication the Secretary of State imposed an embargo on all 'non-essential' documents. Ironically this might have been a document welcomed by teachers as it offered practical advice and exemplar material on how to implement and manage teacher assessment. Realistically though, the information overload imposed on teachers had been unreasonable during this period, although secondary teachers did not encounter this on the same scale as their primary colleagues. There were only two documents, in advance of the SAT material, which were required reading for key stage 3 design and technology teachers, the Order and the non-statutory guidance. It was not uncommon however, for teachers attending INSET sessions during 1992 to ask, for example, how many attainment targets there were in design and technology.

SEAC was clear that teacher assessment was a matter for schools and individual teachers and would not be the subject of a closely monitored national system. The reason for this was clearly expressed in SEAC's first statement on the subject: 'National Curriculum assessment has two main purposes: to indicate what levels of attainment pupils have reached at the end of the key stage and to build a picture of pupils over a period of time in order to help carry them forward in their learning. Teacher assessment is mainly concerned with the latter.'[28]
Teacher assessment was becoming an ever more demanding element of the total assessment picture. For schools applying the letter of the law conscientiously, each year brought the need to start recording another year group's performance. So by 1993, a design and technology department should have had detailed records of all pupils in years 7, 8 and 9 in relation to coverage of the programme of study and performance against each statement of attainment. It addition, it had been suggested that schools should keep portfolios of work so that assessment standards could be justified. The manageability of such a system was highly questionable. Undoubtedly with experience it would have become less onerous, but it was unclear to teachers even if they implemented such a system if anyone would place any importance on the scores which it produced.

In the context of this research it is not surprising that teachers frequently based their assessments on the work undertaken in the pilot or trial. Frequently they used a system, such as the level guide - see appendix 4.4, page 289, which had been developed specifically for other purposes. In 1997 design and technology teachers will be legally required to report teacher assessments in relation to the two attainment targets of the revised Order. The simplified level statements will make this a much simpler procedure. However, the lack of detail in the new attainment targets makes them less than satisfactory for formative assessment, which is the purpose of teacher assessment. The two new attainment targets would have been ideal for summative statutory assessment. Ironically teacher assessment will now

be channelled towards producing a summative score and its central importance to teaching and learning may well be reduced.

A genuine purpose which statutory assessment might have fulfilled was as a calibrating tool for teacher assessment, one which tries to ensure that there is a consistency in standards between schools. However, this is essentially a norm-referencing procedure which does not sit well in a criterion referenced system. For example, to tell a school, following moderation, that its assessments in design and technology are one level too high makes nonsense of each pupil's performance. They will be told that although they have evidenced level 6, they will be reported as having achieved level 5. Only by reporting both scores could fairness have been maintained, although the combined reporting of both statistics would have been an interesting one to have explained to parents. In the context of statutory assessment teacher assessment had a vital role to play, as it was on the basis of these assessments that pupils were allocated to a tier. An independent survey commissioned by SEAC but reporting to SCAA, evaluated the national assessment of Technology at key stage 3. In relation to this issue it concluded:
'Most teachers in case-study schools had considerable difficulty in establishing teacher assessments as a means of allocating pupils to the most appropriate tier for the task. ..... It was noted that many teachers did not use National Curriculum criteria when deciding on the teacher assessment levels.
in addition it commented,
'Observation of teacher assessment suggested that
- implementation of teacher assessment in design and technology at key stage 3 had begun but that the capturing of ephemeral evidence was at a very early stage;
- in particular, there was very little evidence of individual or differentiated assessment.[29]
This confirmed the evidence of this research that teacher assessment was still in its infancy at the time of the first statutory assessment.

The TGAT system, based on a dual approach to assessment, was not new. The inclusion of school based assessments had been justified by the Schools Examination Council in its paper on Course work assessment in GCSE:
'School-based assessment is there to test aspects of attainment which may not easily or adequately be tested by (final) papers.'[30]
The recognition that tests had shortcomings which could be overcome by including the assessment of coursework was central to GCSE examinations. It allowed all aspects of a course to be assessed, not just those which were easily tested but, in addition, it also recognised performance over time and across the full range of the syllabus. This was recognised by the Department of Education and Science in the GCSE General Criteria.[31] Teacher assessment was seen as essential to validate

the assessment process, to ensure a degree of congruency between what was done in school and what was examined. Examinations, in contrast, are highly reliable as they aim to create similar conditions for all the candidates, to ensure fairness. Consequently, in any formal assessment process, there is a trade-off between teacher assessment with a high level of validity and examinations which are reliable. In the context of the National Curriculum, combining the scores does not provide a solution, As noted earlier, it is misleading to bring scores together and aggregate them in a criterion-referenced system as the final score may have little relevance to the performance on which it was based.

This is the contradiction of a system which attempted to produce statistics for too many purposes. Perhaps greater consideration should have been given to just using the SAT scores to produce a performance level for the school. Aggregating scores to produce a school norm would have been acceptable as it was not directly related to what any individual had achieved. Pupil performance in the SAT would have been treated like any other piece of work during the key stage. It could have contributed to the end of key stage teacher assessment, which would have become the only reported score for the individual. Parents and pupils could have compared their performance to the school's norms and those of other schools, if they so desired.

In chapter 4, page 96, mean teacher assessments are compared with mean SAT scores. This gives an overview of the relationship between them, but it would have been of more value if performance could have been monitored at the level of the individual. This would have shown how closely performance in the SAT mirrored what had been expected. The numbers of pupils involved in the trials and pilots prevented such an analysis taking place. Some interesting general conclusions can be made, although they should be treated with caution because of the way in which teacher assessments were derived. Teachers expected pupils to perform better than they actually did by quite a significant amount in 1990, 0.8 of a level, but by a much lesser amount in subsequent years. Teachers' assessments of girls were higher than boys and generally much closer to their performance in the statutory task. In relation to the attainment targets teachers always norm referenced performance correctly. They assessed performance as being at the highest level in Te3, followed by Te2, Te1 and the lowest being Te4. They did though over estimate performance in Te4 by the largest amount. They recognised that performance in this attainment target would be poor but they did not appreciate how poor. There was, though, evidence that teacher assessments and SAT assessments were converging as statutory assessment approached. By 1992 the mean difference between the two measures was approximately 0.2 of a level for each attainment target compared with a range of 0.5 (Te1) to 1.2 (Te4) in 1990.

In an ideal situation, performance should be similar whether it was derived from teacher assessment or end of key stage statutory assessment. Indeed, it can be argued that if the end of key stage assessment actually takes place at the end of the key stage then it should represent each pupil's best opportunity to record their highest level of achievement. In design and technology, there are many specific, as well as more general reasons, why this might not in fact be the case. For example, the material (food, construction material, textiles etc.) in which the task is taken, the context in which the task is placed or the aspect of the programme of study on which it draws.

In 1997 the reported teacher assessments should be valid as the assessments will be based on what pupils have actually done. They will reflect all aspects of the subject, there will be no limit on the range of skills and aptitudes which will be reflected in the end of key stage score. Without the restriction of a timed task, research skills, interactive skills, motor skills, resourcefulness, organisational skills and self motivation, for example, will be recognised in the levels gained. This is extremely healthy for a subject based on positive, purposeful achievement. The reliability of the scores produced may though be compromised. Even if SCAA does not seek to moderate reported scores there will undoubtedly be a need for other organisations, in the cause of fairness, to look critically at the levels recorded and to comment on the standards being applied.

Final conclusions

At every stage of this research recommendations were made to the Schools Examination and Assessment Council concerning the evolution of statutory assessment of design and technology at key stage 3. These recommendations helped formulate the policy decisions which led to the publication of the first statutory tasks in the subject and the associated legislation. However, these recommendations were not the sole advice on which these decisions were based. It is important to isolate the findings from this research, irrespective of the current retreat from statutory assessment, so that they may be of value to others in the future. These conclusions draw on the previous discussion in this chapter and are focused on eight key aspects. They seek to identify issues which might be of importance in the future.

The role of assessment in changing the curriculum

This research has provided sound evidence which supports the view that assessment procedures can be a powerful mechanism for changing curriculum content and teaching style. Without some means of enforcement it is unlikely that a statutory National Curriculum, which defines entitlement, would be fully adhered to.

This can be exemplified in the context of the evaluation process carried out following the 1991 pilot. Of the teachers involved in this pilot 84.2% indicated that the statutory assessment task had changed their approaches to both organisation and teaching style in some respect (59.9% - in some aspects; 24.3% - considerably). This finding was some eighteen months after the introduction of the Order and these tasks were consistent with its intent. Indeed, when expert scrutineers were asked to rate these tasks as effective assessment activities, in relation to the requirements of the Order, an overall rating of 4.4 was achieved on a 1 (poor) to 5 (excellent) ordinal scale. This was sound evidence that the majority of schools and teachers only got to grips with the required changes when compelled to by the assessment requirement. This was reinforced by pupils, 77% of whom when asked: How different was the work you did this term compared to what you normally do?, indicated it was 'quite a bit different.'

The revised National Curriculum became statutory in September 1995. This is a minimum curriculum designed to meet the essential needs of all children. The Dearing review established the principle of parity for end of key stage assessments and teacher assessment. At key stage 3 design and technology is no longer the subject of statutory assessment. The task of ensuring that all pupils are taught the curriculum to which they are entitled will be the remit of the Government's inspection service OFSTED. In 1997 schools will be legally required to report a design and technology teacher assessment for every pupil. The elementary approach adopted in relation to the revised attainment targets will greatly simplify the assessment of design and technology, making it far more manageable for all schools. Of more concern will be coverage of the programme of study, this will challenge the majority of schools at this key stage. If statutory assessment is reintroduced, it will be the coverage of the programme of study which will require scrutiny as well as capability. The revised Order is also more traditional in its aspirations than the first Order. It is likely that many schools will feel more comfortable with the demands which it makes. However, there will be a need to investigate schools' compliance with the Order now that the 'policing' effect of statutory assessment has been removed. There is still much to be done to raise the quality of design and technology to an acceptable level in all the nation's schools.

The style of assessment device

A wide range of assessment devices has been conceived, tested and evaluated to meet changing specifications. The validity and reliability of these devices has been explored and conclusions reached. Each has attempted to assess design and technology capability as defined by the first statutory Order. These assessment tools evolved from ones which attempted to measure achievement as part of the learning process, with pupils' performance determined in a normal learning context,

towards summative assessments undertaken in controlled conditions. This evolution was promoted by a political belief that external, objective tests would take less time, be more manageable and overcome local effects which might skew assessment. This research has shown that valid, reliable and effective assessment tools can be designed to meet different specifications. It is though fully acknowledged that as the specification was modified so was the interpretation of the Order, but control of this was beyond the author of this thesis. However, if the notion is accepted that genuine assessments of design and technology capability will only result from 'authentic tasks' then the following conclusions can be drawn from this research.

A complete holistic design and technology task, as developed during the first two years of this research, is well suited to producing assessments from normal teaching and learning activities as these are undertaken in appropriate conditions. Consequently, the ensuing assessments should accurately reflect each individual's National Curriculum level of design and technology capability. The use of criterion-referenced statements will also allow performance to be accurately described in terms of what the pupil can actually do. However, such a style of task is difficult to constrain within the regulatory framework which statutory assessment requires. Time, resources and teacher involvement might all, for example, influence pupil assessment. Consequently, in a statutory regime the resulting assessments would need to be subjected to external moderation to ensure standardisation of assessment between schools. This would require a costly infrastructure and would replicate at key stage 3 many of the problems encountered by GCSE boards, not least of which are issues relating to the volume/size, scope, and 'shelf-life' of the practical work resulting from design and make activities.

Statutory assessment, which is principally summative in nature requires standard settings and conditions if it is to be fair to all. These are difficult to establish, in relation to design and technology, within an authentic context. Such assessments, therefore, may not provide accurate assessments of each individual's design and technology capability. Assessment procedures of this nature will invariably tend to revert to a traditional marking regime with performance being judged in comparative terms. This type of assessment process is more fitted to a comparative role, one which might use schools' aggregated mean performance for purposes of comparison. Statutory assessment devices can be designed which will focus on either individual performance or school achievement. But the evidence is clear, the purpose of the assessment must be clearly stated before the tools and systems are devised. Assessment procedures which attempt to serve too many purposes frequently fail to satisfy any adequately. Assessments carried out primarily for quality assurance purposes are extremely expensive in relation to both time and cost and may do very little to support teaching and learning.

Satisfactory quality assurance mechanisms in education are still in their infancy and there is much to be done to consolidate recent developments. Any systems which evolve from those currently in use, should be in response to a collaborative process involving all relevant parties. The strategy of recent years has been based on imposition via legal statute. It would be interesting to have seen if market pressures would have encouraged schools to have become involved in a voluntary system of national testing and reporting. Over a similar period of time, this might have been more effective than the compulsory strategy which was adopted.

The nature of assessment criteria

The statutory Order had at its core statements of attainment: attainment descriptors organised into ten hierarchical levels. These were conceived as the criteria against which every child's performance from the ages of 5 to 16 would be judged. These statements were impersonal, complex sentences with considerable embedding, i.e. use of subordinate clauses, and frequent qualification to the main proposition or requirement. The text was also complicated by the use of unspecified verbs (e.g., 'identify' - how and in what way) and nominalisation (the way in which a verb describing an ongoing process is turned into a noun). Such complexity made the statements both difficult to 'unpack' and to reach agreement as to their meaning and interpretation. This made them unreliable for assessment purposes. To increase the reliability of assessments this research proposed and implemented a number of strategies. Evidence from as early as the 1990 trial indicated that these alternative strategies were easier to use and more reliable than the use of the statements in the Order. The two principal alternative strategies involved the use of a stranded matrix and level descriptors (see chapter 4, pages 102 to 105). In a qualitative evaluation of these two devices and the statements of attainment, the approach using level descriptors was ranked first and that using statements of attainment last. This evaluation required teachers to rank the three approaches in order of preference, the mean ranking, first one point and third three points, was as follows; level descriptors - 1.5; stranded matrix - 1.75; statements of attainment - 2.75, see appendix 4.5 and 4.6. This was conclusive evidence which committed this research to pursuing and advocating an assessment regime based on simple, unambiguous level descriptors as the best approach for producing reliable assessments of performance. As has been noted elsewhere, this approach has now been adopted for all National Curriculum subjects.

Generalised level descriptors provide a far more reliable assessment criteria for summative statutory assessment. However, such descriptors are far less useful as judgement points and goals for normal teaching and learning as they provide insufficient detail. The new Order is described as a minimal approach, the basic essentials which a pupil should be taught. Curriculum development, focused on

assessment, will need to be undertaken to provide schools with resources which go beyond the bare necessity. The heightened sense of awareness and importance concerning the value of assessment to teaching and learning was one of the most positive aspects of the National Curriculum. Many design and technology teachers started to see how assessment was an essential tool in improving the quality of their teaching, and consequently of pupils' learning. Care must be taken to ensure that this embryonic skill flourishes. Design and technology has a relatively barren history in this regard and there is much which can be done to nurture this essential component of teaching and learning.

The interpretation of assessment statements

Design and technology is essentially a practical subject, consequently teachers of the subject are more comfortable when assessing performance in terms of practical outcomes. Along with many other teachers of all other subjects they have found the transition to judging pupils' work against theoretical criteria a difficult one to make. This research has demonstrated that once the criteria are contextualised, via a task which pupils are undertaking or exemplified against a pupil's response to a task, their potential as a tool which can be applied reliably is greatly enhanced. This is of particular importance in design and technology because of the variety of backgrounds and range of subject expertise found within the teaching force. This hindered the development of a consensus view of the interpretation of the statements of attainment. There is substantial evidence that this is only resolved by exemplification material based on tasks, whether statutory or non-statutory, accompanied by contextualised assessment criteria, i.e. ones which describe exactly what a pupil has to do in relation to the task to achieve a level.

Both approaches were trialled as part of this research, but with the exception of the marking schemes accompanying the tests, never concurrently. In the 1991 pilot 70% of teachers found the sample assessments useful in making judgements. Because of the open nature of the tasks undertaken in this pilot, these sample assessments were not sufficiently specific. Contextualised assessment criteria had been adopted by the time of the statutory assessment (see chapter 4, page 109 to 111) and these were judged to be extremely beneficial. Ministerial pressure on the Schools Examination and Assessment Council to reduce the amount of written material being sent to schools prevented the publication of assessed pupil portfolios. If, in the future, non-governmental organisations seek to standardise statutory assessment they would be advised to adopt this dual approach of specific tasks, which include contextualised assessment criteria, exemplified by assessed pupil outcomes complete with a justification of the judgements reached.

In addition, curriculum research projects can do a great deal to promote sound and consistent interpretations of level statements. Currently a number of national

projects are producing substantial amounts of curriculum material. These should be accompanied by examples of pupils' work and outcomes in response to the advocated tasks. Assessment commentaries can assist teachers to recognise standards and the key aspects which determine differentiation. Modern technologies, such as CD-ROM and the Internet, provide powerful tools for the transmission and dissemination of such information.

The nature of practical tasks

The statutory assessment of a whole year group, approximately 600,000 pupils, is a unique undertaking especially in design and technology with its practical dimension. The size of the task is far greater than that undertaken by any GCSE examining group, any of which even in 1995, the only year in which pupils were legally required to follow a key stage 4 course, probably had no more than 60,000 entries for any single examination. The difference between a National Curriculum test and a GCSE examination is quite marked. In the context of the National Curriculum at key stage 3, in 1993 the pupils' teachers were also the examiners and the level standards were established prior to the pupil taking the practical task or test. These standards were known in advance to both the pupil and his or her teacher. Unlike GCSE, there was no recourse to post test standardisation and determination of grade boundaries; these were known in advance. Consequently a particular type of task had to be devised. The task had to provoke pupils to demonstrate of what they were really capable. There is good evidence from this research that a single, relatively open-ended task, might not have achieved this, especially as at key stage 3 assessments had to cover all ten levels. Within the one-off strictures of statutory assessment pupils needed to be confronted by tasks targeted on their level of achievement as determined by teacher assessment. Tasks had to be devised which made specified demands on each pupil's level of capability. The model displayed diagrammatically on page 124 features these characteristics. Differentiation by task was novel to this subject but since this research this approach has become more common.

The development of more tasks which place different demands on pupils of varying ability is essential if the ablest are to be fully motivated and stretched, yet every pupil is provided with a task which offers them the opportunity of success. Tasks of this nature place a much greater demand on teachers as they need to know each pupil's level of ability. It is apparent that teachers prefer to set every pupil a similar challenge, an approach which has lost credibility with the majority of other subjects on the curriculum. (In the contexts of the SATs many design and technology teachers were strongly opposed to entering pupils for different level tasks, indicating that it was not their responsibility to make such judgements.) It is right to recognise that many teachers do not have the time or skills to develop tasks of this

nature. If tasks are to successfully administered there must be a variety which recognises the full range of materials and contexts within the design and technology spectrum. The notion of a single task, or even limited task bank, is difficult to accommodate within design and technology. This research has demonstrated that different tasks can be devised which offer pupils fair and equal ways of determining their capability. If statutory tasks had continued there would have been a continuous supply of tasks, differentiated by specification, which would have enriched the curriculum. Others, especially those responsible for curriculum projects, should investigate how this model could be adapted for normal teaching and learning contexts and as a result produce suitable tasks. The subject needs to develop a more sophisticated strategy for the setting of tasks to replace some of the simplistic approaches currently employed.

The standardisation of pupil material

In the context of statutory practical tasks, this research has demonstrated that it is important that some consistency is prescribed in relation to the way in which pupils respond to the task which is set. A framework is essential for a number of reasons, as it:

- ensures that pupils are aware of the criteria against which they will be assessed and allows them to plan their approach to the task;
- provides teachers with a common structure which assists both in the location of evidence and in making comparative judgements: both at the stage of reaching decisions and when undertaking moderation;
- imposes a constraint on the documented evidence, ensuring judgements are taken on the basis of quality rather than quantity;
- can assist the notion of fairness by attempting to equalise resourcing.

This research proposed a number of different frameworks ranging from preprinted project sheets, to labels (used as signposts) and structure diagrams. These are described in chapter 4. Each of these approaches found favour with some teachers and not with others. Perhaps the most popular were the labels (56% very effective rating from teachers in the 1991 pilot). Lack of familiarity, was generally the most common criticism directed at each of these procedures, one which, with time, disappears. Many might feel that such a framework for pupils' responses is restrictive and inhibiting, but it is clear from this research that it is essential within a national statutory regime to promote fairness and reliability.

The nature of tests

The tests developed as part of this research attempted to assess aspects of a practical process driven activity, by asking pupils to write about what they had done or what they might do in an imaginary context. Initially, in the 1992 pilot, assessment judgements were made against the statements of attainment. As this

approach failed to reward pupils for partially correct answers, they were assessed as either having wholly satisfied the statement or not. There was no interim position. Subsequently, procedures were developed which moved closer towards a marked paper. This was in direct response to teachers' evaluation and responses to the 'satisfied/not satisfied' system driven by the statements of attainment. A traditionally marked test with levels awarded on the basis of predetermined mastery points begins to have less direct relevance to the performance criteria the pupil was attempting to evidence. When this criterion described practical performance the integrity of this procedure was difficult to defend. This research suggests that test papers based on theoretical questions, set in hypothetical contexts cannot be guaranteed to reveal a pupil's true level of capability. The conclusion of this research is that tests do have a place in design and technology but they should be used to assess what, traditionally, they have done in the past - understanding and coverage.

The programme of study defines the knowledge and skill which a pupil requires to operate purposefully and effectively. Within the time constraints of statutory assessment, or indeed external examinations, coverage of this knowledge and a pupil's understanding of it is more efficiently assessed via a test. Such a test will, however, be difficult to devise as the programme of study of the revised Order is not defined by level. Tests of this nature will need to be devised, because at key stage 4 the national criteria for shortened and GCSE courses require part of the assessment to be in the form of a test. This research has established some sensible guidelines for those setting tests in the future, see chapter 4 pages 137 to 139. In addition, it has highlighted the need to ask specific questions, appropriate to each pupil's level of capability. This in turn confirms the need for tiered papers so that each pupil is posed questions within his or her ability range and has a reasonable chance of success. Tiered tests allow a pupil to evidence what they know, whereas a single test paper commonly differentiates on the basis of what pupils do not know.

The assessment process

Statutory assessment via end of key stage tests is currently, at key stage 3, restricted to the core subjects. In 1997 Teacher Assessment of design and technology will become a statutory requirement. These assessments will have to be reported but they will not be standardised by any annual formal system. School standards will, as now, be the responsibility of the OFSTED Inspection process, carried out on a four year cycle. It is likely that the Schools Curriculum and Assessment Authority will develop materials to help schools reach assessments. However, without a standard assessment task, being used as a national standardising mechanism, there is no means of validating teacher assessment.

SCAA will, because of financial implications, almost certainly adopt an approach which will be optional. This may include optional tasks and advice on how the attainment targets can be used for assessment purposes. If this is the proposed approach, it is apparent that a key lesson from this research has only partly been absorbed.

If schools are to adopt an assessment procedure it is essential that they have access to all the relevant materials in advance of pupils embarking on the phase of their schooling which is to be assessed. This is one of the key issues leading to the failure of statutory assessment in 1993. Unlike GCSE Examination Boards which are required to publish sample assessment materials when a new exam is introduced, statutory assessment did not observe this principle. The pressure to publish the Order and the demand from politicians for results within a parliamentary time span being the main obstacles. If all schools had received copies of sample statutory assessment tasks at about the same time as the Order became law many of the problems encountered in design and technology would have been surmounted. If following the adoption of the Order there had been a two year period for the development of assessment tasks and for schools to plan and prepare for the implementation of the Order there would have been a far more positive outcome. Indeed the 1990 Order, which many argue was more forward looking than the revised Order, might have gradually been interpreted via assessment tasks and as a result assimilated into the school curriculum.

Statutory assessment is an issue of great complexity which provoked the most intense national debate on education since that concerning the 11+ and comprehensive education. Both of these issues were fuelled by deeply held convictions as to how assessment affects the personal esteem of all those involved and concerned - pupil, student, teacher, parent. Statutory assessment proved to be a highly emotive issue and one which has only been partially resolved by compromise and conciliation - a truce has been negotiated using an impartial intermediary. As a design and technology specialist, committed to the subject's having a central place in every child's education, I have deep concerns about the quality of design and technology experienced by many pupils and students. These concerns focus not just on the standards of their achievements but on the relevance of what they do to the modern world. Design and technology achieved its high profile within the National Curriculum partly because it was seen as a means of achieving a range of educational objectives which were complementary to the subject, especially those linked to enhancing the status of vocational education. These expectations have proved, in the short term, to be wholly unrealistic. The subject, given the innovative nature of the 1990 Order, was highly vulnerable. It required careful nurturing and acclimatisation, rather than the immediate and

sustained exposure to which it was submitted. As Layton (1995) persuasively comments:

'It would be sad if an exciting and radical curriculum innovation, potentially of great significance, should collapse under the weight of the unrealistic responsibilities being placed upon it.'[32]

To prevent this occurring it is essential that teachers of the subject are provided with the necessary support and resources. Over the past thirty years a wide range of strategies has been, and is being, employed to remedy the deficiency. Some curriculum resources of unsurpassed quality have been produced and as a result many teachers have been converted to the new and exciting educational experience of design and technology; many others, however, have been reluctant to change ingrained habits and practices. This was the first attempt at statutory assessment and it failed. It is not the most palatable remedy for changing the curriculum, but if the right strategies had been employed from the outset it might have become a powerful tool for improving an essential experience to which every child is entitled.

I trust this thesis provides a detailed and informative record of the research undertaken in developing statutory assessments in design and technology at key stage three. It has produced some sound outcomes which will make such a task more achievable, if in the future a similar policy is adopted. In addition, it has established a range of strategies and approaches which is applicable to the assessment of design and technology regardless of the context; indeed many of these have already been utilised in GCSE examinations. Until the past fifteen years the assessment of practical, process driven subjects, such as design and technology, has been largely neglected. I hope that this thesis will add to the subject's expanding literature by bringing into the public domain research which might otherwise have been unavailable as a resource for future researchers.

References

1. Daugherty, R. (1995). *National Curriculum Assessment - A Review of Policy 1987 - 1994.* London: Falmer
2. Graham, D. &Tytler, D. (1993). *A Lesson for us all.* London: Routledge, page 53
3. Department of Education and Science and the Welsh Office (1988). *The Interim Report of the National Curriculum design and technology Working Group.* London: HMSO, page 3
4. Department of Education and Science and the Welsh Office (1988). *The Interim Report of the National Curriculum design and technology Working Group.* London: HMSO, page 4
5. National Council for Vocational Qualifications (1991). *Guide to National Vocational Qualifications.* London: NCVQ
6. Goldstein, H. (1994). 'Recontextualising Mental Measurement'. *Educational Measurement: Issues and Practice,* Volume 13,1
7. Gipps, C.V. (1994). *Beyond Testing - 'Towards a Theory of Educational Assessment.* London: Falmer, page 12
8. Resnick, L.R. &Resnick, D.P. (1992). 'Assessing the Thinking Curriculum: new tools for educational reform'. in Gifford, B.R. & O'Connor, M.C. (Eds) *Changing Assessments: Alternative Views of Aptitude, Achievement and Instruction.* London: Kluwer, page 43
9. Department of Education and Science - Assessment of Performance Unit (1982). *Understanding design and technology.* London: Department of Education and Science, page 5
10. Dearing, R. (1993a). *The National Curriculum and its assessment: Interim Report* . London: National Curriculum Council/Schools Examination and Assessment Council, 5.23, page 52
11. Daugherty, R. (1995). *National Curriculum Assessment - A Review of Policy 1987 - 1994.* London: Falmer, page 50
12. Woodhead, C. (1993). 'Do we need a new National Curriculum' in O'Hear, P. and White, J., *Assessing the National Curriculum.* London: Paul Chapman, page 29
13. Department of Education and Science. (1988). *Task Group on Assessment and Testing - Report.* London: DES
14. Glaser, R. (1963). 'Instructional Technology and the Measurement of Learning Outcomes: Some Questions', *American Psychologist,* Volume 18 pp 518-521
15. Marks, J. (1991). *Standards in Schools: Assessment, Accountability and the Purposes of Education.* London: Social Market Foundation
    Marks, J. (1993). *Why there is no time to teach: What is wrong with the National Curriculum 10-level scale.* London: Social Market Foundation
16. Angoff, W.H. (1974). *Criterion-referencing, norm-referencing and the SAT.* The College Board Review, Volume 92
17. Dearing, R. (1993a). *The National Curriculum and its assessment: Interim Report.* London: National Curriculum Council/Schools Examination and Assessment Council
18. Blatch, J. Baroness, (1992) Quoted in *The Guardian* 16th October, London: The Guardian
19. Association of Teachers and Lecturers, (January 1993), statement to the press

20. Skidelsky, Lord, (1993). *The Sunday Telegraph*, 20th March. London: Associated Newspapers
21. Daugherty, R. (1995). *National Curriculum Assessment - A Review of Policy 1987 - 1994*. London: Falmer, page 51
22. Ives, M HMI (1995), Keynote address - TEP Summer School 18th August, Middlesex University - unpublished
23. NFER/Brunel University. (1993). *Report to SCAA on the 1993 National Assessment of English and Technology*. Slough: NFER - unpublished, pages 26 - 29
24. NFER/Brunel University. (1993). *Report to SCAA on the 1993 National Assessment of English and Technology*. Slough: NFER - unpublished, page 80
25. NFER/Brunel University. (1993). *Report to SCAA on the 1993 National Assessment of English and Technology*. Slough: NFER - unpublished page 93 - 112
26. NFER/Brunel University. (1992). *Teacher Assessment in Mathematics and Science at key stage 3*. London: Schools Examination and Assessment Council
27. Schools Examination and Assessment Council (1991). *Teacher Assessment at Key Stage 3*. London: Schools Examination and Assessment Council
28. Schools Examination and Assessment Council (1991). *Key Stage 3 Teacher Assessment in Practice*. London: Schools Examination and Assessment Council
29. NFER/Brunel University. (1993). *Report to SCAA on the 1993 National Assessment of English and Technology*. Slough: NFER - unpublished, page 48 - 49
30. Secondary Examinations Council (1985). *Course work assessment in GCSE*. Working Paper 2 London: Secondary Examinations Counci
31. Department for Education and Science (1985). *GCSE General Criteria*. London: Department of Education and Science
32. Layton, D. (1995). 'Constructing and Reconstructing School Technology in England and Wales'. *International Journal of Technology and Design Education*, Volume 5

1988 Interim Working Party Report - Five Attainment Targets

AT1 - Through exploration and investigation of a range of contexts (personnel, social, environmental, business, industrial) pupils should be able to identify and state clearly needs and opportunities for design and technological activity.

AT2 - Pupils should be able to explore, develop and combine designs and technological proposals, and use their judgements, based on various criteria (economic, technical, aesthetic, ergonomic, environmental, social) to choose an appropriate design for further development.

AT3 - Pupils should be able to develop their chosen design by refining and adding detail, and to produce a plan for making the required artefact or system by identifying tasks and sub-tasks, and ways of undertaking them, and by making judgements of what is realistic, appropriate and achievable.

AT4 - Working to a scheme derived from their previously developed design, pupils should be able to identify, manage and use appropriate resources, including both knowledge and processes, in order to make an artefact or system

AT5 - Pupils should be able to produce a critical appraisal of the processes, outcomes and effects of their own design and technological activity, as well as the outcomes and effects of the design and technological activity of others, both historic and present day. With respect to their own activity, they should be able to use their appraisal to propose and justify modifications to the processes they have used and to the outcomes realised.

1989 Working Party Proposals - Four Attainment Targets

AT1
Through exploration and investigation of a range of contexts (home, school, recreation, community and industry) pupils should be able to identify and state clearly needs and opportunities for design and technological activity.

AT2
Pupils should be able to produce a realistic, appropriate and achievable design by generating, exploring and developing design and technological ideas and by refining and detailing the design proposal they have chosen.

AT3
Working to a plan derived from their previously developed design, pupils should be able to identify, manage and use appropriate resources, including both knowledge and processes, in order to make an artefact, system or environment.

AT4
Pupils should be able to develop, communicate and act constructively upon an appraisal of the process, products and effects of their design and technological activities as well as those of others, including those from other times and cultures.

1989 Statutory Order - Four Attainment Targets

Te1
Identifying needs and opportunities
Pupils should be able to identify and state clearly needs and opportunities for design and technological activities through investigation of the contexts of home, school, recreation, community, business and industry.

Te2
Generating a design
Pupils should be able to generate a design specification, explore ideas to produce a design proposal and develop it into a realistic, appropriate and achievable design.

Te3
Planning and making
Pupils should be able to make artefacts, systems and environments, preparing and working to a plan and identifying, managing and using appropriate resources, including knowledge and processes.

Te4
Evaluating
Pupils should be able to develop, communicate and act upon an evaluation of the process, products and effects of their design and technological activities and of those of others, including those from other times and cultures.

1992 HMI/NCC Proposals - Two Attainment Targets

AT1
Designing

AT2
Making

Pupils should be able to design and make products safely by applying knowledge and skills from the programme of study for technology and where appropriate from other subjects, particularly art, mathematics and science.

Statements of attainment - levels 1- 10

1995 Revised Statutory Order - Two Attainment Targets

AT1
Designing

AT2
Making

Pupils should be able to develop their design and technology capability through combining their Designing and Making skills with Knowledge and Understanding in order to design and make products.

Level Descriptors - level 1 to 8 plus exceptional performance

Attainment target 1 - Identifying needs and opportunities Appendix 2.2

Pupils should be able to identify and state clearly needs and opportunities for design and technological activities through investigation of the contexts of home, school, recreation, community, business and industry.

Level 1

a. Describe to others what they have noticed in familiar surroundings or visualised about imaginery situations
b. Suggest what might be done.

Level 2

a. Describe what they have observed or visualised and found out in their exploration.
b. Suggest practical changes that could be made in response to a need and describe to others why they have suggested the changes.
c. Ask questions which help them identify needs and opportunities for design and technological activity.

Level 3

a. Starting with a familiar situation, use their knowledge and the results of investigations to identify needs and opportunities for design and technological activities.
b. Develop and clarify their ideas about possible needs and opportunities through discussions with those involved.

Level 4

a. Starting with an unfamiliar situation identify needs and opportunities for design and technological activities.
b. Devise ways to gather information in addition to using printed sources.
c. Recognise the points of view of others and consider what it is like to be in another person's situation.
d. Explain that a range of criteria which are sometimes conflicting must be used to make judgements about what is worth doing.
e. Provide oral and written justification for the conclusions they reach as a result of investigation.
f. Know that in the past and in other cultures people have used design and technology to solve problems in different ways.

Level 5

a. Show judgement in the choice of sources of information, both qualitative and quantitative, in the systematic search for a need or opportunity for a design and technological activity.
b. Recognise that economic, social, environmental and technological considerations and the preferences of users are important in developing opportunities.

Level 6

a. Explain how they have identified needs and opportunities for design and technological activities and give a justification of the conclusions they have reached.
b. Explain how different cultures have influenced design and technology, both in the needs met and opportunities identified.
c. Understand how the introduction of new technologies can offer new opportunities and create new demands for design and technological activity.

Level 7

a. Analyse information of several kinds and draw conclusions about the needs and opportunities for a design and technological activity, recognising and resolving conflicting considerations about what is worth doing.
b. Vary methods of investigation to obtain all the information required.
c. Consider both the user and the producer when defining the need for a technological activity.
d. Identify and draw upon sources of expert advice relevant to identification of needs and opportunities for design and technological activity.

Level 8

a. Provide a detailed evaluation, in the light of a range of considerations, of the needs and opportunities for design and technological activity.
b. Plan in detail the various stages of their investigation.
c. Investigate how needs and opportunities have led to design and technological activities in other cultures.

Level 9

b. Review their own knowledge and draw up a strategy to exploit expert sources.

Level 10

a. Convey, using presentation techniques matched to the audience, that the identification of needs and opportunities is justified and worth developing.
b. Elicit and interpret the perceptions, motivations and needs of people in a range of contrasting situations.
c. Make reasoned judgements about what is a subject for design and technological activity and what is better dealt with in other ways.

Attainment target 2 - Generating a design

Pupils should be able to generate a design specification, explore ideas to produce a design proposal and develop it into a realistic, appropriate and achievable design.

Level 1

a. Express their ideas about what they might do to meet an identified need or opportunity.

Level 2

a. Use talk, pictures, drawings, models, to develop their design proposals, giving simple reasons why they have chosen to make their designs.

Level 3

e. Record how they have explored different ideas about a design and technological proposal to see how realistic they might be.
a. Make a design proposal by selecting from their ideas and give reasons for their choices.
b. Apply knowledge and skills to select ideas for different parts of their design.
c. Draw from information about material, people, markets and processes and from other times and cultures to help in developing their ideas.
d. Use models including annotated drawings and three dimensional working models to develop their design.

Level 4

a. Record their ideas as they develop.
b. Review their design proposal to identify where decisions still need to be made; suggest possible courses of action which will improve their proposal.
c. Estimate the resource requirements and check on availability.
d. Describe and edit design proposals.

Level 5

a. Record the progress of their ideas showing how they have clarified and developed them.
b. Extend their first ideas by combining various aspects of them to formulate a design proposal and explain why some ideas were not used.
c. Seek out and organise information to help them develop their ideas and refine their design proposal.
d. Establish and check the availability of the resources required adapting their design as appropriate.
e. Specify what they intend to do and what they will need by using simple plans and flow diagrams.

Level 6

a. Produce a design specification and use it to develop their design proposal.
b. Produce a design proposal recording their decisions and the ways of reaching their chosen outcomes.
c. Make judgements about realistic ways forward by exploring alternative solutions and use these to refine their design proposal.
d. Use specialist modelling techniques to develop design proposals.

Level 7

a. Systematically seek out, appraise, organise, and use information from different sources to develop and combine ideas and judge how realistic they might be.
b. Review the detail of their design using their own experience and that of others, and suggest alternative ways of achieving what is intended.
c. Apply relevant criteria including user requirements, costs, time, skill demands, scale of production and aesthetic considerations to take decisions about the details of the design proposal.

Level 8

a. Record and present, using a range of methods and media, progress of their ideas; detail and refine their design proposal and incorporate modifications; use computer aided design, image generation and desk top publishing techniques where appropriate to explore detail and refine ideas.
b. Plan their activities to take into account multiple constraints which may at times be conflicting.
c. Show willingness to experiment and take risks subject to safety considerations recognising the implications of decisions taken in designing.

Level 9

a. Develop ideas by drawing on information and understanding from a broad knowledge of sources, and showing judgement about the detail required.
b. Refine their design to achieve an optimum practicable outcome demonstrating originality and understanding of constraints in the justification of their design.

Level 10

Provide a substantial account of the full range of ideas they have explored and tyhe startegies used showing: i) how they explored ideas used in existing artefacts, systems or environments and how they used them to develop their own ideas; ii) evidence they have: - identified ways of improving and refining their proposals; - predicted with accuracy the outcomes of possible improvements and refinements; - resolved conflicting demands; - included their decisions in a coherent specification; and using an appropriate range of media and methods.

Attainment target 3 - Planning and Making

Pupils should be able to make artefacts, systems and environments, preparing and working to a plan and identifying, managing and using appropriate resources, including knowledge and processes.

Level 1

a. Use a variety of materials and equipment to make simple things.

Level 2

a. Describe to others how they are going about their work.
b. Use knowledge of the working characteristics of materials and components including construction kits in making artefacts, systems or environments.
c. Show that they can use simple hand tools, materials and components.

Level 3

a. Consider constraints of time and availability of resources in planning and making.
b. Choose resources for making by using their knowledge of the characteristics of materials and components.
c. Use a range of hand tools and equipment, appropriate to the materials and components with some regard to accuracy and quality.
d. Improvise within the limits of materials, resources and skills when faced with unforeseen difficulties.

Level 4

a. Adopt procedures which minimise waste, pay regard to cost and achieve accuracy and finish.
b. Work with others in the planning and apportioning of tasks.
c. Choose tools, equipment and processes suitable for making their design and use these appropriately.
d. Adopt alternative ways of carrying forward their plan when difficulties are encountered and recognise when help is needed.
e. Use drawings, diagrams and models to assist making.

Level 5

a. Identify stages in making and coordinate these into a simple plan to ensure efficient use of time, materials and labour.
b. Use a knowledge and understanding of the properties of a range of materials in their planning and making.
c. Demonstrate by their choice and use of a variety of equipment that they understand the principles upon which these work and the requirements of safety and accuracy.
d. Apply knowledge of materials, components and processes to overcome problems as these arise.

Level 6

a. Plan and organise making in order to achieve the desired outcome.
b. Combine knowledge of the properties of a range of materials and processes to identify those most suitable for their design.
c. Demonstrate by their choice and use of a variety of tools and equipment, that they understand the limitations of them and the need for safety and accuracy.
d. Use knowledge of materials, components, tools, equipment and processes, to change working procedures to overcome obstacles as making proceeds.
e. Show judgement in seeking advice and information.
f. Use knowledge of technical and symbolic representations of materials, components and processes to assist making.

Level 7

a. Plan and carry out working procedures to match the constraints on making to overcome problems and to achieve the desired quality.
b. Demonstrate competence in the use of general planning and making skills as a result of understanding the materials, components, tools and equipment and the scale of production.
c. Use a range of technical, symbolic and other means of representation to assist in planning, organising, making and incorporating necessary modifications.

Level 8

a. Review how to make best use of materials, procedures, tools and equipment.
b. Show evidence of knowledge of making processes and devise and implement procedures for quality assurance.
c. Identify and incorporate modifications during making.

Level 9

a. Make judgements about the quality and usefulness of sources of advice and information consulted during planning and making.
b. Demonstrate how they have overcome constraints in planning and making to achieve a quality product.
c. Use knowledge of specialist conventions to assist making, to introduce improvements and explain what they are doing.

Level 10

a. Use a range of techniques, processes and resources with confidence, safety and creativity to achieve high quality work.

b. Review the design proposal during planning and making and show resourcefulness and adaptability in modifying the design in the light of constraints to make a high quality product.

Attainment target 4- Evaluating

Pupils should be able to develop, communicate and act upon an evaluation of the process, products and effects of their design and technological activities and of those of others, including those from other times and cultures.

Level 1

a. Describe to others what they have done and how they have done it.

b. Describe to others what they like and dislike about familiar artefacts, systems or environments.

Level 2

a. Discuss with teachers and others how satisfied they are with their design and technological activities, taking into account their original intention and how they went about their task.

b. Make simple judgements about familiar artefacts, systems or environments, including those from other times and cultures.

Level 3

a. Discuss their design and technological activities and outcomes with teachers and others, taking into account how well they have met the needs of others.

b. Comment on the materials and processes used and how the task was tackled.

Level 4

a. Review the ways in which their design has developed during the activity, justifying decisions and appraising results in relation to intentions.

b. Review the decision making process they used in producing their final artefact, system or environment.

c. Comment upon existing artefacts, systems or environments, and those from other times and cultures, including appearance and use of resources.

d. Understand the social and economic implications of some artefacts, systems or environments.

Level 5

a. Evaluate their product in relation to the design intentions and to the original needs or opportunities, taking into account users views, cost effectiveness and scale of production.

b. Justify the ideas, materials, components, procedures, techniques and processes used and indicate possible improvements.

c. Understand that artefacts, systems or environments from other times and cultures have identifiable characteristics and styles, and draw upon this knowledge in design and technological activities.

Level 6

a. Review the original needs and opportunities originally identified and decide if they are appropriate.

b. Devise and carry out ways of testing the extent to which the product satisfies the design intentions.

c. Evaluate the ways in which materials have been used.

d. Evaluate the procedures, techniques and processes used and indicate possible improvements.

e. Illustrate the economic, moral, social and environmental consequences of design and technological innovations including some from the past and other cultures, using specific examples.

Level 7

Present an evaluation of their activities against the original need, drawing on information gathered about the product and the reactions of users. Evaluation should include suggestions for improvement.

Level 8

a. Present an evaluation of their activities, including suggestions for improvements and a discussion of: (i) the relationship between the materials chosen and the procedures, techniques and processes used; (ii) justification of possible improvements; (iii) the suitability of the product for manufacture; (iv) an estimate of the effects and consequences, including environmental and economic ones

b. Understand that artefacts, systems or environments reflect the circumstances and values of particular cultures and communities.

Level 9

Demonstrate that they have applied knowledge and understanding derived from evaluations of their own and others' design and technological activity.

Level 10

a. Demonstrate through their choice of working methods and discernments and flair in decision taking, the quality of their design and technological capability.

b. Evaluate artefacts, systems or environments to show the interaction of influences on their developments and use this knowledge in their work.

'A Course of Studies in Design' - CSE Examination

North West Examinations Board (NWEB)
This examination was devised by the Schools Council Project in Design and Craft Education it was based on the following abilities:

Twelve general abilities:

1. Analyse a situation in order to identify a need or problem area.
2. Identify specific factors affecting the need or problem area (factors may include social, ergonomic, functional, aesthetic, material and cost elements).
3. Relate factors identified (synthesis) to define precisely a brief, specification or design problem.
4. Gather from a variety of sources specific information related to a brief, specification or design problem.
5. Produce outline solutions which satisfy the requirements identified in a design problem.
6. Make valid and logical selections from all known alternative solutions.
7. Specify a procedure necessary for the production of a solution.
8. Apply relevant motor skills to produce a solution.
9. Judge a solution in terms of the design brief or speculation.
10. Record information and evidence of observations, investigations and decisions.
11. Communicate ideas/solutions clearly in graphic, written or 3D form.
12. Apply knowledge of tools, materials, techniques and principles.

GCSE : Craft, Design and Technology courses

Statement of Educational aims

1. To foster awareness, understanding and expertise in those areas of creative thinking which can be expressed and developed through investigation and research, planning, designing, making and evaluating, working with materials and tools.
2. To encourage the acquisition of a body of knowledge applicable to solving practical/technological problems operating through processes of analysis, synthesis and realisation.
3. To stimulate the development of a range of communication skills which are central to design, making and evaluation.
4. To stimulate the development of a range of making skills.
5. To encourage students to relate their work, which should demand active and experiential learning based upon the use of materials in practical areas, to their personal interests and abilities.
6. To promote the development of curiosity, enquiry, initiative, ingenuity, resoucefulness and ingenuity.
7. To encourage technological awareness, foster attitudes of cooperation and social responsibility and develop abilities to enhance the quality of the environment.
8. To stimulate the exercising of value judgements of an aesthetic, technical, economic and moral nature.

Assessment Objectives

Candidates should be able to:

1. describe and apply facts, principles and concepts related to artefact and or/or systems design, realisation and evaluation;
2. demonstrate graphical and other communication skills necessary to give, in a clear and appropriate form, information about an artefact or system;
3. identify problems which can be solved through practical/technological achievement;
4. analyse problems which they have identified, or which have been posed by others, and produce appropriate design specification taking into account technical and aesthetic aspects;
5. identify the resources needed for the solution of practical/technological problems;
6. identify the constraints imposed by knowledge, resources availability and/or by external sources which will influence proposed solutions;
7. gather, order and assess the information relevant to the solution of practical/ technological problems;
8. produce and or/interpret data (eg diagrams, flow charts, graphs, experimental results);
9. produce and record ideas as potential solutions to problems;
10. appraise solutions to a design problem relative to the initial specification;
11. select and develop a solution after consideration of the constraints of time, cost, skill and resources;
12. plan the production of the selected solution;
13. demonstrate appropriate skills, make or model the artefact or system;
14. propose or make modifications to a product or system against its specification;
15. compare and evaluate the performance of an artefact or system against its specification;
16. satisfy all mandatory and other necessary safety requirements during the planning and making of an artefact or system;
17. describe the interrelationship between design/ technology and the needs of society.

Issues relating to aggregation

Any aggregation process will discard information. The best aggregation system is therefore one which keeps as much information as possible through each stage of the aggregation process. To discuss this further it is necessary to outline different types of aggregation and the criteria by which it is suggested any aggregation system should be evaluated.

Methods of aggregation

Essentially aggregation methods fall into two types:

1. Arithmetic methods: These rely on arithmetical computations on the mark distributions of components. These computations include calculating totals and averages. It is important to realise that additivity is not an automatic property and careful consideration needs to be made to ensure like is being added with like. When components are aggregated the standard deviations of the mark distribution for each component mean that certain marks lend more weight in the aggregation process than others. Such effects need to be ironed out. This is often where scaling methods are used to equate components before final aggregation.

The important assumption underlying these methods however is that the component scores represent quantities which can be arithmetically manipulated and where addition is valid.

2. Combination Methods: Combination methods usually depend on a set of rules for arriving at a final total. It is important to realise that these methods carry forward each sub-score as a single descriptor, (number or grade) thus discarding much information. The problem here however is to justify the set of rules used to arrive at a final description of performance. These might take the form of a points system where a score is carried forward as a point score, totalled and then read from a table of fixed boundaries (which itself needs careful examination as to the effects it produces in outcome distributions, and therefore justification); or they could take the form of algorithms which say that certain profiles or sub scores should receive allocated final scores according to the pre-judged quality of the profiles. The judgements on the quality of the different profiles need also to be carefully examined and justified. For example on what criteria should the award of a B be given to the profile abbd? The outcome is not geared to arithmetical computations which take into account standard deviations or any other factors, and therefore is virtually completely arbitrary. That is why combination methods need careful justification. It is also the case that combination methods operate quite wide levels of compensation between components allowing poor performance on one component to be offset by good performance on another to differing degrees. Regression to the mean should be allowed for.

Criteria for evaluating aggregation methods - 'Limen-referencing'

French et al (1987) in the DAATE project discuss grading in the context of 'Limen referencing' (the term was first used by Christie & Forest, 1981). Limen referencing refers to a mental proces of individual judgement. Taken from the arena of psychological testing it refers to the point at which the person taking the test no longer responds to a certain stimuli. The point at which the stimuli disappears IN THAT PERSON'S JUDGEMENT, is known as the 'Limen' and the stimulus is now sub-limenal.

Limen referencing provides an important characterisation of the assessment process because it emphasises the individual nature of the judgements being made. The process of assessment by markers is therefore a judgemental process. The assessment is an expert judgement of the quality of the pupil's outcome and it should be the function of any aggregation system to carry forward as much information as possible about this judgement through to the final summary descriptor.

Each time an aggregation occurs information is thrown away about the original judgements on the quality of pupil outcomes. It seems best therefore to reduce the number of aggregation stages in arriving at the final score or descriptor.

Design and Technology SATs and aggregation

The issue of agregation is a difficulty one to resolve. It is essential that any aggregation is sound and valid. The aggregation issue relating to a criterion referenced system needs careful investigation. The arithmetic methods are recommended in being best able to carry forward through the aggregation process the maximum amount of information about awarder's limen-referenced judgements, and yet, because the SAT assessments are criterion referenced the Arithmetic Methods are not necessarily appropriate. National Curriculum levels (ie component scores) do not of course have underlying mark distributions and strictly speaking they are ordinal data. The situation is therefore to make the limen-referenced judgements as valid and reliable as possible (through standardisation procedures and experience) and then to validate the system of aggregation. In this sense the final aggregated outcome is made up of two components - a limen-referenced judgement and a 'system referenced' total.

The Stranded Matrix

The stranded matrix inserts an extra layer of aggregation from strands to ATs. For other reasons the stranded matrix may be the best to use, but the aggregation system must take into account the fact that the final descriptor is, using the 50% mastery rule, two steps away from the original expert judgements at strand level.

The AT Instrument

This is more straightforward in relation to aggregation as the judgements are being made at the level of AT and there is only one aggregation process to validate. However it does limit the diagnostic assessments.

The Significance of the PC

The importance of validating the aggregation system is clear when it is considered that the PC has special significance as a representation of the outcome derived from holistic activity. The report, Design and Technology for ages 5 to 16 (June 1989 Department of Education and Science) recommends,

"the use of a single profile component called design and technology capability, which reflects the holistic activity in which pupils have engaged."

This has given this PC special significance which must be taken into account when it is determined. However the PC is not an assessment which can be made on the basis of criteria referencing as none are provided. Indeed it would be extremely difficult to provide criteria as there are many ways in which a level may be achieved, pupils will demonstrate strengths and combinations of strengths in a wide variety of ways across the attainment targets. The PC must be fair to all.

The June 89 report also offers advice on how a PC level might be determined (4.26). It gives an example of how rules might be applied to determine a PC level on the basis of AT scores, the level at which assessment is required by the Education Reform Act. This two tier level of ruling would be difficult to operate, it is unduly harsh on pupil performance and will consequently depress pupil levels of achievement significantly. It requires pupils to have a mastery of at least 75% at the level at which the PC is assigned. It will be difficult for a pupil to reflect this level of achievement in a SAT. The aggregation rule operated during the 1990 Summer trial assigns a PC at the level of 50% mastery. This is the rule which is now statute at key stage 1 (except in English where the ATs are weighted and the PC is the average of the levels obtained). From the outset it is recognised that the 50% mastery rule raised some questions but it must be seen if these were realised in operation.

Measuring within-profile variance

One approach to validating aggregated PC outcomes is to take into account the way in which pupil's performance across ATs can vary. The underlying assumption here is that the evenness of performance at AT level is desired as a reflection of holistic achievement. To test the evenness of performance we have used a measure of variance (MV). This we have defined as the sum of NC levels of AT performance which vary from the PC score awarded by the aggregation system used in the summer trial (50% mastery) - a measure of dispersion about the PC. For example:

AT scores	*PC(50% mastery)*	*MV*
4 5 5 6	5	2
4 3 5 6	5	4
3 6 5 2	5	6
2 2 5 7	5	8

These examples achieve the same PC but the measure of variance differs considerably. The premise might be made that the greater the measure of variance the more dubious is the PC award as an assessment of design and technology capability. If the second lowest score is taken (75% mastery) the measure of variance will be identical.

It has been suggested that the majority of pupils will tend to perform reasonably consistently across the ATs. The measure of variance can be used to test this hypothesis. The results from the 1990 Summer trial show that 90.3% of pupils are within a measure of variance of 4, on average one NC level per AT. When the frequency is high the measure of variance is low. However it is important to arrive at a PC which, if possible, reflects accurately the performance of every pupil. For example the pupil with a profile 0 0 5 5 would be assigned a PC of 5 using the 50% mastery rule which clearly might not represent his or her level of capability.

However, including the possibility of recording no achievement which has been part of our analysis to date, there are 1,936 unique profiles of 4 AT scores possible. The figures above are based on 197 profiles which occured in the summer trial. Although many of the 1,936 profiles will not occur in practice, and it is possible that pupils' AT profiles will tend towards evenness (ie low measures of variance scores), further investigation of the possible outcomes across possible profiles is reported below.
The AT profiles produced in the 1990 trials have been investigated using a variety

of aggregation rules and the results are presented in tabular form in appendix B.

Other methods of aggregation

1. Arithmetic methods

- The average

The use of the average in a criterion referenced system will always be contentious. This might more accurately reflect capability but it creates the problem of the PC not being a whole number - which makes no sense as a reporting device in a criterion referenced system at the level of individual reporting. To achieve a whole number a second tier of rules would need to come into operation. The debate would revolve around .5. The rule at key stage 1 indicates that, " the fraction of one half should be rounded up." We show the effects of rounding up and down in appendix B. To overcome some of the problems related to the average we have investigated the use of a conversion table.

- Conversion table

A conversion table might be less controversial and a more easily operated system as it does not require calculations to be made. Our table is based on the sum of AT scores. A proposed table is shown below which has the effect of rounding down one half fractions.

1	2	3	4	5	6	7	8	9	10
1 to 6,	7 to 10,	11 to14,	15 to18,	19 to 22,	23 to 26,	27 to 30,	31 to 34,	35 to38,	39 to 40

This table has the effect of rounding up one half scores

1	2	3	4	5	6	7	8	9	10
1 to 5,	6 to 9,	10 to13,	14 to17,	18 to 21,	22 to 25,	26 to 29,	30 to 33,	34 to37,	38 to 40

AT scores	***PC(Conver.Tab.)***	***MV***
4 5 5 6	5	2
4 3 5 6	5	4
3 6 5 2	4	6
2 2 5 7	4	8

This method of achieving a PC we believe has merit as the table has the effect of smoothing. The PC , when derived in this way, might be described as reflecting pupils' average level of capability across the attainment targets - a measure of consolidated performance. This clearly gives the PC educational significance. A conversion table might also be used to obtain both AT and PC levels from strand assessments, hence the PC score would also be only one step away from the expert judgement. It is worth examining how it deals with the attainment target profiles produced in the summer trial which have a high the effects of the rounding up and rounding down methods are reported below.

(It can be argued that PC scores based on the average do appear to more fairly represent a pupil's capability. However, it can also be argued that the aggregated score is not criterion referenced and throws away too much information about positive achievements. There is the argument that extreme low levels awarded by the SAT are an accident of the SAT, for example that a pupil has chosen an inappropriate project title or in some other way has not been provided with the opportunity to show their achievement. In this case an aggregation system which disregards extreme low levels of achievement might actually improve the reliability of the assessments. This issue needs further debate and exploration).

2. Combination Rules
The Mastery rules might loosely be considered combination methods in that they do not rely on numeric computations, but select a final PC award from the scores offered in the ATs according to to a desired level of mastery.

There are many possibilities, one suggestion is to consider each profile as a whole, and produce an expert consensus on what such a profile is worth in terms of an overall grade. This sort of approach can become quite complicated with different parameters governing the distances allowed between levels in the profile before you are allowed to move up to the next overall result and so on. The method given in the Parkes Report is of this type. We report a modified version of this rule which requires the PC to be equal to or lower than 2 of the AT scores and not 3.

Combination methods are more appropriate to Criterion referenced sysytems than to norm-referenced systems because the (AT) outcomes can be validly carried forward as single descriptions of performance as they subsume the measurement of mastery at SoA level, however such an assertion needs closer scrutiny.

Mastery can be determined at 25%, 50%, 75% or a 100% the decision for which is chosen might have an educational reason but in reality is fairly arbitrary as a good case can be made for each level. However which ever is chosen will effect the scores produced and the perceived success of individual pupils, teachers and schools. It is a commonly held view that the 25% and 50% mastery rule will lead to teachers focusing their teaching to narrow goals either because of their approach to the subject or in attempt to inflate the achievement levels of their pupils.

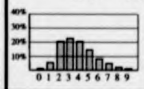
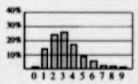
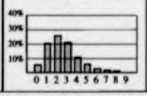
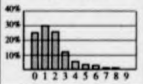
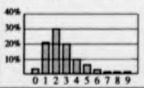
When judgements are made by parents, governors and others about the performance of subjects in a school is it reasonable to expect them to understand that the PCs have been calculated in different ways? Despite the criterion referenced nature of the National Curriculum judgements of this nature will be normative. We believe that the PC should be produced in a way which is fair to pupils and recognises their strengths. The performance of the vast majority should not be depressed because of the need to police the few. There are other mechanisms charged with overseeing the curriculum at both local and national levels.

Weighting attainment targets

There has been much discussion amongst Design and Technology specialists about the weighting of the attainment targets in determining the PC. Many advocate the importance and uniqueness of Te3 to the subject and hence believe it should have extra importance in arriving at the PC. There are many aspects of this argument we would support, not least that if assessment should in someway reflect the time spent on aspects of the task then Te3 should be more significant. A possible weighting might rank the ATs, in order of importance Te3, Te2 and Te1 and Te 4 equal (40%, 30% and 15% for the other two ATs).

Aggregation - possible methods and the implications

Attainment target levels to PC

	Numeric methods				Combination Rules				
	The average			The cumulative percentage	25% Mastery	50% Mastery	75% mastery	100% Mastery	**Modified Parkes rule** - PC one level above lowest AT score
The rule	There are many numeric methods which might be considered, the mean appears simple enough to be investigated. It would require a rule in relation to rounding up or down, both possibilities are shown below.			This rule expresses the sum of the AT scores as a percentage of the possible total - which in Design and Technology is 40	A 25% mastery rule identifies a pupil's very best achievement - this means that the PC is the same as the highest AT score.	A 50% mastery rule focuses on the better half of what a pupil has achieved, - this means that 2 or more ATs must be at the PC score.	A 75% mastery rule in effect ignores the lowest AT score, - this means that 3 or more ATs must be at the PC score.	A 100% mastery rule awards a PC at the lowest AT level achieved by the pupil - all 4 ATs are at the PC score or above.	The PC is one level above the lowest AT score provided that 2 or more AT scores are at or above the proposed PC.
Sample AT profiles with PC and MV calculation		Rounding .5 up	Rounding .5 down		Ats PC MV 4.5.5.4 = 5 2 2.3.4.5 = 5 6 0.4.5.2 = 5 9 2.3.6.7 = 7 10 5.2.2.5 = 5 6 3.4.5.4 = 4 4 2.3.6.3 = 6 10 6.3.2.5 = 6 8	Ats PC MV 4.5.5.4 = 5 2 2.3.4.5 = 4 4 0.4.5.2 = 4 7 2.3.6.7 = 6 8 5.2.2.5 = 5 6 3.4.5.4 = 4 2 2.3.6.3 = 3 4 6.3.2.5 = 5 6	Ats PC MV 4.5.5.4 = 4 2 2.3.4.5 = 3 4 0.4.5.2 = 2 7 2.3.6.7 = 3 8 5.2.2.5 = 2 6 3.4.5.4 = 4 2 2.3.6.3 = 3 4 6.3.2.5 = 3 6	Ats PC MV 4.5.5.4 = 4 2 2.3.4.5 = 2 6 0.4.5.2 = 0 11 2.3.6.7 = 2 10 5.2.2.5 = 2 6 3.4.5.4 = 3 4 2.3.6.3 = 2 6 6.3.2.5 = 2 8	Ats PC MV 4.5.5.4 = 5 2 2.3.4.5 = 3 4 0.4.5.2 = 1 9 2.3.6.7 = 3 8 5.2.2.5 = 3 6 3.4.5.4 = 4 2 2.3.6.3 = 3 4 6.3.2.5 = 3 6
	4.5.5.4 = 4.5 2.3.4.5 = 3.5 0.4.5.2 = 2.75 2.3.6.7 = 4.5 5.2.2.5 = 3.5 3.4.5.4 = 4 2.3.6.3 = 3.5 6.3.2.5 = 4	4.5.5.4 = 5 2.3.4.5 = 4 0.4.5.2 = 3 2.3.6.7 = 5 5.2.2.5 = 5 3.4.5.4 = 4 2.3.6.3 = 4 6.3.2.5 = 4	4.5.5.4 = 4 2.3.4.5 = 3 0.4.5.2 = 3 2.3.6.7 = 4 5.2.2.5 = 3 3.4.5.4 = 4 2.3.6.3 = 3 6.3.2.5 = 4	4.5.5.4 = 45% 2.3.4.5 = 35% 0.4.5.2 = 28% 2.3.6.7 = 45% 5.2.2.5 = 35% 3.4.5.4 = 40% 2.3.6.3 = 35% 6.3.2.5 = 40%					
Mean PC and *MV* if this rule is applied	**3.8**	**4.3** ***5.1***	**3.4** ***5.0***	**38%**	**5.5** ***6.9***	**4.5** ***4.9***	**3.0** ***4.9***	**2.2** ***6.6***	**3.2** ***5.2***
If this rule had been used in the 1990 trials (1551 cases) — **PC/*MV***	**2.71**	**2.83** ***/2.07***	**2.60/*2.1***	Not considerd as reporting would not be in NC terms.	**3.73** ***4.1***	**3.07** ***2.78***	**2.50** ***2.78***	**1.52** ***4.76***	**2.52** ***3.49***
Profile of PC scores									
Implications General If NC assessment is to be intelligable to those outside education the following two questions must be considered: What will be the effect of different aggregation rules for different subjects when determining the PC from ATs? If subjects do have different aggregation rules what would be the implications of having different rules at different key stages?	Numeric rules which treat the AT scores merely as numbers which can be manipulated, rather than professional judgements, are perhaps more appropriate to a norm-referenced system.				Any combination rule is more appropriate to a criterion referenced system as it is derived from the professional judgements made by teachers and the relationship to the ATs is apparent.				
	The average focuses attention on the level at which a pupil can typically be expected to perform.			Assuming that the ultimate objective is to reach 10, a PC as a percentage expresses the amount of the curriculum in which the pupil has proved their capability. All achievement is reflected in the score.	Any of these rules might be thought to allow either teachers or pupils to effectively, 'drop' one or more of the attainment targets. This belief is based on the notion that the SAT is the principle method of policing the National Curriculum when in reality there is an infra-structure at both local and national levels charged with doing this.				Two tier rules might be more difficult to operate however data handling systems can overcome practical issues.
	These methods encourage the maximising of AT totals and hence effort in every AT.					Selecting the second highest At score is a method which disregards extreme data either high or low	These rules result in low mean PCs. It might be claimed that this protects the integrity of the subject and that the resulting PC reflects holistic capability. They all emphasise a pupil's weaknesses and disregard their strengths/areas of expertise. If Design and Technology adopts a rule different to other subjects, which depresses PC performance, there may be serious consequences when mean aggregated PC sores enter the public domain. Will parents, governors etc. understand scores in one subject which are significantly lower than in other subjects? A rule which focuses on pupils' weaknesses makes no allowance for the individual who was not motivated by the SAT, missed a number of lessons (a fairly common occurrence), or for some other reason did not do themselves justice in a SAT		
						These two rules produce the lowest MV if evenness of performance across Ats is the criteria			

Introduction

1.1 This proposal is made by the Design and Technology centre at Middlesex Polytechnic and is informed by the findings of MEGNAP's two year Technology SATs development programme from September 1989 to October 1991.

1.2 The proposal is not being made under MEGNAP auspices since MEG now has formal obligations within the wider Joint Council bid. However, if the Middlesex Polytechnic bid succeeds, we would wish to continue our close and productive collaboration with both MEG and WJEC as well as with the individual examiners and examination board officers named in 7.1.

1.3 All schools reported benefits from their involvement with our previous work because the trial SATs and associated INSET provided the first and, in many cases, *only* in-service opportunity to accommodate Technology as a new curriculum subject. Many of the original problems and challenges of two years ago remain firmly in place. The past two years have provided us with an in-depth perspective on the issues concerning NC assessment. Technology, unlike the core subjects, requires more support during the early years of implementation but as it becomes firmly embedded in the curriculum the necessity for certain materials may disappear. There are already some teachers who do not require additional support but they are in the minority. This submission is based on what we see as the current reality and the capability of a typical teacher operating in this area of the curriculum.

1.4 If the timetable for national curriculum implementation is to be adhered to, further SATs development must continue to lend teachers and pupils support in coming to terms with Technology as well as providing for a straightforward, realistic and manageable way of assessing it. Our work clearly shows that if SAT development is to proceed smoothly it is essential that both teachers and pupils are provided with a structured framework. This must be *neither elaborate nor expensive* to implement yet must offer pupils the degree of autonomy implied by SoAs within the Order. All of the proposals that follow are also guided by the need for SATs to be *manageable* and *workable* by teachers *without entailing an unrealistic assessment load* during or at the end of the Key Stage.

1.5 There has been little, if any, prior research in relation to Technology and special educational needs. In terms of pupil numbers and the variety of special needs, this is clearly a major issue and one which we will continue to address.

1.6 The objective of the Middlesex Polytechnic team is to write differentiated test materials for both the long task and short test for D&T and IT. In addition, materials to promote standardisation and mark schemes which illustrate the operationalisation of the SoAs, will need to be produced to support exemplification material produced by SEAC. Our objective for 1992 is to operate a system which will have all the characteristics of the full national pilot in 1993.

1.7 We believe that the assessment of technological capability should encourage pupils to be active participants. Assessment should focus on the skills, knowledge and capabilities generic to technology and their application to genuine contexts. Our aim is to develop SATs which will enrich pupils' technological experience.

1.8 It is proposed for the long task that there will be differentiated materials serving levels 1-4; 3-7 and 6-10. This pattern is repeated in the short task. The bands provide for sufficient overlap to avoid cliff and ceiling effects while at the same time providing teachers with three manageable categories of classification. Information from our pilot indicates that the majority of pupils are currently operating below level 6. In the future we would expect the majority, at the end of KS3, to be undertaking tasks in the 3 - 7 range. Although two of these tests cover 5 levels, we believe this is appropriate given current levels of performance and the need to assess capability. This banding may need to be reviewed in the future.

1.9 We advocate the use of a stranded system of SoAs for both design and technology and information technology. These matrices/templates will also be valid for TA of the long task. By organising the SoAs into generic strands, both systems make it possible to assess and diagnose achievement. In design and technology the stranded matrix we have

developed organises the SoAs into eleven *generic* strands. In information technology, our assessment template has five strands.

1.10 The development team will aim to capitalise on the expertise acquired during the previous development phase. Details of the personnel involved is provided in appendix 1. The team will be based at Middlesex Polytechnic. The infrastructure used during the 1991 pilot will be employed in the production and distribution of SATs during the 1992 pilot. Our proposed programme for the 1992 pilot is shown in appendix 5.

Basis for the 1992 Pilot

2.1 This bid is based on the following number of schools for 1992

Design and technology		Information technology	
SCU	63	SCU	53
Wales	12	Wales	12
Special	10	Special	10
Total	**85**	**Total**	**75**

2.2 This bid includes no provision for 'Warm schools'. We are not convinced of the value of including such schools in the context of this pilot or the FUR. Schools willing to undertake this role, have been identified and can be included, at an additional cost, if requested.

2.3 In addition, a cohort of schools including City Technical Colleges, Independent and Grant maintained schools will be invited to take part in the pilot. This will add a further 8 schools to the above total. This is being done to provide representative evaluation data.

Content of Technology SATs

Design and technology ATs 1 - 4

3.1 **The Long task** - The differentiated element of the long task will be a procedure based on our label system developed over the past two years. It is not envisaged that pupils will be supplied with labels but rather with an operational framework, appropriate to their capability, from which to work. Teachers will be provided with clear and simple guidance about which level band should be undertaken by pupils. Appendix 3 shows the three differentiated frameworks, additions from one band to the next are shown by a different colour.

3.2 In principle, it is important that any chosen theme should be capable of 'ownership' by specialists within the Design & Technology federation. Middlesex Polytechnic has produced a number of potential themes that meet this criterion and are capable of operationalisation to ensure rigorous and standardised activities for Key Stage 3 SATs. Since it has proved difficult to set a single theme that provides a manageable focus for both teachers and pupils, it is proposed that within each theme a set of theme-related *tasks* is prescribed and that either teachers or pupils will select one of these as the basis for the SAT.

3.3 A theme which we propose for 1992 is **Identity.** This has been chosen because we believe it offers scope in all areas of design and technology. It also has a degree of novelty - and repeating a theme previously undertaken by pupils is undesirable. In our survey of themes being used by schools, this title did not occur. 'Identity' also allows attention to be drawn to certain aspects of the PoS and SoAs which currently schools find difficult to integrate into their programme. We are firmly convinced from the 1991 pilot that within a theme a task bank should be provided. An example of the tasks associated with 'Identity' are listed below; these also have the merit in asking pupils to undertake a task from a designerly perspective. The eight items of the PoS identified apply to all tasks but under each we have noted the two key aspects which underpin that task.

Product identity

Select a product range and investigate possible new products which could be added to the range.

- *recognise that the preference of consumers can change;*
- *investigate existing solutions to design and technological problems when developing ideas for new ones.*

National Identity

Choose a country and investigate ways of promoting its national identity for tourists.

- *consider the influence of advertising on consumers*
- *use information sources in developing their proposals*

Cultural Identity

Investigate ways in which the achievements of a particular culture can be shared more widely.

- *know that aesthetic qualities influence consumers' choices;*
- *take account of human scale and proportion when designing.*

Corporate Identity

Choose a team, club or organisation you belong to and investigate ways in which its identity could be improved.

- *analyse a system to determine its effectiveness and suggest improvements;*
- *identify markets for goods and services and recognise local variations in demand*

3.4 The accompanying long task - appendix 4 - project folder illustrates how the theme and tasks would be introduced to pupils. This folder, if the bid is successful, will be used for the Autumn trial. The long task has been designed so that it will require teachers to address specific aspects of the programmes of study.

3.5 The programmes of study largely involve the specification of behavioural outcomes and the majority of these are generic to design and technology. The principal non-generic items are related to knowledge of 'mechanisms' and 'energy'. Because of this, it is proposed that PoS identification prior to a SAT should *emphasise* generic headings where appropriate and indicate which items "when working towards higher levels" should be considered in relation to the differentiated tasks. Appendix 2 illustrates clearly how this has been constructed.

3.6 **The Short task** - There will be three differentiated tests covering the same bands as the long task. Each test will consist of questions relating to the theme. In the short task a strand will be selected in each AT which will provide the focus for assessment. The four strands selected would be associated with four groups of questions each designed to occupy about 20 minutes of time. We do not propose that the short task should require pupils only to recount and reflect on what took place during the long task - although it will provide a firm base of experience on which to draw. The short task can create new contexts within the theme which will motivate pupils and encourage performance.

3.7 This is best illustrated using an example. This example also shows how a skill such as interviewing might be dealt with in a task of this type.

Short Task - levels 3 to 7 AT1 strand - investigation SoAs 3a, 4 b and e, 5 a and b, 6 c and 7 b and d

Situation

- *People often lose things in school.*
- *Everyone has decided a new lost property system is required.*
- *A key aspect will be knowing the* ***identity*** *of the owner when something is handed in as lost.*

Activity

Before thinking of solutions more information is needed about the problem.

What information, do you think, will be important?
How will you gather the information?

Who will you talk to and consult about the issues?
How will you collect information from other people?

How would you decide which information was important?
How could technology be important in providing a solution?

3.8 The short task will exploit graded questions and timed sections. Breaking the activity down into a series of timed phases will allow information to be revealed as the activity develops. This will focus pupils thinking to address similar issues which will in turn result in more comparable answers.

3.9 The long task will reveal most about *capability* by assessing *operational performance* - especially in relation to AT3. The pencil and paper test, on the other hand, will provide the opportunity to assess *intelligent behaviour* through reflection on what has been done during the long task and applying this experience to new situations. The team's trialling of short pencil and paper tests has shown some possible ways forward. The example above provokes essentially written responses, although diagrams and flow charts could also be used. Other questions set might require drawings and sketches and might be initiated from photographs and graphic images. Practical skills such as making can be addressed only by questions which assess, theoretically, a pupil's ability to organise procedures and devise strategies to meet defined objectives.

3.10 The concept of tasks within a theme means that the SAT can be more prescriptive of context, materials and processes. Although some teachers were in favour of extremely open briefs during trialling and piloting, many specialists felt that they had been 'disenfranchised' by the single theme and that too much freedom for pupils in the beginning had caused the SAT to become unmanageable. We see no problem in allowing teachers to select one of the four tasks for their teaching group. A degree of flexibility and choice is, we believe, essential for successful delivery.

3.11 To ensure a degree of fairness and to create some degree of 'standard conditions', it will be important for all pupils in a school to undertake the long task during the same time window. We would advocate that the time window should be reasonably close to the date of the short task. For this to happen, all design and technology teachers will need to participate in delivery and assessment and all resources will need to be available. Consequently, we would define the materials required as those normally available to pupils of that age in each of the specialised environments to which they have access. Any closer definition will create serious implications for overall resource management or will create subject-specific SATs.

3.12 All ATs will be covered by the combination of long and short tasks. The short task will focus on one strand in each AT. Performance in this strand will be taken as an indication of attainment across the AT. The long task will assess capability in all 11 strands and determine a levelness of operational capability in relation to each AT. The long task must be assessed by teachers working with the pupils to ensure that 'witness evidence' is considered. This is essential if behavioural aspects are to be assessed fairly.

Ensuring that the SATs are manageable, valid and reliable

Assessment - Design and technology

4.1 Different rules can be applied to the operation of the stranded matrix (eg. 1-n or n/strand), and it is therefore extremely flexible. Despite the uneven distribution of SoAs throughout the levels, the matrix smooths out the requirements at each level and allows performance to be recognised fairly in a single activity. By requiring achievement in each strand for a level to be achieved, coverage is at a minimum 50%.

4.2 Manageability of assessment is of paramount importance and potentially the most difficult aspect of NC for teachers to come to terms with. MEGNAP has demonstrated through its developmental work that clear-cut strategies and procedures will facilitate manageability and result in standardised and consistent awards of levels. Exemplification of what a level means in the context of pupil outcomes is essential, as our pilot materials clearly show. SEAC materials will therefore be important in the context of teacher assessment. We are keen to be involved in the preparation of such materials, but undertaking this important work has not currently been costed into the project. We propose further development of the *level guide* as a means of helping teachers make these assessments. It provides an efficient first step for teachers making NC assessments for the first time.

4.3 In relation to the short test, we would anticipate that teachers will be assigned to mark only one band of tests ie. level 1 to 4 or 3 to 7. This is possible as 'witness evidence' is not an issue in relation to the short test. Teachers will, as a consequence, be making decisions about work which falls into a narrow band of achievement and thus efficiency will be maximised. Clear instructions will be provided in relation to each of the differentiated tests.

4.4 Clearly there is no guarantee that AT5 will be marked by the same teacher marking ATs 1 to 4. The target marking time of 15 minutes will be divided one third to IT and two thirds to D&T. This is clearly an extremely tight target to meet.

4.5 Our understanding of the specification is that the SAT includes both long and short tasks and that the end of key stage assessment is based on a combination of these assessments. In the course of both activities, pupils will be required to cover at least 50% of statements at the AT level which they are deemed to have achieved. We do not believe it is possible to provide valid assessments in the short task, for all ATs, at all levels, if a minimum 50% coverage of SoAs is required. (For example level 4 has 19 SoAs - 27 items - across the ATs, hence a minimum of 10 SoAs - 14 items - will need to be evidenced.) Consequently, we propose that the long task provides an assessment based on all 11 strands and the short task based on 4 strands - one from each AT.

4.6 The SAT score will only be the end of key stage score if a pupil has achieved those levels or higher, via TA of the long task in the other 7 strands.

Validity and reliability

5.1 Using the definitions agreed by phase 1 developers with EMU, the validity and reliability of the new tasks and tests will be approached using a similar range of techniques as employed by MEGNAP in the pilot. These were able to demonstrate acceptable levels of validity and reliability.

Descriptive validity measures will be derived from materials validation (undertaken by experts - see KS 4 scrutineers and an evaluation of the pupils' use of the operational framework); INSET and guidance validation and SAT process/management validation using interview, observation and questionnaire techniques. Teacher and coordinator evaluation questionnaires will also provide evidence for assessing descriptive validity.

Construct validity will be addressed using evidence from an inter-AT correlational analysis.

Reliability will be addressed using a mark-remark exercise which will provide evidence for assessing consistency of assessment between and within markers using the stranded matrix in terms of the three differentiated tests. The design will need to indicate the accuracy of the banding of the test in terms of differentiation and will no longer need to monitor witness evidence effects as markers will be assigned by banded test and not by teaching group.

ATTAINMENT TARGET 1 : Identifying Needs and Opportunities appendix 4.1

Pupils should be able to identify and state clearly needs and opportunities for design and technological activities through investigation of the contexts of home, school, recreation, community, business and industry.

Can the pupil:

	IDENTIFICATION	INVESTIGATION	IMPLICATIONS
1	a. Describe to others what they have noticed in familiar surroundings or visualised about imaginary situations.		b. Suggest what might be done.
2	a. Describe what they have observed or visualised and found out in their exploration.	c. Ask questions which help them identify needs and opportunities for design and technological activity.	b. Suggest practical changes that could be made in response to a need and describe to others why they have suggested the changes.
3	a. Starting with a familiar situation, use their knowledge and the results of investigations to identify needs and opportunities for design and technological activities.	b. Develop and clarify their ideas about possible needs and opportunities through discussion with those involved.	
4	a. Starting with an unfamiliar situation identify needs and opportunities for design and technological activities. c. Recognise the points of view of others and consider what it is like to be in another person's situation.	b. Devise ways to gather information in addition to using printed sources. e. Provide oral and written justification for the conclusions they reach as a result of investigation.	d. Explain that a range of criteria which are sometimes conflicting must be used to make judgements about what is worth doing. f. Know that in the past and in other cultures people have used design and technology to solve problems in different ways.
5	a. Show judgement in choice of information in the systematic search for information, both qualitative and quantitative, in the systematic search for a need or opportunity for design and technological activity. b. Recognise that economic, social, environmental and technological considerations and preference of users are important in developing opportunities.		
6	a. Explain how they have identified needs and opportunities for design and technological activities and give a justification of the conclusions they have reached.	c. Understand how the introduction of new technologies can offer new opportunities and create new demands for design and technological activity.	b. Explain how different cultures have influenced design and technology, both in the needs met and opportunities identified.
7	d. Identify and draw upon sources of expert advice relevant to identification of needs and opportunities for design and technological activity.	c. Consider both the user and the producer when defining the need for a technological activity. b. Vary methods of investigation to obtain all the information required.	a. Analyse information of several kinds and draw conclusions about the needs and opportunities for a design and technological activity, recognising and resolving conflicting considerations about what is worth doing.
8	a. Provide a detailed evaluation, in the light of a range of considerations, of the needs and opportunities for design and technological activity.	b. Plan in detail the various stages of their investigation.	c. Investigate how needs and opportunities have led to design and technological activities in other cultures.
9	b. Review their own knowledge and draw up a strategy to exploit expert sources.	a. Demonstrate how they have devised and implemented a strategy for the investigation of unfamiliar situations which draw on their previous experience of design and technology.	
10	a. Convey, using presentation techniques matched to the audience, that the identification of needs and opportunities is justified and worth developing.	b. Elicit and interpret the perceptions, motivations and needs of people in a range of contrasting situations.	c. Make reasoned judgements about what is a subject for design and technological activity and what is better dealt with in other ways.

ATTAINMENT TARGET 2 : Generating a Design Proposal

Pupils should be able to generate a design specification, explore ideas to produce a design proposal and develop it into a realistic, appropriate and achievable design.

Can the pupil:

	DEVELOPING DESIGNS	DECISION MAKING	COMMUNICATING IDEAS
1	a. Express their ideas about what they might do to meet an identified need or opportunity.		
2	a. Use talk, pictures, drawings, models, to develop their design proposals giving simple reasons why they have chosen to make their design.		
3	b. Apply knowledge and skills to select ideas for different parts of their design. c. Draw from information about material, people, markets and processes and **from other times and cultures** to help in developing their ideas.	a. Make a design proposal by selecting from their ideas and give reasons for their choices.	d. Use models including annotated drawings and three dimensional working models to develop their design. e. Record how they have explored different ideas about a design and technological proposal to see how realistic they might be.
4	b. Review their design proposal to identify where decisions still need to be made; suggest possible courses of action which will improve their proposal.	c. Estimate the resource requirements and check on availability.	a. Record their ideas as they develop. d. Describe and edit design proposals.
5	b. Extend their first ideas by combining various aspects of them to formulate a design proposal and explain why some ideas were not used. c. Seek out and organise information to help them develop their ideas and refine their design proposal.	d. Establish and check the availability of the resources required adapting their design as appropriate.	a. Record the progress of their ideas showing how they have clarified and developed them. e. Specify what they intend to do and what they will need by using simple plans and flow diagrams.
6	a. Produce a design specification and use it to develop their design proposal. d. Use specialist modelling techniques to develop design proposals.	c. Make judgements about realistic ways forward by exploring alternative solutions and use these to refine their design proposal.	b. Produce a design proposal recording their decisions and the ways of reaching their chosen outcomes.
7	a. Systematically seek out, appraise, organise, and use information from different sources to develop and combine ideas and judge how realistic they might be.	b. Review the detail of their design using their own experience and that of others, and suggest alternative ways of achieving what is intended.	c. Apply relevant criteria including user requirements, costs, time, skill demands, scale of production and aesthetic considerations to take decisions about the details of the design proposal.
8	b. Plan their activities to take into account multiple constraints which may at times be conflicting.	c. Show willingness to experiment and take risks subject to safety considerations recognising the implications of decisions taken in designing.	a. Record and present, using a range of methods and media, progress of their ideas; detail and refine their design proposal and incorporate modifications; use computer aided design, image generation and desk top publishing techniques where appropriate to explore detail and refine ideas.
9	a. Develop ideas by drawing on information and understanding from a broad knowledge of sources and showing judgement about the detail required. b. Refine their design to achieve an optimum practicable outcome demonstrating originality and understanding of constraints in the justification of their design.		
10	a. Provide a substantial account of the full range of ideas they have explored and the strategies used showing: (ii) Evidence that they have: - identified ways of improving and refining their proposals; - predicted with accuracy the outcomes of possible improvements and refinements; - resolved conflicting demands; - included their decisions in a coherent specification; and using an appropriate range of media and methods.		

ATTAINMENT TARGET 3 : Planning and Making

Pupils should be able to make artefacts, systems and environments, preparing and working to a plan and identifying, managing and using appropriate resources, including knowledge and processes.

Can the pupil:

	COMPETENCY WITH MATERIALS	MANUFACTURING CAPABILITY	ORGANISING AND PLANNING
1	a. Use a variety of tools and equipment to make simple things.		
2	b. Use knowledge of the working characteristics of materials and components including construction kits in making artefacts, systems or environments.	c. Show that they can use simple hand tools, materials and components.	a. Describe to others how they are going about their work.
3	b. Choose resources for making by using their knowledge of the characteristics of materials and components.	c. Use a range of hand tools and equipment, appropriate to the materials and components with some regard to accuracy and quality. d. Improvise within the limits of materials, resources and skills when faced with unforeseen difficulties.	a. Consider constraints of time and availability of resources in planning and making.
4	a. Adopt procedures which minimise waste, pay regard to cost and achieve accuracy and finish.	c. Choose tools, equipment and processes suitable for making their design and use these appropriately. e. Use drawings, diagrams and models to assist making.	b. Work with others in the planning and apportioning of tasks. d. Adopt alternative ways of carrying forward their plan when difficulties are encountered and recognise when help is needed.
5	b. Use a knowledge and understanding of the properties of a range of materials in their planning and making. d. Apply knowledge of materials, components and processes to overcome problems as these arise.	c. Demonstrate by their choice and use of a variety of equipment that they understand the principles upon which these work and the requirements of safety and accuracy.	a. Identify stages in making and coordinate these into a simple plan to ensure efficient use of time, materials and labour.
6	b. Combine knowledge of the properties of a range of materials and processes to identify those most suitable for their design. e. Show judgement in seeking advice and information.	c. Demonstrate by their choice and use of a variety of tools and equipment, that they understand the limitations of them and the need for safety and accuracy. d. Use knowledge of materials, components, tools, equipment and processes, to change working procedures to overcome obstacles as making proceeds.	a. Plan and organise making in order to achieve the desired outcome. f. Use knowledge of technical and symbolic representations of materials, components and processes to assist making.
7	b. Demonstrate competence in the use of general planning and making skills as a result of understanding the materials, components, tools and equipment and the scale of production.	a. Plan and carry out working procedures to match the constraints on making to overcome problems and to achieve the desired quality.	c. Use a range of technical, symbolic and other means of representation to assist in planning, organising, making and incorporating necessary modifications.
8	a. Review how to make best use of materials, procedures, tools and equipment.	b. Show evidence of knowledge of making processes and devise and implement procedures for quality assurance.	c. Identify and incorporate modifications during making.
9	b. Demonstrate how they have overcome constraints in planning and making to achieve a quality product.	a. Make judgements about the quality and usefulness of sources of advice and information consulted during planning and making.	c. Use knowledge of specialist conventions to assist making, to introduce improvements and explain what they are doing.
10	a. Use a range of techniques, processes and resources with confidence, safety and creativity to achieve high quality work.	b. Review the design proposal during planning and making and show resourcefulness and adaptability in modifying the design in the light of constraints to make a high quality product.	

ATTAINMENT TARGET 4 : Evaluating

Pupils should be able to develop, communicate and act upon an evaluation of the process, products and effects of their design and technological activities and of those of others, including those from other times and cultures.

Can the pupil:

	EVALUATING OWN OUTCOME	APPRECIATING AND APPRAISING
1	a. Describe to others what they have done and how they have done it.	b. Describe to others what they like and dislike about familiar artefacts, systems or environments.
2	a. Discuss with teachers and others how satisfied they are with their design and technological activities, taking into account their original intention and how they went about their task.	b. Make simple judgements about familiar artefacts, systems or environments, including those from other times and cultures.
3	a. Discuss their design and technological activities and their outcomes with teachers and others, taking into account how well they have met the needs of others. b. Comment on the materials and processes used and how the task was tackled.	
4	a. Review the ways in which their design has developed during the activity, justifying decisions and appraising results in relation to intentions. b. Review the decision making process they used in producing their final artefact, system or environment.	c. Comment upon existing artefacts, systems or environments, and those from other times and cultures, including appearance and use of resources. d. Understand the social and economic implications of some artefacts, systems or environments.
5	a. Evaluate their product in relation to the design intentions and to the original needs or opportunities, taking into account users views, cost effectiveness and scale of production. b. Justify the ideas, materials, components, procedures, techniques and processes used and indicate possible improvements.	c. Understand that artefacts, systems or environments from other times and cultures have identifiable characteristics and styles, and draw upon this knowledge in design and technological activities.
6	a. Review the original needs and opportunities originally identified and decide if they are appropriate. c. Evaluate the ways in which materials have been used. d. Evaluate the procedures, techniques and processes used and indicate possible improvements.	e. Illustrate the economic, moral, social and environmental consequences of design and technological innovations including some from the past and other cultures, using specific examples. b. Devise and carry out ways of testing the extent to which the product satisfies the design intentions.
7	a. Present an evaluation of their activities against the original need, drawing on information gathered about the product and the reactions of users. Evaluation should include suggestions for improvements.	
8	a. Present an evaluation of their activities, including suggestions for improvements and a discussion of: (i) the relationship between the materials chosen and the procedures, techniques and processes used; (ii) justification of possible improvements; (iii) the suitability of the product for manufacture; (iv) an estimate of the effects and consequences, including environmental and economic ones	b. Understand that artefacts, systems or environments reflect the circumstances and values of particular cultures and communities.
9	a. Demonstrate that they have applied knowledge and understanding derived from their own and others evaluations of their design and technological activities.	
10	a. Demonstrate through their choice of working methods and discernments and flair in decision taking, the quality of their design and technological capability.	b. Evaluate artefacts, systems or environments to show the interaction of influences on their developments and use this knowledge in their work.

Middlesex Polytechnic/MEGNAP
National Curriculum Project for Design & Technology
Key Stage 3 - Standard Attainment Task (SAT)

Assessment - a model for trialling

This assessment model was produced by mapping across the statements of attainment, from the Statutory Orders, which revealed distinct areas of technological competency. An assessment matrix was then produced with these strands arranged as vertical columns and the levels of attainment as the horizontal bands. The matrix consists of a definition of achievement against every strand at every level. We feel such a scheme will be easier to operate, as teachers will be making judgements in a consistent fashion about these eleven aspects. It will also be more helpful as a diagnostic tool. The identified strands of competency are:

	Attainment Target
Identifying possibilities Carrying out investigation Recognising implications	AT1
Developing designs Decision making Communicating and modelling ideas	AT2
Competency with materials Manufacturing capability Organisation and planning	AT3
Evaluating own technological activity Appreciating and appraising technology	AT4

In each strand of competency there is a reasonably unambiguous question at each level, to which, with experience and training, teachers will be able to answer either 'yes' or 'no'.

Assessment - what is involved?

Your task will be to decide on the level of achievement each pupil has reached in relation to each of these eleven strands. In most cases you can begin by looking at levels 3 - 7, the levels where most Y9 pupils will be. Once these levels have been determined the next step is to aggregate these levels to obtain the Attainment target level. For active research you should award the highest level achieved in a majority of strands for each attainment target.

For example, if a pupil is awarded levels of 7, 5 and 4 for the three strands in attainment target 1, their aggregated level for AT 1 would be 5.

In practice this means that the level awarded will be the middle of the three figures for Attainment targets 1, 2 and 3. For attainment target 4, which only has two strands, the level awarded will be the level awarded for the lower of the two strands.

When aggregating from Attainment targets to Profile Component (PC), apply the principle adopted by SEAC for producing PCs in all subjects - the 50% mastery rule. The PC indicates the level at which the pupil has achieved half or more of the attainment targets. To apply this rule take the lower of the two highest AT levels and this is the achieved PC level.

For example if a pupil achieves levels of 5, 7, 4 and 3 , for the four respective attainment targets, the pupil will have obtained a PC of 5. Or if a pupil achieves levels 3, 4, 4 and 2, the pupil will have obtained a PC of 4.

In practice this means that the PC awarded will be the second highest of the levels obtained by the pupil for the four attainment targets. It may seem unlikely that a pupil will address all of the relevant statements of achievement in a single SAT. However if a SAT is a design-based long task - as the SAT you will be trialling is - there should be evidence of a pupil's achievement in each of the eleven strands.

As with any criterion related system it is the interpretation of the statements which is the key to standardisation. This will be the most critical aspect of the assessment process. The work which is undertaken during the active research phase will be the first opportunity we will have to examine pupils' current levels of achievement at KS3. Your pupils' work will form the starting point for case studies of what can, in reality, be achieved.

Assessment Procedure - recorded on Megnap data forms
An essential principle to remember is that each pupil is being assessed against the criteria on the assessment matrix. Care must be taken to avoid judging pupils against each other. The progression guide will help you to achieve an overview and offers a simplification wherever you are uncertain about what is required.

It is important to remember that you will be observing what your pupils are doing and how well they are progressing throughout the SAT. Some of these observations will be important in helping you decide on the levels which you award. Some of the statements in the assessment matrix deal with verbal communication. These are indicated with a symbol depicting a teacher talking to a pupil. They are generally found in levels 1 to 4 and in attainment targets 1 and 4. Others rely on your observation of the way the student is operating. This is the case with some of the statements which relate, for example, to manufacturing capability; these are indicated by the eye symbol.

Whenever possible ask a pupil to record conversations with you in their project diary, this will avoid you having to keep a separate record. If a pupil has difficulty in doing this you can record the pupil's responses in the diary for them, this will possibly only apply for SEN pupils. Evidence of a pupil's manufacturing capability is usually fairly obvious in what they produce. However there are aspects, such as safety, which are not. If you need to intervene in what a pupil is doing again ask them to record what has occurred. If a pupil has difficulty or you feel the issue needs to be recorded formally then make the appropriate comment in the project diary. The folder on 'Record Keeping' in your teacher's kit offers further advice .

Initially we would suggest that, as a first step, you assess each pupil's work independently. You might then wish to discuss your assessment with a colleague or seek guidance from your LEA Development Officer. Start by ensuring that you have all the pupil's work.

As with any design project the first step is to gain an holistic view of what the pupil has achieved. Start by familiarising yourself with the contents of the pupil's project folder, trying to gain an insight into what the pupil has set out to do, how s/he has gone about it, and what s/he thinks of what s/he has done.

Each of the eleven strands then needs to be assessed independently. You have a mark sheet in the Megnap data forms. Look through the folder for evidence of the particular strand which you are assessing. When you are satisfied that you have seen all the evidence for the pupil's achievement in that strand refer to the assessment matrix. Look through the statements and identify the one which describes most closely what the pupil has done.

The importance of the labels
You will probably find the headings and 'structure' created by the project labels useful when you come to assess and when you begin to search for evidence of performance in a particular strand. There is not a direct correlation between labels and Attainment Targets or strands but some labels relate more directly to particular Attainment Targets or stages in the design process, and the labels are an attempt to standardise the issues pupils might address as their projects develop.

When looking for evidence for assessment purposes, the labels will sometimes help you identify the relevant material. At other times, a pupil's failure to use a label may indicate that a certain issue has not been addressed.
Much of the time, you will be assessing by using the assessment matrix and searching for evidence of achievement at a particular level within a strand. At times, however, you may be able to find accurate indications of performance level by working from the material under a label heading to the assessment matrix.

Although evidence found under a particular label is likely to address issues within a number of strands, labels will help you to locate categories of evidence. A response to some of the more 'open' statements, may give you a broad impression of the level on which the pupil is operating. The strand selected should be marked. Be prepared to adjust your selection until you have confirmed the highest level of achievement which the pupil has reached.

Carry out the same procedure for each of the other strands. When you have the levels attained in each of the strands record them on the Megnap data forms. Following assessment of the eleven strands the aggregation rules can be applied to establish the National Curriculum levels achieved for ATs and PC and these results also should be recorded on the Megnap data form.

Appendix 4.2

Attainment Target 1

appendix 4. 2

Identifying Needs and Opportunities

	Identification - can the pupil:	Implications - can the pupil:	Investigation - can the pupil:
1	- describe to others what has been seen in familiar surroundings or visualized about imaginary ones?	- suggest a practical change to what they have seen?	- ask questions that show a curiosity about familiar surroundings?
2	- following exploration, describe what has been observed or visualized?	- suggest practical changes that could be made in response to a need?	- ask questions that help to identify needs and opportunities?
3	- as a result of investigation into something familiar, use their knowledge to identify needs and opportunities?	- give reasons to others about changes they have suggested?	- develop and clarify their ideas about possible needs and opportunities through discussion with those involved?
4	- as a result of investigation into something unfamiliar, use their knowledge to identify needs and opportunities?	- recognise the points of view of others and think about what it is like to be in another person's situation?	- devise ways of gathering information and provide oral and written justification for conclusions reached as a result of their investigations?
5	- explain how they have identified needs and opportunities?	-recognise that economic, social, environmental and technological preference of users are important in developing opportunitiest?	- show judgement in choosing information (qualitative and quantitative) in systematically searching for a need or opportunity?
6	- justify needs and opportunities they have identified?	- understand how the introduction of new technologies can offer new opportunities and create demands?	- vary the methods of investigation to obtain all the information required ?
7	- recognise and resolve conflicting considerations after analyzing different kinds of information about needs and opportunities?	- consider both the user and the producer when defining the need for a technological activity?	- identify and draw upon sources of expert advice relevant to their investigations?
8	- provide a detailed evaluation, in the light of a range of considerations, of needs and opportunities?	- investigate how needs and opportunities have led to design and technological activities in other cultures?	- plan in detail the various stages of their investigation?
9	- demonstrate that consideration has been given to cost-effectiveness, optimising production, and potential sales?	- make reasoned judgements about what is a subject for design and technological activities and what is better dealt with in other ways?	- review their knowledge base and devise a strategy to exploit expert sources?
10	- match presentation techniques to their audience when justifying needs and opportunities worth developing?	- make reasoned judgements about the moral and ethical implications of design and technological activities?	- identify and interpret the perceptions, motivations and needs of people in a range of contrasting situations?

Pupils should be able to identify and state clearly needs and opportunities for design and technological activities through investigation of the contexts of home, school, recreation, community, business and industry.

Level Achieved			

Generating a Design Proposal

Pupils should be able to generate a design specification and explore ideas to produce a design proposal and develop it into a realistic, appropriate and achievable design.

	Attainment Target 2		
	Developing Designs - can the pupil:	**Decision making - can the pupil:**	**Communicating Ideas - can the pupil:**
1	- make progress on the basis of trial and error?	- make a choice about what they are going to do?	- express their ideas about what they might do?
2	- use gathered information?	- give simple reasons why they have chosen to make something?	- use talk, pictures, drawings and models to develop their design proposals?
3	- use and apply information gathered including that from other times and cultures to help develop their ideas?	- make a realistic design proposal by selecting from their ideas and giving reasons for their choices?	- record how they have explored different ideas using annotated drawings and working models to develop their design proposals?
4	- investigate, use and apply information so as to suggest courses of action to improve their original proposals?	- estimate resource requirements and check on availablity?	- record systematically how their ideas have developed and been edited?
5	- refine and clarify their design proposal explaining why some of their ideas were not used?	- establish and check on the availablity of the resources, adapting their designs as appropriate?	- produce simple scale diagrams or flow diagrams to show what they intend to do?
6	- use specialist modelling techniques to develop and refine their design proposal ?	- make judgements about realistic ways forward by exploring alternative solutions?	- produce an outline design specification and use it to develop their design proposal ?
7	- systematically appraise information to develop and combine ideas?	- devise and apply criteria to make detailed decisions?	- formalise a detailed design specification recording their decisions and ways of reaching their chosen objectives?
8	- review their activities to take into account multiple, conflicting constraints?	- manage experimentation effectively?	- present systematically their detailed ideas and modifications using a range of methods and media?
9	- know when they have enough information of sufficent accuracy and detail for the next stage of development of their design proposal?	- show effective decision making to achieve an optimum practical outcome?	- communicate ideas in a way that uses media realistically and with originality?
10	- carry out a thorough investigation of existing artefacts, systems or environments to aid the development their own ideas?	- provide comprehensive, reasoned evidence to justify any decisions taken?	- provide a substantiated account of the full range of ideas they have explored and the strategies used?

Pupils should be able to prepare a plan to achieve their design, and to identify, manage and use appropriate resources, including knowledge and process, in order to make artefacts, systems and environments.

Planning and Making

Attainment Target 3

	Competency with Materials - can the pupil:	Manufacturing Capability - can the pupil:	Organisation and Planning - can the pupil:
1	- use a variety of everyday materials?	- make simple things with due regard to safety?	- describe how they are going about their work?
2	- use their knowledge of materials characteristics, including construction kits, to make things?	- show that they can use simple hand tools, materials and components?	- improvise within the limits of materials, resources and skills?
3	Choose resources for making by using their knowledge of the characteristics of the materials and components?	-use a range of hand tools and equipment, with some regard for accuracy and quality?	- consider constraints of time and availability of resources in planning and making?
4	- apply knowledge of materials, components and processes to overcome problems as they arise?	Use drawings, equipment and processes suitable for making their design appropriately?	Adopt alternative ways of carrying forward their plan when difficulties are encountered and recognise when help is needed?
5	- adopt procedures which minimise waste, pay regard to cost and achieve accuracy and finish?	- use their knowledge of materials and processes to make things efficiently?	- identify stages in making and co-ordinate these into a simple plan to ensure sufficient use of time, materials and labour?
6	- combine knowledge of the properties of a range of materials and processes to identify those most suitable for their design?	- demonstrate choice in the use of a variety of equipment which enhances the speed and accuracy of their work?	- use knowledge of technical and symbolic representations of materials, components and processes to assist making?
7	- demonstrate competence in the use of general skills as a result of understanding the materials, components, tools and equipment and the scale of production?	- implement working procedures to match the constraints on making to overcome problems and to achieve the desired quality?	- use a range of technical, symbolic and other means of representation to assist in planning, organising, making and incorporating necessary modifications?
8	- review how to make the best use of materials, procedures, tools and equipment?	- show evidence of knowledge of making processes and devise and implement procedures for quality control?	- identify and incorporate modifications during making?
9	- demonstrate how they have overcome constraints to achieve a quality product?	- make judgements about the quality and usefulness of sources of advice and information?	- use knowledge of specialist conventions to assist making, to introduce improvements and to explain what they are doing?
10	- use a range of techniques, processes and materials with confidence and creativity to achieve high quality work?	- use a range of techniques and manufacturing processes with confidence and creativity to achieve high quality work?	- review the design proposal in the light of constraints and show resourcefulness and adaptability in modifying the solution?

Evaluating

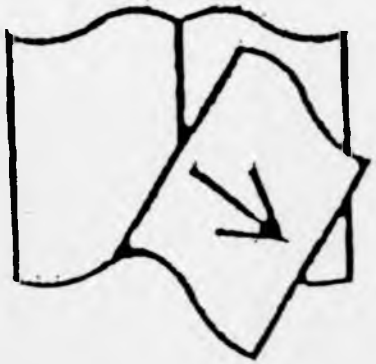
Pupils should be able to develop, communicate and act upon an evaluation of the processes, products and effects of their design and technological activities and of those of others, including those from other times and cultures.

Attainment Target 4

	Evaluating own technological activity - can the pupil:	Appreciating and appraising technology - can the pupil:
1	- describe to others what has been done and how well ?	- describe to others what they like and dislike about familiar artefacts, systems and environments?
2	- discuss their satisfaction with the outcome, taking into account original intentions and how they went about the task?	- make simple judgements about familiar artefacts, systems and environments, including those from other times and cultures?
3	- discuss their work, taking into account how well they have met the needs of others, commenting on the materials and processes in tackling the task?	- understand that an artefact, system or environment results from the culture or time from which it comes?
4	- review design development, justifying decisions and appraising results in relation to intentions and reviewing the decision making process.	- comment upon existing artefacts, systems or environments and those from other times and cultures, including appearance and use of resources, in a way that shows an understanding of the social and economic implications?
5	- justify resources and procedures used and evaluate the result taking into account users' views, cost-effectiveness and scale of production?	- understand that artefacts, systems or environments from other times and cultures have identifiable characteristics and styles, and draw upon this knowledge in design and technological activities?
6	- review the original needs and opportunities; then devise and carry outways of testing the extent to which the product satisfies the design intentions?	Illustrate the economic, moral, social and environmental consequences of design and technological innovations, including some from the past and other cultures, using specific examples?
7	- make appropriate tests of consumer response to the product, system or environment to determine value for money, effectiveness and style?	- present an evaluation of an artefact, system or environment against the original need, drawing on information gathered and the reactions of users, including suggestions for improvement?
8	- present an evaluation including discussion of the suitability of the product for manufacture and an estimate of the effects and consequences, including environmental and economic ones?	- understand that artefacts, systems or environments reflect the circumstances and values of particular cultures and communities?
9	- demonstrate the application of knowledge and understanding derived from their own and others' evaluations of their design and technological activities?	- understand that a number of interacting factors influence artefacts, systems or environments of various cultures and communities?
10	- demonstrate through the choice of working methods, and discernment and flair in decision-taking, the quality of their design and technological capability?	- evaluate systematically artefacts, systems or environments to show the interaction of influences?

PAGES MISSING
NOT AVAILABLE

283 + 284



ATTAINMENT TARGET 1: IDENTIFYING NEEDS & OPPORTUNITIES appendix 4.3

Pupils should be able to identify and state clearly needs and opportunities for design & technological activities through investigation of the contexts of home, school, recreation, community, business and industry.

FROM LEVEL 5 ONWARDS, PUPILS ARE REQUIRED TO CONSIDER EXISTING ARTEFACTS, SYSTEMS OR ENVIRONMENTS IN THE IDENTIFICATION OF THEIR OWN NEEDS/OPPORTUNITIES (see AT4 for details)

To achieve a level, pupils should be able to:

1	describe what they have noticed (familiar surroundings) or visualized (imaginary situations) and suggest what might be done.
2	describe what they have **observed** (familiar) or visualized (imaginary) or **discovered (using their senses and asking questions)** and suggest what might be done **and why.**
3	describe/record what they know (including observation, visualisation, exploration) and have discovered **through investigation (printed sources, discussion with those involved)** and identify what might be done and why.
4	describe/record what they know (in **unfamiliar** as well as familiar situations) and have discovered through investigation (including printed sources, discussion with those involved, **original sources, -ie., by devising a questionnaire to collect information)**, identify what might be done, **taking into account implications that may influence possible solutions (others' points of view, any conflicting considerations**, and an awareness that **different solutions have been reached in other times and cultures)** and **justify orally and in writing** what has been decided.
5	describe/record their investigation, **showing judgement in the systematic choice of sources of information**, identify what might be done, taking into account implications that may influence possible solutions (others' points of view, **economic, social, environmental, technological, temporal and cultural considerations)** and justify what has been decided.
6	describe/record their investigation, showing judgement in the systematic choice of sources of information, identify what might be done, taking into account implications that may influence possible solutions (others' points of view, economic, social, environmental considerations, and **particularly the influence different cultures have had on design & technology and new technologies are having on opportunity and demand)** and justify what has been decided.
7	describe/record their investigation, showing judgement in the systematic choice of sources of **varied** information and **methods (to include sources of expert advice)**, identify what might be done taking into account implications that may influence solutions (economic, social environmental, cultural, technological, but **particularly considering the user and producer)** and justify what has been decided.
8	describe/record and **plan in detail the stages of their investigation**, identify what might be done, taking into account implications that may influence solutions **(particularly considering how needs and opportunities have led to design & technological activity in other cultures)**, evaluate and justify what has been decided.
9	describe/record **how they have devised a plan/strategy for investigation, which draws upon their previous knowledge and experience of design & technology (i.e., as outlined in levels 1 - 8) and exploits expert sources**, identify what might be done taking into account implications that may influence solutions, and justify what has been decided.
10	**use presentation techniques appropriate to their audience to convey and justify what they have identified, taking into account a range of contrasting implications, and show an ability to decide whether a need is suitable to be developed into design & technological activities.**

ATTAINMENT TARGET 2: GENERATING A DESIGN PROPOSAL

Pupils should be able to generate a design specification, explore ideas to produce a design proposal and develop it into a realistic, appropriate and achievable design.

THE TERM 'VERBALLY' IS NOT USED TO DENOTE ORAL EXPRESSION BUT THE USE OF WORDS.
FROM LEVEL 5 ONWARDS, PUPILS ARE REQUIRED TO CONSIDER EXISTING ARTEFACTS, SYSTEMS OR ENVIRONMENTS IN GENERATING THEIR DESIGN PROPOSAL (see AT4 for details)

To achieve a level, pupils should be able to:

1	**express their ideas about what they might do to meet an identified need or opportunity.**
2	use talk, pictures, drawings, models, to develop their design proposals, **giving simple reasons why they have chosen to make their design.**
3	describe/record the development of their idea verbally and visually (in pictures, **annotated drawings**, models), **select from their own ideas, other designs and gathered information, use this selection, and consider how realistic their idea is.**
4	describe/record developments verbally and visually (in pictures, annotated drawings, models), select from ideas and information, **critically review their design proposal looking for improvements, and estimate resource requirements/availability.**
5	describe/record developments verbally and visually (in pictures, annotated drawings, models), select from ideas and information, critically review their design proposal looking for improvements, **refine by combining elements of their ideas with gathered detailed information, and carry the design forward into a realistic plan of resources and actions.**
6	describe/record developments verbally and visually (in pictures, annotated drawings, models **including specialist modelling techniques**), select, critically review, refine, **explore alternative solutions and further refine against those**, and make a realistic plan of resources and actions.
7	describe/record developments verbally and visually (in pictures, annotated drawings, models including specialist modelling techniques), select, critically review, refine own ideas, explore alternative solutions to further refine, **still further refine applying practical/economic/ functional and aesthetic considerations**, and make a realistic plan of action.
8	describe/record developments verbally and visually (in pictures, annotated drawings, models including specialist modelling techniques, and **using a range of methods and media, e.g., computer aided design, image generation, desktop publishing)**, select, critically review, refine own ideas, explore alternative solutions and apply practical and aesthetic considerations to further refine, **show a willingness to experiment**, and make a realistic plan **taking into account conflicting constraints.**
9	describe/record developments verbally and visually (in pictures, annotated drawings, models including specialist modelling techniques, and using a range of methods and media), select, critically review, refine own ideas, explore alternative solutions and apply practical and aesthetic considerations to further refine, show a willingness to experiment, **demonstrate originality**, and make a realistic plan **showing an understanding of constraints.**
10	**provide a thorough, coherent, substantiated record/presentation of all aspects of development** as detailed at Level 9.

ATTAINMENT TARGET 3: PLANNING AND MAKING

Pupils should be able to make artefacts, systems and environments, preparing and working to a plan and identifying, managing and using appropriate resources, including knowledge and processes.

FROM LEVEL 5 ONWARDS, PUPILS ARE REQUIRED TO CONSIDER EXISTING ARTEFACTS, SYSTEMS OR ENVIRONMENTS, WHERE APPROPRIATE, IN PLANNING AND MAKING (see AT4 for details)

To achieve a level, pupils should be able to:

1	use a variety of materials and equipment to make simple things.
2	use **knowledge of the working characteristics of materials and components (including construction kits) and simple hand tools and equipment to make artefacts, systems or environments, and be able to describe the procedure.**
3	use knowledge of the working characteristics of materials and components **to choose resources, bearing in mind constraints of time and availability**, and use a **range of appropriate** hand tools and equipment, **improvising where necessary**, to make artefacts, systems or environments **with some regard for accuracy and quality**.
4	use knowledge of characteristics of materials to choose resources, bearing in mind constraints of time and availability, **work with others in planning, adopting procedures which minimise waste**, and use a range of appropriate tools and equipment, **utilising designs (i.e., drawings, diagrams, models)**, improvising and **seeking help** where necessary, to make artefacts, systems or environments with regard for **cost**, accuracy and quality/ **finish**.
5	use knowledge of characteristics of materials to choose resources, **identify stages in making and co-ordinate these into a simple plan ensuring efficient use of time, materials, labour, demonstrate by choice and use of a variety of equipment an understanding of its working principles and safety requirements**, utilising designs, improvising and seeking help, to make artefacts, systems or environments with regard for accuracy and finish.
6	use knowledge of characteristics of a **range of** materials and of **processes to identify the resources most suitable for their design, plan and organise making**, demonstrate by choice and use of a variety of equipment an understanding of its **limitations** and safety requirements, utilising designs **(including technical and symbolic representations)**, improvising, seeking **advice and information**, to make artefacts, systems or environments with regard for accuracy and finish.
7	use knowledge of materials, processes and equipment **to demonstrate planning and making skills**, showing an understanding of **constraints** and **scale of production**, utilising designs (including **a range of** technical, symbolic **and other means of** representations), improvising and **incorporating modifications**, seeking advice and information, to make artefacts, systems or environments with regard for accuracy and finish.
8	**review how to make best use** of materials, processes and equipment, demonstrate planning and making skills, showing an understanding of constraints, scale of production, utilising designs, improvising and incorporating modifications, seeking advice and information, to make artefacts, systems or environments with regard for accuracy and finish, **devising and implementing a procedure for quality assurance.**
9	**review how to make best use of materials**, processes and equipment, **demonstrate planning** and making skills, **showing how constraints have been overcome**, utilising designs and **specialist conventions** (i.e., a report to explain what is being done or to introduce improvements), **making judgements about the usefulness of sources of information and advice consulted**, to make quality tested artefacts, systems or environments with regard for accuracy and finish.
10	**use a range of techniques, processes, and resources with confidence, safety and creativity, reviewing the design proposal during planning and making and showing resourcefulness and adaptability in modifying the design in the light of constraints, to achieve high quality work.**

ATTAINMENT TARGET 4 : EVALUATING

Pupils should be able to develop, communicate and act upon an evaluation of the processes, products and effects of their design and technological activities and of those of others, including those from other times and cultures.

FROM LEVEL 5 ONWARDS, PUPILS ARE REQUIRED TO DRAW UPON EXISTING ARTEFACTS, SYSTEMS OR ENVIRONMENTS THROUGHOUT THEIR WORK.

To achieve a level, pupils should be able to:

1	evaluate their own work, describing what they did, and say what they like and dislike about familiar artefacts, systems or environments.
2	evaluate their own work, **taking into account original intention and methods, and make simple judgements** about familiar artefacts, systems and environments (**including those from other times and cultures**).
3	evaluate their own work, taking into account original intention, **success in meeting others' needs, materials and processes**, and make simple judgements about familiar artefacts, systems and environments (including those from other times and cultures).
4	evaluate their own work, in relation to original intentions, others' needs, **reviewing their decision-making process, justifying decisions, materials and processes, and comment upon** existing artefacts, systems and environments (including those from other times and cultures), **evaluating appearance, use of resources, and understanding the social and economic implications.**
5	evaluate their own work, in relation to original intention, others' needs,**users' views, cost-effectiveness and scale of production**, justifying decisions, **ideas**, materials, processes and **indicating possible improvements, and draw upon knowledge of identifiable characteristics and styles of existing artefacts, systems and environments** (including those from other times and cultures).
6	evaluate their own work, **in relation to appropriateness** of original intention, others' needs, users' views, cost-effectiveness, scale of production, justifying decisions, ideas, **way materials were used**, processes, **devising tests of satisfaction** and indicating possible improvements, and draw upon existing artefacts, systems and environments (including those from other times and cultures), **illustrating an understanding of the economic, moral, social and environmental consequences of specific innovations.**
7	**present** an evaluation of their own activities (intentions, ideas, materials, processes, cost-effectiveness, scale of production) against the original need, **drawing on information gathered about the product and the reactions of users**, suggesting possible improvements, and draw upon existing artefacts, systems and environments (including those from other times and cultures), illustrating an understanding of the economic, moral, social and environmental consequences of specific innovations.
8	present an evaluation of their own activities (intentions, ideas, **relationship between materials and processes**) against the original need, drawing on information gathered about the product and the reactions of others, suggesting possible improvements, **discussing the suitability for manufacture and estimating the economic, moral, social and environmental consequences of their product**, and draw upon existing artefacts, systems and environments (including those from other times and cultures) **illustrating an understanding that they reflect the circumstances and values of their particular cultures and communities.**
9	**demonstrate that they have applied knowledge and understanding derived from evaluations of their own and others' work** (as detailed in levels 1-8).
10	**demonstrate the quality of their design and technology capability, through discernment and flair in decision taking, choice of working methods, and through applying knowledge of the interaction of influences evidenced in existing artefacts, systems or environments in their own work.**

Level Guide

Each level under the four ATs is consolidated into a single cogent statement that attempts to pick out the most distinctive feature of that level. It is assumed that all higher levels subsume statements below them and so features are not repeated over and over again with vague qualifications (eg., better than... more precisely.....).

In practice this means always starting at level 1 when using this guide. Work through the statements until you reach the one which has not as yet been achieved by the pupil. This guide is only intended to provide a starting point for making assessments.

It is also assumed that competencies described by each statement in an AT will have been developing through lower level work, and that by level X each emerges as a clearly demonstrable capability. For example, AT 1 level 7 - *"Analyse information and know where to find it"* - is the level at which one expects clear and unamabiguous evidence that pupils can locate and sort out information about needs and design opportunities even though there may have been earlier but less complete indications of this.

NB. Key words in BOLD.

design & TECHNOLOGY

Key Stage 3 Level Guide

AT 1

Pupils should be able to identify and state clearly needs and opportunities for Design and Technological activities through investigation of the contexts of home, school, recreation, community, business and industry.

1. **Describe, suggest and imagine** things that might be done.
2. **Ask questions** and **give reasons** for what might be done.
3. **Investigate** and **discuss familiar** situations.
4. **Understand unfamiliar** situations and **justify** doing things.
5. **Show judgement** and **recognise influences** on what might be done.
6. **Explain** thinking and **appreciate** impact of new technology.
7. **Analyse** information and know where to find it.
8. **Plan** and **evaluate** ideas in detail.
9. **Review** work and show **evidence of a strategy**.
10. Make **reasoned judgements** and **convey** these to audiences.

AT 2

Pupils should be able to generate a design specification, explore ideas to produce a design proposal and develop it into a realistic, appropriate and achievable design.

1. **Express** ideas in some form.
2. **Use various media** to develop designs - giving reasons.
3. **Develop** a design proposal **drawing on** knowledge and information.
4. **Review** and **reflect on** design proposals and resource implications.
5. **Extend** design work througha **combination** of ideas and research.
6. **Push design work forward realistically** and use **specialist modelling** methods.
7. Be **systematic** in **organising** and **combining** ideas and apply design criteria.
8. Use a **comprehensive** range of media to develop ideas and show **willingness to experiment.**
9. **Develop** and **refine** ideas to **optimise** solutions.
10. Provide a **reasoned**, **substantiated** account of design thinking.

AT 3

Pupils should be able to make artefacts, systems and environments preparing and working to a plan and identifying, managing and using appropriate resources, including knowledge and processes.

1. **Make** simple things.
2. Draw on **knowledge to make** things and describe how this is done.
3. **Use a range of** tools and equipment - recognising and **working within their limits** and those of materials.
4. **Plan making** and proceed with the aid of **drawings** and **models.**
5. **Co-ordinate making** activities and **use equipment safely** and effectively.
6. Bring knowledge and skills together **comprehensively** to facilitate making.
7. Use **specialist techniques** to assist planning and making.
8. **Review making** to **optimise** use of resources and maintain quality.
9. Make **reasoned** judgements and give **explanations** about the making process.
10. Produce **quality work** with**confidence** and **safety** and account for the design modifications during making.

AT 4

Pupils should be able to develop, communicate and act upon an evaluation of the process, products and effects of their design and technological activities and those of others, including those from other times and cultures.

1. **Describe** likes and dislikes and what has been done.
2. Discuss personal design work and **make simple judgements** about other designs.
3. **Discuss** design activities - especially meeting the **needs of others.**
4. **Review** personal design development and **comment about** influences on design in general.
5. **Evaluate** and **justify** personal design work drawing on a **wider knowledge of design.**
6. **Review** personal design work against **criteria** and **appreciate** the **wider consequences** of designing.
7. **Review** design outcomes **with reference** to the views of others.
8. **Present** a detailed and **comprehensive evaluation** of personal design work.
9. **Derive** knowledge and understanding from design evaluation.
10. Demonstrate **discernment** and **flair**in decision taking and judgement.

Summer remark

The activity was planned to assess the reliability of the stranded matrix and to evaluate and compare the three assessment instruments which had been developed:

The statements of attainment placed in strands (SoA)
The stranded assessment matrix (as used in the summer trial)(SAM)
The attainment target instrument (appendix 3b interim report)(ATI).

The three assessment instruments were used by two teachers in each of four schools. The Stranded Matrix was crossed on two groups of pupils in the design. The AT Instrument and the SOA matrix were re-marked on only one group of pupils each. The work of 80 pupils was remarked and analysis of the results allows direct comparisons using pairs of assessment devices on the same groups of pupils within schools. The assessment operation is described in the diagram below, showing the cross-matching of teachers and pupils. The detailed instructions which the teachers received are available if required.

Teacher	SAT Activity	Remark Activity		
	Stranded Matrix	Statements of Attainment	AT Instrument	Stranded (Remark)
A on Pupils A	A on A	A on A		
A on Pupils B			A on B	A on B
B on Pupils B	B on B		B on B	
B on Pupils A		B on A		B on A

It is important to emphasise the exploratory nature of this exercise in terms of its scale and the constraints of the small sample of two teachers and two groups of ten pupils in four schools. The Stranded Matrix and AT Instrument are being trialled more extensively in the Autumn (1990) but it is not a controlled study of the matrices like this one.

It is possible to approach three main issues through the re-mark exercise:

1. The 'Witness Evidence' effect.
2. The levels of agreement between marker and re-marker produced by each matrix and the correlations between them.
3. Comparison and evaluation of the three assessment instruments.

Although it was clear that some inter-school differences did exist, given the nature of the data, the tables below report the results overall for the 8 teachers and the 4 schools.

The 'Witness Evidence' effect

The table below gives the mean differences between the marker and re-marker using the Stranded Matrix for each AT and PC. Two seperate groups of pupils were involved.

The Stranded Matrix can only be used for this comparison as the first step in the re-mark activity required the teachers to enter the levels they had awarded at the end of the SAT using the Stranded Matrix. It is only these scores which therefore involve 'witness evidence'. None of the other matrices were used at the end of the SAT and the scores produced are a mixture of remembered (and forgotten) evidence after a time lapse, and the effect of using the matrix.

If the mean differences between the marker and the re-marker are positive then the marker's level is usually higher than the re-marker's level for the same pupil's work and this would illustrate that the original marks do reflect witness evidence.

	Te1	**Te2**	**Te3**	**Te4**	**PC**
Pupil Group A	0.47	0.38	0.53	0.55	0.42
Pupil Group B	0.80	0.40	0.17	0.25	0.33

Mean Differences Between Markers Using The Stranded Matrix

As all of these are positive, we can conclude that there is a 'witness evidence' effect.

Agreement Levels and Correlations Between Markers

The table below shows the level of agreement between markers using different assessment instruments on the same groups of pupils. The correlations between markers are also given for each.

Matrix		**Te1**	**Te2**	**Te3**	**Te4**	**PC**
Stranded Matrix *(Group A)*	% Agreement between marker and re-marker	50.0%	47.5%	32.5%	47.5%	37.5%
	Correlation between marker and re-marker (Kendall)	0.618	0.438	0.389	0.388	0.525
Stranded Matrix *(Group B)*	% Agreement between marker and re-marker	37.5%	45.0%	42.5%	47.5%	47.5%
	Correlation between marker and re-marker (Kendall)	0.668	0.535	0.716	0.613	0.601

Matrix		Te1	Te2	Te3	Te4	PC
AT Instrument *(Group B)*	% Agreement between marker and re-marker	45.0%	40.0%	42.5%	67.5%	40.0%
	Correlation between marker and re-marker (Kendall)	0.654	0.483	0.747	0.749	0.580
SOA Matrix *(Group A)*	% Agreement between marker and re-marker	57.5%	40.0%	25.0%	42.5%	55.0%
	Correlation between marker and re-marker (Kendall)	0.681	0.459	0.322	0.603	0.490

Percentage Agreement

The percentage agreement index shows the extent to which the marker and re-marker agreed when they used the same matrix on the same group of pupils. The SOA matrix at PC level produces the best agreement (55% of the time). However, the PC levels used in calculating the agreement are based on the aggregation system, and it may be more informative to report the mean percentage agreement over the four attainment targets. The table below gives the adjustments.

Matrix	Average Percentage Agreement over 4 AT's
Stranded (Group A)	44.4%
Stranded (Group B)	43.1%
AT Instrument	48.8%
SoA Matrix	41.3%

Mean percentage agreement between marker and re-marker over the four attainment targets.

This is a more accurate reflection of the agreement at AT level. Note that the AT Instrument produces the most agreement and the SOA the least. However, the levels of agreement do not differ widely between the matrices although in most cases the agreement is less than half the time. (See below for the number of levels difference involved in the disagreements).

Correlations

All of the correlations are significant at the .01 level and beyond. (Kendall coefficients by their nature tend to produce low looking levels, but the absolute value of the correlation is not important and attention should be paid to the significance levels).

Two further pieces of evidence are useful. First a look at the mean PC level produced by each assessment instrument (on the same group of pupils' work) and secondly a look at the range of levels difference between re-markers.

Matrix	Group A		Group B	
	Mark	***Re-mark***	***Mark***	***Remark***
Stranded	2.47	2.05	2.65	2.33
AT			2.58	2.25
SoA	2.40	2.30		

Mean PC Levels for Each Matrix on Seperate Pupil Groups

This table again illustrates the witness evidence effect between stranded matrix re-marks in that each re-mark is lower. These differences are systematic effects and not due to the matrix alone.

The differences between the other matrices mark and re-mark levels cannot be so clearly interpreted as the 'original' mark in each case is in fact a 'remark', ie, the AT and SoA instruments were not used at the end of the SAT and cannot strictly include a 'witness' effect.

It seems from the evidence of percentage agreement between markers and re-markers and their correlations, and taking the 'witness evidence' effect into account and the limited confidence it is possible to place on the data, that no major difference exists between the matrices in terms of the levels awarded. Each matrix used in this exploratory study seemed capable of producing levels more or less reliably on a remark.

Range of Levels Difference Between Marker and Re-marker

The following table shows that for the majority of the time the re-markers did not exceed one level upwards or downwards from the marker. (The instance of the re-marker awarding higher is much less than awarding lower).

	Te1	Te2	Te3	Te4	PC
Stranded (Group A)	87.5%	95.0%	65.0%	87.5%	95.0%
Stranded (Group B)	82.5%	90.0%	85.0%	97.5%	92.5%
AT Instrument	97.5%	90.0%	87.5%	97.5%	95.0%
SOA Matrix	95.0%	92.5%	77.5%	87.5%	90.0%

The percentage occurence of the re-marker disagreeing by one level (higher or lower).

This table illustrates that the re-markers rarely differed by more than one level. It is interesting to note that at PC level the SOA matrix produces the least agreement.

If the re-marker can be considered as a moderator this study seems to suggest (given the limitations of the data) that the number of appeals might run above 5% using the 50% mastery rule.

Qualitative evaluation of the different Assessment Instruments

In addition to the remark data generated by the study the markers and remarkers were asked to complete an evaluation form and record their reactions to using each matrix.

Placing the matrices in order of preferences

Each teacher was asked to place the three instruments in order of preference (first, second and third) in relation to four aspects

1. Ease of use 2. Workload 3. Layout 4. Fairness

Finally they were asked to indicate their order of preference overall. The results are shown in the table below.

Mean rank order of preference for each instrument under the 5 evaluation headings. The lower mean the higher preference.

The table shows that in most categories the AT instrument is preferred; however, the complexity of this exercise was such that only a few teachers participated (eight in four schools - each teacher spending in excess of eight hours). If only one teacher had chosen differently the rank order of the first and second places would have been different. What is conclusive is the difficulty and low level of choice placed on the Statement of Attainment device as an assessment instrument; it is the least favoured in all five categories.

This validates the averaged percentage agreement figures reported above for the matrices where the AT instrument produced the most agreement and the SoA the least across all four ATs. This is also reflected in the table showing the percentage occurence of re-marker disagreement at one level where the AT instrument (and the Stranded matrix on group A of the pupils) produce the lowest disagreement - 5% of 1 level up or down, and the SoA matrix produces 10% disagreement of more than one level.

		Mean	***Rank Order***
	Str. Matrix	1.75	2
Ease of Use	AT. Inst	1.50	1
	St. of AT	2.75	3
	Str. Matrix	2.125	2
Workload	AT. Inst	1.00	1
	St. of AT	2.875	3
	Str. Matrix	1.75	2
Layout	AT. Inst	1.50	1
	St. of AT	2.75	3
	Str. Matrix	1.625	1
Fairness	AT. Inst	2.00	2
	St. of AT	2.375	3
	Str. Matrix	1.75	2
Overall	AT. Inst	1.50	1
Preference	St. of AT	2.75	3

Mean rank order of preference for each instrument - lower figures denote higher preference

Comments

The teachers were asked to comment on their use of the different instruments.

Statement of Attainment Matrix

Only one teacher felt that the SoA matrix was preferred. The general consensus was that it was, " very difficult to work with" due to several factors including, the level a child was thought to have achieved had to be remembered whilst looking at other levels; the slightest interruption meant restarting; there was a lack of progression; indefinite distinctions between levels; and that it was time consuming. One of the teachers stated, "I was really tearing my hair out with one." All of this seems to illustrate the way in which the SoAs were not designed and cannot be used successfully as an assessment instrument.

Stranded Matrix

All teachers had used this first for the actual assessing of their pupils' SAT scores. There was support for the fact this matrix produces a number of small assessment decisions which meant that in some ways it was easier to use and interruption proof. This also provides more formative information about a pupil's performance. There was a feeling that the calculations required (taking the middle of three figures) were time consuming but it was felt that this was not a long term problem.

AT Instrument

This was liked by many. Teachers were attracted to the simplicity of making only four assessments. The use of bold type for critical indicators was found most useful and the teachers felt that in this form a logical sequence or progression was now evident. One teacher felt that although this matrix was easiest to use, "it can be unfair." This teacher felt that stranding was important, "to allow a pupil's achievement to be recorded." Another teacher commented, "definitely the scheme I would choose as I feel it to be the fairest all round for pupil and teacher."

The effectiveness of assessment the assessment devices

Assessment procedures during the 1991 Summer pilot have been undertaken within a rigorous framework. The levels secured by pupils have been determined on the basis of what the statements of attainment require. Two assessment procedures have been operated and it is essential to understand the premise underlying our reason for piloting two approaches. Trials in 1990 convinced the development team that generally design and technology teachers carry out assessments in one of two ways:

Those with an analytical/scientific background indicate a preference for itemised assessments. Examining the detailed assessment evidence, giving credit where appropriate and determining the AT scores and capability on the basis of this detailed review.

Those with an artistic/lateral perspective favour an approach which firstly establishes a general level of capability within an AT. This is followed by a closer examination which ensures the accuracy of the level judgement.

Valid assessments are more likely to be produced if teachers are able to operate in a way most sympathetic to their way of thinking. This belief that teachers do, broadly, approach assessment from these two perspectives was fully supported by the expert scrutiny of the materials. Clearly within a criterion referenced system, assessments should be congruent whether one starts from the detailed and works up or from an overview and works down. The approach taken should not affect the statements which a pupil has evidenced in their response to a task.

The two systems have been described as:

stranded assessment matrix AT assessment device

It is important that the rigour attached to each process is clearly demonstrated. If a simpler rule such as 50% (3 or more SoAs) is used the assessments produced may indeed have been higher but probably less sound. Similarly if an n - 1 rule had been administered, key or driving statements may not have been secured. Both systems though have been designed to ensure that when a level is attached to a pupil it: "represents a pupil's typical and consistent capability."

The levels which were awarded are in the context of NC design and technology which was certainly not the received curriculum for the majority of pupils taking part in the pilot. Consequently many of the teachers did not believe that the assessments fairly represented pupils' ability. Their judgements were however based on a narrower perception of the curriculum which typically these schools and teachers were delivering - eg. Te 3 focused.

The objective of both systems was to determine each pupil's operational level of capability. Consequently teachers were not asked to record achievement at the level of SoAs. This would have increased the assessment burden to an unreasonable level. Previous experience has shown that at SoA level many teachers produce rogue assessments, way beyond an individual's capability, because as yet they cannot recognise the progression which is within the statements. We appreciate this problem since the SoAs are not easy to interpret and, taken at face value, it is easy to see how erratic assessments are produced. Both devices were used by following the same procedure.

The stranded matrix

If AT profiles are examined, the comparability of this system in terms of rigour can be demonstrated.

% of statements required to secure level in all 4 ATs			
Rule Level	**n-1(3 or more SoAs)**	**Stranded Matrix**	**50% (3 or more SoAs)**
1	100%	100%	100%
2	77%	100%	78%
3	82%	75%	68%
4	78%	64%	52%
5	71%	65%	65%
6	78%	62%	62%
7	72%	90%	64%
8	72%	100%	72%
9	100%	100%	87%
10	100%	100%	87%
average	83%	85%	73%
av. KS3 (3 to 7)	76%	72%	62%

It can be seen that this matrix offers a level of rigour equal to a procedure requiring an n - 1 rule. In addition, however, it guarantees that a pupil has demonstrated ability in each of the eleven generic strands of AT - a sound guide to capability. In comparison with the 50% rule, it demands greater rigour in addition to demanding evidence in eleven generic strands.

AT assessment device

The rigour attached to this system cannot be demonstrated in the same way as with the stranded matrix due to the nature of the approach. We have, however, impressed on teachers that they must adhere to the procedure. When awarding a level using this system teachers need to be convinced that the majority of aspects within the sentence defining the level have been evidenced - the description which most closely matches the pupil's performance. This has been interpreted generally that if more than one aspect is missing the level cannot be awarded. This can be more difficult than an n - 1 rule as some statements consist of several items eg.

AT1 level 4 19 statements - 27 assessment items. Fifteen statements could secure the level across all ATs but possibly 23 assessment items would be required using AT assessment procedure.

The rigour of this procedure can also be confirmed by comparing the assessments produced with those using the stranded matrix where the rigour can be mathematically demonstrated. In our experience (all schools were visited by the central team for a moderation exercise) teachers were extremely demanding particularly in awarding the lower levels; concrete evidence was required in a project folio even for level 1 statements. There is clear evidence that many teachers do not appreciate the 5 to 16 nature of the assessment framework.

Comparison of assessment devices

	% of pupils achieving							
	AT Instrument				***Stranded***			
	Te1	**Te2**	**Te3**	**Te4**	**Te1**	**Te2**	**Te3**	**Te4**
1	16	12	11	34	12	8	8	19
2	29	30	21	29	22	30	21	28
3	30	29	32	22	35	30	30	32
4	13	17	17	10	15	15	20	12
5	7	9	12	3	10	10	11	4
6	2	2	5	1	2	4	7	1
7	1	2	2	1	2	2	3	2
8	0.3	0.2	0.3	0	1	0.4	1	0.3
9	0.2	0.1	0.1	0	0.2	0.5	0.4	0.3
10	0.1	0	0.1	0	0.2	0.1	0.2	0.2
mean	**2.8**	**2.98**	**3.23**	**2.27**	**3.1**	**3.17**	**3.44**	**2.68**

Despite the anticipation that the AT instrument, which asks teachers to make judgements against a level statement, might provide teachers with greater latitude to make higher assessments this did not prove to be the case. Mean assessments made using the stranded matrix were between 0.2 and 0.4 of a level higher in relation to the ATs. The AT instrument produced a greater number of assessments at level 1, especially in AT4.

Due to the way in which the pilot was designed, it was possible to compare clusters in which pupils have tackled the same SAT but the teachers will have used different assessment devices. In one, the teachers will have used the stranded matrix and in the other, the AT assessment device. The comparisons are shown by SAT and percentage of pupils achieving levels.

Exhibition	% of pupils achieving levels									
	Stranded matrix					*AT assessment device*				
	Te1	Te2	Te3	Te4		Te1	Te2	Te3	Te4	
1	9	5	6	18		13	6	9	35	
2	21	25	19	27		26	31	21	21	
3	36	33	29	32		31	29	34	25	
4	17	17	20	14		16	19	20	14	
5	12	12	12	5		10	11	13	2	
6	2	5	9	2		2	2	3	1	
7	2	3	3	0.1		1	2	1	1	
8	0.4	0.2	1	0.1		0.2	0.2	0.4	0	
9	0.1	0.1	0	0		0	0	0	0.2	
10	0	0	0	0		0	0	0	0	
mean	3.21	3.35	3.59	2.76	pc 3.2	2.93	3.09	3.20	2.33	pc 2.9

Measurement	% of pupils achieving levels									
	Stranded matrix					*AT assessment device*				
	Te1	Te2	Te3	Te4		Te1	Te2	Te3	Te4	
1	20	13	15	24		19	13	11	41	
2	23	35	23	33		36	36	24	31	
3	33	28	30	33		32	29	37	20	
4	15	12	15	8		8	15	13	6	
5	5	7	9	1		4	5	9	1	
6	1	2	5	0.4		1	1	5	1	
7	1	3	2	0.8		0.1	1	1	0.4	
8	0.4	0.4	2	0.1		0.2	0.1	0.4	0	
9	0	0.3	0	0		0	0.1	0.1	0	
10	0	0	0	0		0	0	0	0	
mean	2.75	2.86	3.09	2.36	pc 2.8	2.45	2.70	3.05	1.97	pc 2.5

Public Places	% of pupils achieving levels									
	Stranded matrix					*AT assessment device*				
	Te1	Te2	Te3	Te4		Te1	Te2	Te3	Te4	
1	17	10	9	28		17	14	15	35	
2	24	36	23	26		29	30	24	29	
3	28	26	29	22		28	25	30	18	
4	11	10	17	9		13	15	18	9	
5	8	6	7	4		8	10	9	6	
6	3	3	6	2		3	3	3	2	
7	5	3	2	3		0.2	3	2	1	
8	3	2	2	1		0	0	0.2	0	
9	1	2	3	3		0	0	0.2	0	
10	1	1	1	1		0	0	0	0	
mean	3.27	3.15	3.51	2.80	pc 3.2		2.82	2.98	3.02	2.29 pc 2.8

Higher assessments were achieved using the stranded matrix for each of the SATs by on average 0.3 of a level. Public Places and Exhibition produced very similar results whilst Measurement, as in previous analysis, produced results 0.3 of a level lower. All other aspects conform to the normal pattern of performance. The conclusion can be drawn that pupils who took Measurement and were assessed using the AT Instrument may have under performed by 0.7 of a level or the converse, other pupils' performance was inflated.

To check this conclusion, it is worth comparing the best achieving cluster - either Exhibition or Public Places assessed using the stranded matrix with the lowest - Measurement assessed using the AT instrument by both TA and SAT.

Public Places - SM	Te1	Te2	Te3	Te4	PC
TA mean scores	4.18	4.13	4.57	3.86	4.1
SAT mean scores	3.27	3.15	3.51	2.80	3.2
Measurement - AT I	Te1	Te2	Te3	Te4	PC
TA mean scores	2.96	2.96	3.45	2.79	3.04
SAT mean scores	2.45	2.70	3.05	1.97	2.5

In both cohorts, expectations prior to the SAT were not realised. However the discrepancy is greatest in the cluster with the highest performance. The teacher assessments were almost one level higher than the SAT performance whilst in the lowest scoring cluster the difference was about half a level.

Remark instructions

Nine schools (one from each cluster) and two teachers in each schools agreed to take part. Each transferred 10 of their own pupils' SAT scores awarded using the device originally assigned to the school and then remarked that group using the alternative device. They also used both devices on a group of 10 pupils which they did not take for the SAT.

Notation

The notation used to identify each subsection of the remark data is as follows:

Upper case 'A' and 'B' represent 2 teachers.
Lower case 'a' and 'b' represent the two different groups of 10 pupils in each school.
The 'a' and 'b' show which teacher originally assessed the group for the SAT.

Therefore:

A/a = Teacher A marking pupil group a (own pupils)
A/b = Teacher A marking pupils group b (colleague's pupils)
B/b = Teacher B marking group b (own pupils)
B/a = Teachers B marking pupils group a (colleague's pupils)

		Stranded matrix		AT instrument	
	Teacher	**mark**	**re-mark**	**mark**	**re-mark**
pupil group 'a'	**A** on pupils 'a'	A/a		A/a	
	B on pupils 'b'		B/a		B/a
pupil group 'b'	**B** on pupils 'a'	B/b		B/b	
	A on pupils 'b'		A/b		A/b

Reliability - markers

The design of the study and the differences calculated to assess the differences between markers is given over page.

Mean differences between markers using the same device					
AT instrument	**Te1**	**Te2**	**Te3**	**Te4**	**PC**
pupil group a	0.06	0.17	0.18	-0.08	0.03
pupil group b	0.12	-0.10	-0.14	-0.05	-0.01
Stranded matrix	**Te1**	**Te2**	**Te3**	**Te4**	**PC**
pupil group a	0.40	0.22	0.23	0.08	0.27
pupil group b	0.29	0.26	0.42	0.50	0.37

Mean differences over all teachers and pupils (pupils n=90)

If both markers agreed completely, the difference between them would be zero. The mean differences between markers on the same group of pupils were therefore subjected to a correlated T test analysis to see if they differed significantly from zero. The results are reported below:

Significance levels for correlated T test on mean differences between markers

	Stranded		**AT**	
	group a	**group b**	**group a**	**group b**
AT1	0.0016*	0.5880	0.6407	0.2563
AT2	0.1028	0.1086	0.2050	0.3536
AT3	0.0811*	0.0185	0.1990	0.3082
AT4	0.5692	0.0019**	0.5468	0.6572
PC50	0.0290*	0.0128*	0.8021	0.9088

* significant at the 0.05 level, (ie., only 5% difference attributable to chance)
** significant at the 0.01 level (ie., only 1% difference attributable to chance)

The stranded matrix produces significant differences between the means of the marker and remarker - ie. the differences do not occur just by chance. Generally the significances are much lower indicating that the differences between markers on the stranded matrix are more marked than in the AT instrument. This seems to indicate that the stranded matrix produces less reliable results between markers, whereas the AT instrument consistently produced high non-significant differences between markers. The conclusion from this analysis is that the AT instrument is better at producing consistent results from different teachers. All the significance levels are highly non-significant, and the PC mean-up aggregation system eradicates the small differences that there are for the AT instrument on the 'b' group of pupils to produce a significance level of 1.0000 (or no difference at all).

The AT instrument allows teachers to produce more consistent results, but at the same time, as stated earlier, it produces more conservative levels. This may be due to the requirement to make more holistic-type decisions with the AT instrument where the teacher is being required to carry much more information about the pupil performance. The stranded matrix on the other hand produces higher levels, but is less reliable. As can be seen below, it also produces a wider spread of levels disagreement between the markers.

Range of levels difference between the markers

The table below gives the percentage occurrence of markers agreeing within plus or minus one level.

- Percentage agreements to plus or minus one level

	Pupil group a					Pupil group b				
	Te1	**Te2**	**Te3**	**Te4**	**PC**	**Te1**	**Te2**	**Te3**	**Te4**	**PC**
stranded	77.7%	77.8%	73.3%	80.0%	80%	73.3%	70.0%	61.1%	71.1%	71%
(std.dev.)	1.17	1.28	1.25	1.29	1.14	1.43	1.50	1.67	1.48	1.4
AT inst.	82.3%	80.1%	76.7%	84.4%	82%	84.4%	86.7%	75.6%	92.2%	91%
(std.dev.)	1.13	1.24	1.30	1.22	1.26	1.01	1.02	1.34	0.95	0.9

In almost all cases, the standard deviation of the AT instrument is showing less spread; moreover, the majority of levels awarded for the AT instrument are within the 1 level bound. The implication is that the stranded matrix produces differences of more than 1 level more frequently and this may have implications for appeals. However, the percentage occurrence of plus or minus 2 or more levels difference is in the region of 20 - 50% for both devices.

The details of these differences are provided by school below. Notice that in some cases the agreement levels are very low between marker and re-marker, but there are also noticeable differences in agreement between, as well as within, schools.

On average, over all 9 schools each assessment device produced the following percentage levels of agreement between markers at AT level. The AT level only is considered because it is at that level the markers were making their judgements. The PC level is the result of an aggregation system and not the direct outcome of a limen referenced judgement guided by criteria.

Percentage levels of agreement between markers over all schools by AT

	AT Instrument		***Stranded matrix***	
AT	**group a**	**group b**	**group a**	**group b**
1	50.0%	48.9%	43.3%	35.6%
2	46.7%	47.8%	41.1%	37.8%
3	45.6%	38.9%	41.1%	40.0%
4	42.2%	53.3%	41.1%	34.4%
Av. over 4 ATs	46.1%	47.2%	41.6%	36.9%

Clearly the AT instrument produces better agreement between markers both at AT level and overall. However, when the patterns of agreement are inspected at school level variations are apparent under both assessment systems:

***Percentage levels of agreement (ie, difference = 0) between markers at each school**

	AT instrument					stranded matrix					AT instrument					stranded matrix				
	group a%					*group a%*					*group b%*					*group b%*				
school	**AT1**	**AT2**	**AT3**	**AT4**	**PC**	**AT1**	**AT2**	**AT3**	**AT4**	**PC**	**AT1**	**AT2**	**AT3**	**AT4**	**PC**	**AT1**	**AT2**	**AT3**	**AT4**	**PC**
1	20	10	30	50	20	20	10	40	20	20	10	40	—	40	10	10	—	—	10	—
2	40	40	—	20	10	30	50	30	—	50	40	30	20	40	20	20	10	20	10	—
3	70	50	90	60	80	90	50	30	60	10	60	60	60	100	70	60	50	50	30	50
4	70	70	90	70	60	100	100	100	90	100	80	60	70	50	70	90	90	100	100	80
5	100	100	100	100	100	30	40	50	20	20	100	90	100	100	100	40	30	40	10	40
6	30	20	—	10	10	30	10	—	30	20	50	50	10	40	60	20	30	30	30	30
7	50	50	40	20	60	60	70	60	50	60	60	50	20	20	60	30	60	40	40	70
8	40	40	10	20	40	30	30	40	20	20	30	40	40	50	50	30	40	40	40	20
9	30	40	50	20	30	—	10	20	80	10	10	10	30	40	—	20	30	40	40	20
average **%**	50	47	46	42	46	43	41	41	41	41	49	48	39	53	49	36	38	40	34	34

This presents the same findings as the remark study carried out in the Autumn which also showed that the AT instrument produced the highest average agreement over 4 ATs (48.8%). However, it should also be borne in mind that the overall level is not high - on average teachers produce the same results using either device for less than half the time. This is due to the novelty of the subject and the complexity of the assessment criteria. This also reinforces comments made earlier regarding the reliability of teacher assessment.

Rank order correlations (Spearman) between markers by device

The rank order correlations show the extent to which the markers' judgements move together in a certain direction.

AT instrument	**Te1**	**Te2**	**Te3**	**Te4**	**PC**
pupil group a	0.64	0.63	0.63	0.66	0.54
pupil group b	0.72	0.72	0.67	0.76	0.75
stranded matrix	**Te1**	**Te2**	**Te3**	**Te4**	**PC**
pupil group a	0.64	0.60	0.63	0.56	0.61
pupil group b	0.55	0.45	0.54	0.50	0.54

All correlations are significant at the 0.01 level, indicating a higher level of correlations between markers' decisions under each device. The AT instrument produces higher correlations generally, but very similar to those produced by the stranded matrix on pupil group a. Those produced by the stranded matrix on pupil group b are markedly lower.

Conclusions from the remark study

It seems clear from the remark study, given the restricted scope of the exercise , given the novelty factor, that differences do exist between the devices when used to mark the same groups of pupils, and that the AT instrument produces the most consistent results with no significant differences between markers and higher levels of agreement. The number of levels difference is also less spread out with the AT instrument.

Qualitative evaluation of the assessment devices

In addition to marking and remarking groups of pupils, the teachers involved in the remark study were also asked, having spent some time in using both devices on familiar and unfamiliar pupils, to rank the two devrices on the same five criteria as used in to the 1990 trial, see appendix 4.5.

ease of use, workload, layout, fairness, overall preference

The mean results are shown below. The lower the mean, the higher the ranking.

	stranded	AT instrument
ease of use	1.56	1.44
workload	1.56	1.44
layout	1.56	1.44
fairness	1.33	1.67
overall preference	1.44	1.56

The table demonstrates that AT instrument was preferred on most occasions in terms of ease of use, workload and layout. The stranded matrix is preferred on most occasions for fairness and overall.

It was expected that the format of the AT instrument might gain favour in terms of the first three criteria. However, teachers seem to feel that the detail which the stranded matrix contributes to the final level assessment produces a fairer result. It is probable that this outcome influences the decision about preference over all. The pattern of preferences is interesting, in that of 19 teachers, 16 of them preferred one device against the other in all 5 categories. Seven always preferred the AT instrument and 8 always preferred the stranded matrix. Only three of the teachers produced mixed rankings. This confirms the view described at the outset, that teachers preferred to operate In one of the two modes.

Conclusion on comparison of devices

There are significant differences between the two devices in terms of the results that they produce. The AT instrument provides more conservative, yet more consistent results between teachers. The stranded matrix provided a wider range of levels. In terms of the qualitative evaluation, it is clear that the AT instrument is preferred for its ease of use, layout and lesser demands on time. The stranded matrix, however was preferred overall and considered to be fairer. Teacher preferences are very equally distributed, and the differences marginal.

Assessment Criteria Te 2 - designing and communicating task 1 - construction materials	
statements of attainment	**' what to look for'**
1a	There is some **understanding that the components hold the workpiece by applying pressure**. One **possible way of designing such a clamp is sketched**. The design will not work without significant modifications .
2a	Simple sketches show an **understanding of how the clamp applies pressure to the workpiece**. A possible design is illustrated or modelled and **reasons are given for the choices made**. The design will need some modification if it is to work.
3a 3d 3e	The principle of the **clamp's geometry is understood**. The **development of the design is recorded** and shows that at least **more than one solution** has been considered **for some aspects** of the design. Reasons are given for the decisions taken about the final form of the design, which is represented either by **annotated drawings or a simple working model**. The design proposed **will be able to hold small things**.
4a 4d	Drawings show that the **principle of the clamp is clearly understood**. There is evidence that the final solution has been developed from a **range of ideas** and that these ideas have been **sequentially recorded.** Reasons are given for the decisions taken and the **final design is described in detail,** either by annotated drawings or a simple working model. The design **meets all the requirements of the task.**
5a 5b 5e	A **record of ideas** shows how the design has been **developed and refined**. Possible solutions to the various aspects of the design have been **integrated to produce a sound proposal** which is clearly presented using either detailed annotated drawings or a working model. The design proposed will operate successfully as a clamp. **Unlocking of the clamp has been tackled but the solution does not work well. A plan is provided which describes how the clamp will be made.**
6b 6d	**A design proposal**, which details all aspects of the design, shows **how decisions have been reached in arriving at a sound and feasible solution. Specialist modelling techniques** (orthographic projection or a formal 3D drawing system for example) have been used both in arriving at a solution and in describing the final proposal.The design proposed will operate successfully as a clamp which can be **unlocked easily using one hand.** A plan is provided which describes how the clamp will be made.
7a 7b	A **fully supported** design proposal, **detailing the design in orthographic projection** or equivalent shows how **ideas have been developed and combined in arriving at a coherent design.** The proposal realistically appraises the merits of the design. The design proposed will operate successfully as a clamp, can be unlocked easily using one hand and **has the possibility of being adjusted**. A **clear plan** is provided which shows how the clamp will be made.
8a 8b	A **comprehensive** design proposal, **demonstrating competence with a range of 2 and 3D communication skills** (which might include CAD), shows how ideas have been combined and refined in arriving at a coherent design. The **proposal realistically appraises the merits of the design.**The design proposed will operate successfully as a clamp, can be unlocked easily using one hand and **can be easily adjusted up to 20mm**. A **detailed** plan is provided which shows how the clamp will be made.
9a	A comprehensive design proposal, **demonstrating flair** with a range of 2 and 3D communication skills (which might include CAD), shows how ideas have been combined and refined in arriving at an **optimum practicable design.** The proposal realistically appraises the merits of the design whilst **recognising its constraints.**The design proposed will operate successfully as a clamp, can be unlocked easily using one hand, can be easily adjusted up to 20mm and **most parts are fastened or pivoted together**. A detailed plan is provided which shows how the clamp will be made.
10a	A comprehensive and **coherent** design proposal, demonstrating flair with a range of 2 and 3D communication skills (which might include CAD), shows how ideas have been combined and refined in arriving at an optimum practicable design. The proposal realistically predicts the performance of the design and details its constraints.The design proposed will operate successfully as a clamp, can be unlocked easily using one hand, can be easily adjusted up to 20mm and **all parts are fastened or pivoted together.** A detailed plan is provided which shows how the clamp will be made.

Assessment Criteria Te 3 - Planning and making task 1 - construction materials	
statements of attainment	**' what to look for'**
1a	An **attempt has been made to make a clamp** but due either to defects in the design, lack of precision or failure to complete the task, the clamp does not work.
2b	A clamp has been made which will **hold small objects** and demonstrates that **simple hand tools have been used with some success.** The materials chosen to make the clamp **demonstrate a knowledge of their properties.**
3b 3c	A clamp has been made which **works satisfactorily**. **Materials and equipment have been chosen** such that the clamp has been made **to a satisfactory level of accuracy and quality.**
4a 4c 4e	A clamp has been made which **meets all the requirements** of the task. The **making** of the clamp has been **achieved efficiently** (appropriate choice of tools, equipment and processes) with **minimal waste** resulting from inaccurate work.
5a 5c	A clamp has been made which **operates efficiently** with the exception of the **unlocking device which is difficult to operate**. The making was **well organised,** evidence of a plan is provided, to ensure the **efficient use of time, materials, equipment and labour.**
6a 6b	A clamp has been made which operates efficiently and **can be unlocked easily using just one hand.** There is clear evidence that: the materials from which the clamp is made have been **chosen because of their properties;** the making process was **planned (with assistance) in detail.**
7a 7b	A clamp has been **made precisely** (as detailed in the design proposal) so that it operates efficiently, can be unlocked easily using just one hand and has **some facility for adjustment,** however, this aspect does not operate particularly well. With some advice, the making has been planned efficiently, but with guidance, so that the materials, tools and equipment have **matched the scale of production and the time available.**
8a 8b	A clamp has been made precisely (as detailed in the design proposal) so that it operates efficiently, can be unlocked easily using just one hand and can be **easily adjusted to accommodate items up to 20mm.** The making has been planned, with advice, to make the **best use** of the materials, procedures, tools and equipment available. There is evidence that the **quality of production has been monitored at every stage of manufacture.**
9b 9c	A **quality outcome has been achieved** (as detailed in the design proposal). The clamp operates efficiently, can be unlocked easily using just one hand, can be easily adjusted to accommodate items up to 20mm and **attempt has been made to ensure that no parts can be detached.** Manufacture has been **planned independently** to make sensible use of the materials, procedures, tools and equipment available. There is evidence that the quality of production has been monitored at every stage of manufacture.
10a	A **high quality** outcome has been achieved (as detailed in the design proposal). The clamp operates efficiently, can be unlocked easily using just one hand, can be easily adjusted to accommodate items up to 20mm and **no parts can be detached.** Manufacture has been planned independently to ensure the best use of the materials, procedures, tools and equipment available. There is evidence that the quality of production has been monitored at every stage of manufacture.

Assessment Criteria Te 2 - designing and communicating task 2 - control materials	
statements of attainment	**' what to look for'**
1a	Simple sketches indicate an understanding that the LED has to be energised by the two cells. A simple means of securing the cells has been considered, but the design is unlikely to work without significant modifications.
2a	Sketches show an understanding that the two cells must be **connected in series** and **connected as part of a circuit to the LED**. A simple means of securing the cells to make **electrical contact** has been considered but may need modification to be made to work.
3a 3d 3e	Understanding of the basic LED circuit is shown by at least one clear diagrammatic sketch illustrating a **workable circuit** including cells connected in series, an LED and a series resistor. There is at least one **realistic and workable idea for securing the cells** to make electrical contact, and **some reasons** for design decisions are evident.
4a 4d	Understanding of the basic LED circuit is shown by a clear circuit diagram illustrating a workable circuit including cells connected in series, an LED and a series resistor. This is accompanied by **sketches or plans for an actual circuit layout**. There is more than one realistic and workable idea for securing the cells to make electrical contact, and reasons for design decisions are evident.
5a 5b 5e	A **record of ideas** shows how the design work for the circuit, cell holder and switch has been developed and refined. Elements of design work **have been integrated** to produce a sound proposal which is clearly presented as annotated drawings and an **intelligible circuit diagram**. The switch design has been attempted but the favoured solution requires refining. **A plan is provided** showing how the beacon will be completed.
6b 6d	A design proposal, which details all aspects of the design, shows how decisions have been reached in arriving at **a sound and feasible solution**. **Specialist modelling techniques** (eg., formal orthographic drawing and circuit diagram) have been used both in arriving at a solution and in describing the final proposal. A realistic switch design is proposed and a plan is provided showing how the beacon will be completed.
7a 7b	A **fully supported design proposal** - detailing the design elements through orthographic projection (or equivalent) and a circuit diagram shows how ideas have been combined in arriving at a **coherent design**. The proposal **realistically appraises** the merits of the design elements. A clear plan is provided which shows how the beacon will be completed. The proposed design will operate successfully as a beacon and will be capable of switching using a screwdriver.
8a 8b	A comprehensive design proposal, demonstrating competence with a range of 2D and 3D communication skills (which might include CAD) shows how ideas have been combined and refined in arriving at a coherent design. The proposal realistically appraises the merits of the design. The proposed design will operate successfully as a switched beacon whose **flash rate can be adjusted within the limits stated**. A detailed plan is provided showing how the beacon board will be made.
9a	A comprehensive design proposal, demonstrating flair with a range of 2D and 3D communication skills (which might include CAD) shows how ideas have been combined and refined in arriving at an optimum practicable design. The proposal realistically appraises the merits of the design whilst **recognising its constraints.** The proposed design will operate successfully as a **manually or automatically switched** beacon whose flash rate can be adjusted within the limits stated. A detailed plan is provided showing how the beacon board will be made.
10a	A comprehensive and coherent design proposal, demonstrating flair with a range of 2D and 3D communication skills (which might include CAD), shows how ideas have been combined and refined in arriving at an **optimum practicable design**. The proposal **realistically predicts** the **performance of the design** and details its constraints. The proposed design will operate successfully as a manually or automatically switched beacon whose flash rate can be adjusted within the limits stated. A detailed plan is provided showing how the beacon board will be made.

Assessment Criteria Te 3 - Planning and making task 2 - control materials	
statements of attainment	**' what to look for'**
1a	An attempt has been made to make the beacon board (with cell accommodation) but due to either defects in the design, lack of precision or failure to complete the task, the LED is not energised.
2b	A beacon board has been made which is **capable of energising** the LED. It demonstrates that **simple hand tools have been used with some success.** The materials chosen to make the board and cell attachment **demonstrate a knowledge of their properties.**
3b 3c	A beacon board has been made which **works reliably**. Materials and equipment have been chosen and used, such that the board has been made **to a satisfactory level of accuracy and quality.**
4a 4c 4e	A beacon board has been made which **meets all the requirements** of the task (eg., including ease of replacing cells). The **making** of the board has been **achieved efficiently** (appropriate choice of tools, equipment and processes) with **minimal waste** resulting from inacurrate work.
5a 5c	A beacon board has been made which **operates effectively** with the **exception of the switch which is difficult to operate**. The making was **well organised** and a plan in evidence - ensuring the **efficient use of time, materials, equipment and labour.**
6a 6b	A beacon board has been made which operates effectively and can be **switched on easily** using a screwdriver. There is clear evidence that: the materials and components from which the board is made have been **chosen because of their properties;** the making process **was planned (with assistance) in detail**.
7a 7b	A beacon board has been made **precisely** (as detailed in the design proposal) so that it operates effectively, can be switched on and off and **flashes continuously**. However, the flash rate is not easily adjusted. With some advice, the making has been planned efficiently, but with guidance, so that the materials, components, tools and equipment have **matched the scale of production and the time available.**
8a 8b	A beacon board has been made precisely (as detailed in the design proposal) so that it operates effectively, can be switched on and off, and flashes continuously. **The flash rate is easily adjusted.** The making has been planned, with advice, to make the best use of the materials, components, procedures, tools and equipment available. There is evidence that the **quality of production has been monitored at every stage.**
9b 9c	A **quality outcome has been achieved** (as detailed in the design proposal). The beacon board operates effectively, can be switched on and off and has an easily adjustable flash rate. **A reasonable attempt has been made to provide for automatic switching** in reduced light. Manufacture has been **planned independently** to make sensible use of the materials, components, procedures, tools and equipment available. There is evidence that the quality of production has been monitored at every stage of manufacture.
10a	A high quality outcome has been achieved (as detailed in the design proposal). The beacon operates effectively, can be switched on and off and has an easily adjustable flash rate. It also switches on automatically in reduced light conditions. Manufacture has been planned independently to ensure the best use of materials, components, procedures, tools and equipment available. There is evidence that the quality of production has been monitored at every stage of manufacture.

Assessment Criteria Te 2 - designing and communicating task 3 - food materials	
statement of attainment	**' what to look for'**
1a	Pictures and/or writing show(s) that **the task has been understood** and **some action is being attempted in response to it**.
2a	Pictures and writing **describe what is to be made** in response to the task and **some basic understanding is revealed in the choice of ingredients**.
3a 3d 3e	Pictures and notes **describe how energy giving ingredients will be combined** to create a **compact bar** and **why they have been chosen.**
4a 4d	A record, using drawings and notes, shows **how various combinations of ingredients have been considered and reconsidered** for suitability.
5a 5b 5e	**Information on** sources of **carbohydrate, protein, vitamins, minerals, etc. has been used** and a record **shows clearly how and why the solution developec as it did**, including a **list of ingredients** as they would appear on the packaging.
6b 6d	**Information** on various sources of carbohydrates, protein, vitamins, minerals, etc. **has been sought and used critically** to provide a **balanced** product. A record shows clearly how and why the solution developed as it did, and lists the ingredients as they would appear on the packaging.
7a 7b	**A range of sketches, notes and diagrams** records how **relevant information and different combinations of ingredients have been explored and considered** to develop a solution providing the **required kilojoules and content**. A **scale drawing** of the finished product shows how **layering and/or coating** will enhance the product, and how it will be **divided into smaller pieces.**
8a 8b	A range of sketches, notes and diagrams records how **a variety of combinations** of ingredients, **perhaps targeting different tastes**, have been explored and considered. **Different possible coatings and methods of layering have been detailed**, taking into account the need to divide the bar into smaller pieces. A scale drawing shows the final design and the **whole record reveals preparedness to experiment.**
9a	A **comprehensive** record, including CAT/DTP/use of software where appropriate, demonstrates how **critical knowledge of commercial products** and a **wide range of experiments** have attempted to identify ingredients which could form the basis of a **range of similar energy bars**, each containing approximately the correct number of kilojoules and the same ratio of carbohydrate, protein, vitamins, etc. Ways of **overcoming conflicts** have been sought. A scale drawing shows the final design which is to be made, fulfilling all requirements of the task.
10a	A **substantial** and **fully justified** account, **including supporting data** and using CAD/DTP/software where appropriate, presents a full range of **existing** and **potential** ideas and **feasible design solutions to a prospective manufacturer.** Possible combinations of ingredients and attempts at resolving conflicting demands have been **continually refined** as a result of **testing**. The final design, which fully achieves the requirements for a range of energy bars, provides **an original and elegant solution.**

Assessment Criteria Te 3 - Planning and making task 3 - food materials	
statement of attainment	**' what to look for'**
1a	Food ingredients have been combined to produce **an attempt at a simple 'bar' which someone could carry with them** (though the ingredients may not be appropriate and the resulting shape may not resemble a bar).
2b	A simple and reasonably compact energy bar has been produced, **showing ability to use basic tools and equipment**, though with little consideration for suitability of content.
3b 3c	A compact bar has been made which shows **some attempt to achieve a quality finish**, though the content may be high in sugar and fat.
4a 4c 4e	A compact bar has been made which **achieves a quality finish** and shows that **thought** has gone into **providing energy in a form that would be palatable and portable**, though not necessarily healthy, for the user.
5a 5c	A **plan for making** has been provided and the finished high energy bar, **containing some protein, vitamins and minerals, as well as carbohydrates**, provides a compact solution. There is an **attempt** to enable the user **to divide the bar into smaller pieces.**
6a 6b	A **step-by-step plan** for making has been provided and the resulting bar is compact, **tasty** and offers a **healthy combination of ingredients**. It can be **easily divide into smaller pieces.**
7a 7b	Co-ordinating time and resources competently and **incorporating necessary modifications**, an attractive and compact high energy bar has been produced with the **correct kilojoule value and ratio of ingredients**. The bar can be easily divided into smaller pieces and **includes layering and/or coating.**
8a 8b	Improvements and modifications have been **identified** and incorporated during making and **tests for quality** have been carried out on the compact, appealing and highly nutritious energy bar which has been produced. The bar can be easily divided into smaller pieces and **effectively** utilises layering **and** coating.
9b 9c	Knowledge of **methods of manufacture** and of **existing commercial products** have been included in producing a compact bar which uses layering and coating, divides easily into smaller pieces, and is **presented as part of a coherent range of possible bars** which all fulfil the stated criteria. Quality tests have been carried out, focusing particularly on the **physical** and **nutritional** characteristics.
10a	**Creativity** and **confidence** have been evident during making, where **a range of techniques** and **continual testing and modifications** have resulted in a high quality energy bar which fulfils all the stated criteria, together with practical experiments to arrive at a **possible series of bars** which all fulfil the physical and nutritional criteria and could be **effectively marketed commercially.**

ASSESSMENT SCHEME - national pilot 1992

Instructions to the marker

This test is to determine a pupil's level in two of the four attainment targets.
The test differentiates by task in making assessments of Te 1 and Te 4. Two statements of attainment are tested at each level for each attainment target except where there is only one statement.

The first set of questions, which relate to part of Te 4, asks pupils to review their project relating to the charity fair.
The remaining questions, which relate to Te 1 and part of Te 4, ask pupils to use the knowledge and experience of the long task in a fresh context.
The mark scheme is in the same order as the tests.

Using the mark scheme

The mark scheme states:

- the statements of attainment being tested;
- the questions which will provide the evidence;
- what to look for when carrying out an assessment.

The "what to look for" column indicates the evidence required to satisfy the SoAs in the context of this test. The requirements indicate what evidence it is reasonable to expect in the test but this is obviously not as comprehensive as would be expected in teacher assessment.

Assessment procedure

- Each question is focused on a specific SoA at a particular level. In each test, every pupil starts at the lowest level and works towards the highest level in each section.
- Mark all the questions which the candidate has attempted.
- Tick the coloured statement box on the test paper if the answer is satisfactory; cross the box if the answer is unsatisfactory.
- Record your assessment on each pupil's OMR form.
- The column 'Que.' on the mark scheme shows the question which tests the statement.
  The test is also identified beside the question number.
  Both the question number and the test are colour coded in line with the test papers.

Determining a level

The rules for deciding the level achieved in each attainment target and the method for aggregating these with the scores from the long task have yet to be determined.

An indication of the level at which children performed in the test can be obtained in the following way:
Te 1 Section B - the average, rounded down, of the highest two questions answered correctly. For example, if a pupil's two highest correct answers are at levels 4 and 6, level 5 has been achieved.

Te 4 - Section A and C - the average, rounded down, of the highest correct answer in each section. For example, if a pupil's highest correct answer in section A is 6 and section C is 3, level 4 has been achieved.

Attainment target Te 4		
Section C	**Que.**	**What to look for**
1b describe to others what they like and dislike about familiar artefacts, systems or environments.	test 1 **13**	One sensible like & one dislike are given, showing that they have looked at the drawing & understood what it represents *e.g.*, - I like the picture of the donkey. - I dislike pins - they are dangerous
2b make simple judgements about familiar artefacts, systems or environments, including those from other times and cultures.	test 1 **14**	a) One possible reason is given, *e.g.*, it is the most attractive. b) One possible reason is given, e.g., the pin would secure it safely or most people would have a button hole they could fix it to. N.B. An answer in the personal ("I like donkeys") is not acceptable at this level.
3b comment on the materials and processes used and how the task was tackled.	test 1 **15** test 2 **16**	One reason is given for each, *e.g.*, - a) speed of production / donkey didn't need to be drawn by hand each time. - b) speed of production/ produces identical images.
4b review the decision-making process they used in producing their final artefact, system or environment.	test 1 **16** test 2 **17**	Three considerations are given, *e.g.*, - cost of materials; - what people are likely to pay; - promotion or display costs.
5b justify the ideas, materials, components, procedures, techniques and processes used, and indicate possible improvements.	test 2 **18**	a) Two practical reasons are given, eg.; - relatively cheap; easily cut/worked; can be vacuum formed; b) an appropriate material is named, eg.; - thin plywood, metal, plastic, thick card; c) one suitable method is given, eg.; - lino, silk-screen, block, potato printing d) two reasons are given, eg.; - bold clear image, - consistent reproduction of image.
6b devise and carry out ways of testing the extent to which the product satisfies their design specification.	test 2 **19** test 3 **16**	a) Two tests are given, *e.g.*, - a practical test to check how it is attached; - questionnaire / interview to test people's reactions; b) two reasons are given, *e.g.*, - in identifying manufacturing problems; - in identifying stages/sequence of manufacture.
7a present an evaluation of their activities against the original need, drawing on information gathered about the product and the reactions of users. Evaluation should include suggestions for improvements.	test 2 **20** test 3 **17**	a) Three aspects are identified, e.g., - visibility - badges are not very visible; - preventing damage - badges have not been packed to protect them; - creating an image - name of the charity is not on the outside of box; b) Two feasible improvements are drawn and described. The answer must include information on the materials which would be used and the manufacturing procedures and techniques which would be used.
8b understand that artefacts, systems or environments reflect the circumstances and values of particular cultures and communities.	test 3 **18**	A practical and sensible approach is discussed under each heading. The suggestions must be in keeping with the local scale of the event, eg., newspapers - ask the paper to carry an article about the work of the charity; place an advert in the, "coming events", column. advertising posters - ask shops/house owners to put posters in their windows; talk to the local authority about putting up banners. exhibitions - ask the local library/supermarket/hotel/tourist information office if a small exhibition explaining the work of the charity can be mounted in the foyer/ entrance hall.
9a demonstrate that they have applied knowledge and understanding derived from evaluations of their own and others' design and technological activities	test 3 **19**	The implications of mass production rather than batch production should be clearly illustrated in the answer. Reference should be made to aspects such as: quantity purchase of materials, use of alternative materials which will reduce costs, faster production methods such as injection moulding for donkey badge and silk screen printing for cat badge. The modified design should take all such aspects into account. The answer should detail the implications of the changes to the design on the manufacturing processes to be used.
10b evaluate artefacts, systems or environments to show the interaction of influences on their development and use this knowledge in their own work	test 3 **20**	A detailed description of how these three key aspects could be evaluated by the charity, eg., organisation - an internal review - probably qualitative - which examines all aspects of the event from the promotional campaign to the collection of takings. Evaluate the systems used in relation to both efficiency and appropriateness. Discuss the organisation of the event with those taking part - badge makers, sellers etc. cost effectiveness - a critical review of the return on investment. All elements of expenditure should be examined independently looking for possible savings and better value for money. public awareness - use of questionnaire and interviews to discover how people responded to the work of the charity and if this affected their willingness to contribute. How the public responded to raising money via a badge day and if they would have responded more generously to other kinds of events. Need for ongoing evaluation - before, during and after the event - should be evidenced in the answer.

Attainment target Te 4		
Section A	**Que.**	**What to look for**
1a describe to others what they have done and how well they have done it.	test 1 **1**	A sensible response to both a) and b) which refers to the pupil's work for the charity fair project.
2a discuss with teachers and others how satisfied they are with their design and technological activities, taking into account their original intention and how they went about their task.	test 1 **2**	Both sentences are completed with an intelligent response which refers to their project. The second may refer to either skills, knowledge or operational aspects, such as time management.
3a discuss their design and technological activities and their outcomes with teachers and others, taking into account how well they have met the needs of others.	test 1 **3** test 2 **1**	Both sentences are completed with plausible responses. For example the first may refer to meeting a need or the quality of their outcome. The second response might refer to making money or helping the charity with its work.
4a review the ways in which their design has developed during the activity, justifying decisions and appraising results in relation to intentions.	test 1 **4** test 2 **2**	Two reasons are given which refer to their outcome and batch production. For example a template was made so that shapes could be repeated accurately.
5a evaluate their product in relation to the design intentions and to the original needs or opportunities, taking into account users' views, cost-effectiveness and scale of production.	test 2 **3**	The relationship between expenditure and income. Answer must show that both sides of cost effectiveness have been taken into account, eg., in the context of the charity fair, production costs should be outweighed by direct profit from a sold product or indirect profit from a promotional product.
6c evaluate the ways in which materials have been used.	test 2 **4** test 3 **1**	a) Two valid reasons are given for the choice of the materials. Answers may refer to physical properties, availability, or cost, for example. b) Two valid reasons are given for the choice of materials. Answers may refer to visual properties such as colour or pattern, or the effect of combining materials. A personal statement such as, "because I like them," is not adequate at this level.
7a present an evaluation of their activities against the original need, drawing on information gathered about the product and the reactions of users. Evaluation should include suggestions for improvements.	test 2 **5** test 3 **2**	a) Three relevant comments are listed. For assessment purposes, however, the second part is of more importance. b) One improvement is clearly and accurately described and illustrated with notes detailing the effect of the improvement on the key issues stated in the question.
8a present an evaluation of their activities, including suggestions for improvements, and a discussion of: (i) the relationship between the materials chosen and the procedures, techniques and processes used (ii) justification of possible improvements (iii) the suitability of the product for manufacture (iv) an estimate of the effects and consequences, including environmental and economic ones.	test 3 **3**	Answers to each of the three parts must relate clearly to batch production. There may be a degree of repetition in answering the three parts. For example a material may have been selected because it was suitable for the production process. If this is the case look for an additional comment in the evaluation of each aspect.
9a demonstrate that they have applied knowledge and understanding derived from evaluations of their own and others' design and technological activities.	test 3 **4**	Three experts are listed together with appropriate and detailed questions. Experts must be relevant to the task a pupil worked on. Questions must recognise why they are consulting that expert. Examples, expert offering design input, manufacturing input, financial knowledge, patent information, legal advice, trading standards information, the charity concerned.
10a demonstrate, through their choice of working methods and discernment and flair in decision taking, the quality of their design and technology capability.	test 3 **5**	Any answer to be credited must evidence flair in designing, and a quality of communication skill which is commensurate with this level. The answer must refer to the project and the assessor must be confident that the claims in the answer are supported by the evidence from the long task.

TASK 1 **CONSTRUCTION MATERIALS**

Making things out of sheet material or wire (up to 1mm thick) can be difficult without a clamp to hold them.

A quick lock/release clamp would be especially useful to someone making jewellery.

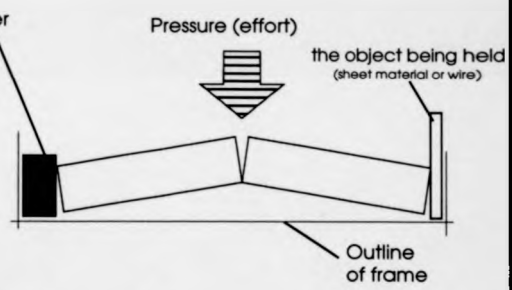
The picture shows a clamp which can be locked and released quickly.

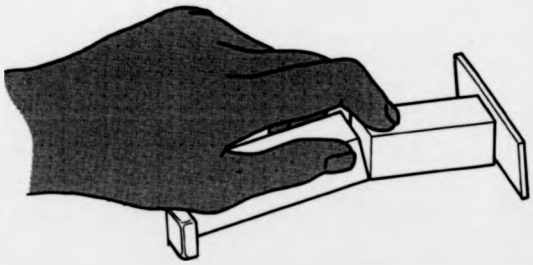
It uses two moving parts inside a frame.

When they are pushed down in the centre, they will go flat and exert pressure at both ends.

Any thin object placed as shown will be trapped and held. A small rubber block (e.g. pencil eraser) at the opposite end helps the system to work as a clamp.

How the clamp works.





The pictures show the basic idea which you need to develop.

- The locking pieces may be shapes other than rectangular blocks.
- The locking pieces may be made from wood, metal or plastics

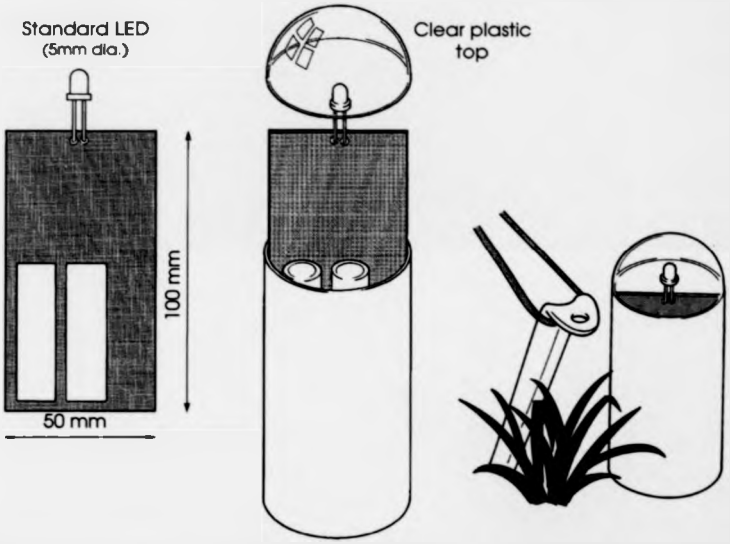
level	
1-4	**Design and make** a clamp to hold thin objects, which uses two moving parts. The clamp should be no longer than 180mm in length. The clamp must be easily moved around on a surface. When pressure is applied with a hand, the clamp should lock and stay locked when the hand is taken away.
3-6	It should be possible to unlock the clamp easily, using just one hand.
5-8	The clamp should be adjustable so that components up to 20mm in diameter or thickness can also be held - e.g. small jewellery parts for soldering or sheet acrylic for laminating.
7-10	All parts of the clamp should remain fastened or pivoted together so that nothing can ever be separated and lost.

TASK 2 **CONTROL MATERIALS**

The picture shows the outline of a small warning beacon used when camping.

The beacon might be placed next to tent pegs which people might trip over at night.

The warning beacon consists of a moulded plastic case containing an LED circuit and at least two 1.5 volt cells (HP7 size).



The size of the board (1.5 to 3mm thick) *The board being inserted in the moulded plastic case* *the warning beacon in use*

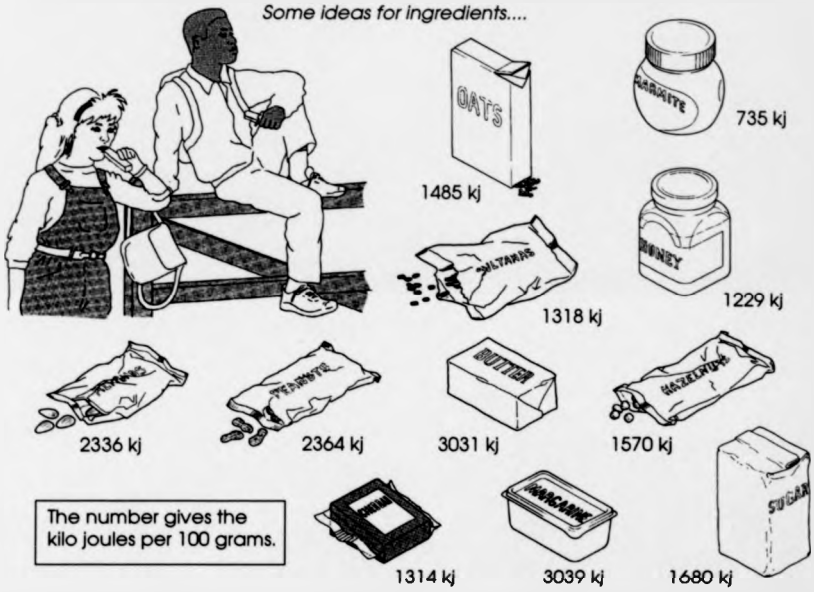
level	
1-4	**Design and make** an LED circuit. The circuit must be fitted onto, or be a part of, a board which would slide into the case as shown. The LED should be mounted in the position on the board, as shown in the picture. You should also design and make a way of holding the cells onto the board and connecting them to the circuit. It must be easy to replace the cells when necessary.
3-6	The beacon should be operated by a switch fitted to, or as part of, the board, for turning the LED on or off. It must only be possible to operate this switch through a 6mm diameter hole in the side of the lamp case.
5-8	In operation, the LED should flash on and off continuously at a frequency of between 1 and 5 Hz. This frequency should be adjustable through a second 6mm hole in the case.
7-10	The warning beacon should light up automatically when darkness falls.

You MUST NOT make the case.
You MUST NOT solder onto the cells.
You MUST NOT use a ready made cell holder or the parts taken from one.
You MUST NOT use a ready made switch or the parts taken from one.

Task 3 **Food materials**

The energy and other nutrients which our body needs come from the food we eat during the day, both from meals and from snacks. Snacks are often used to satisfy hunger between meals but they may also be eaten instead of a meal.

Some ideas for ingredients....



The number gives the kilo joules per 100 grams.

level	
1-4	**Design and make** a snack bar to be eaten in place of a meal during a day's hiking or a sponsored walk. The bar must be convenient for someone to carry with them, so it must be as compact as possible.
3-6	The bar should contain not only carbohydrates, but also protein, vitamins and minerals. It should be designed so that it can easily be divided into smaller pieces. Your design notes must identify and justify your decision to include each ingredient. State the ingredients as they would appear on the packaging and give the total kilojoules/kilocalories in the bar.
5-8	The bar should provide a minimum of 1470 kilojoules/350 kilocalories and it should have a maximum volume of 250 cubic centimetres. Layering and coating are used to make products look appealing. Include at least one of these methods in your product.
7-10	In addition to designing and making one version of the bar, you should produce a concise written report to a prospective manufacturer. The report should demonstrate how this bar could be one in a range of bars appealing to different tastes but having similar physical and nutritional characteristics. You should make use of desktop publishing/ word processing in your report.

Section C

The charity fair in your school was a success.
The charity still needs more money.
You and your friends have decided to raise more money by holding a badge day.
The badges you make will be sold in your local high street.

Level 3

6 When deciding what kind of badge to make, you will need to talk both to the people who run the charity and to people who might buy the badges.
Write a question you would ask:

a) people who run the charity _____

_____

b) people who might buy the badge _____

_____

te1.3b

Level 3

7 Suggest **THREE** other ways in which a charity might raise money apart from a fair or a badge day.

First way _____

Second way _____

Third way _____

te1.3a

Level 4

8 When designing the badge, which of the following would you choose?
Tick your choice.

a) badges which are easy to make from expensive materials. ☐
b) badges which are difficult to make from cheap materials. ☐

Give **TWO** reasons for your choice.

Reason one _____

_____

Reason two _____

_____

te1.4d

Attainment Target Te1 - Identifying needs and opportunities

Level 4

9 You need to find out about possible materials from which the badge could be made. List **THREE** things you might do and describe what information you might find. Answer by completing the following sentences.

First thing

I would _____

_____

I might find out _____

_____

Second thing

I would _____

_____

I might find out _____

_____

Third thing

I would _____

_____

I might find out _____

_____

te1.4b

Level 5

10 How would you take the following into account in developing a design for a badge? Give **ONE** example under each heading.

A functional aspect _____

_____

_____

A visual aspect _____

_____

_____

An environmental factor _____

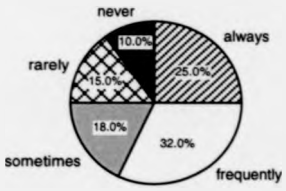
_____

_____

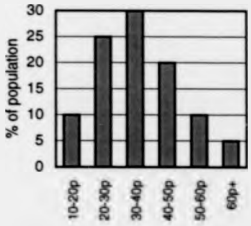
te1.5b

Attainment Target Te1 - Identifying needs and opportunities

11 From a questionnaire you discover the following:



People's attitude to wearing badges



"How much would you be prepared to pay for a badge being sold to raise money for a charity?"

How much are they prepared to pay?

Level 5

a) Which information would be the most useful in deciding on the selling price of your badge?

Information I would use _____

_____

_____

b) Give a reason for your choice.

Reason for choice _____

_____

_____

te1.5a

Level 6

12 Using the charts in the previous question, what would be your target selling price for the badges?

My target selling price would be _____ pence.

Give **TWO** reasons for selecting this selling price.

Reason one _____

_____

Reason two _____

_____

te1.6a

Attainment Target Te1 - Identifying needs and opportunities

Part 2

1. **Look at picture A**. These comics, magazines and toys get very untidy, so that people using the waiting area cannot find what they want. If you had to design a better way of keeping these things tidy, you would need to do some research.

   a) Describe **two** things you would need to find out from people using this waiting area.

   i) _____

   _____

   ii) _____

   _____

   b) Describe **two** things you would need to find out from the cleaner of this waiting area.

   i) _____

   _____

   ii) _____

   _____

4 marks

1/5a
1/5b

2. a) As part of your research into designing a better way of keeping the magazines and toys tidy, you decide to talk to the person in charge of the waiting area. The person tells you the following:

- people complain that the waiting area is always too cold; ☐
- people never put magazines or toys back because they don't know where to put them; ☐
- some people sit listening to personal stereos;
- there is very little money available to improve the waiting area. ☐

Tick the **two** comments which would be most useful in designing your solution.

b) How would each comment you have chosen help you to design a better way of keeping the magazines and toys tidy?

Comment one _____

_____

_____

Comment two _____

_____

_____

4 marks

☐

1/5a
1/5b

3. **Look at picture B**. You have been asked to design a better chair. Describe **two** important things about chair design for each of the following.

a) The comfort of the people sitting on the chairs.

i) _____

_____

ii) _____

_____

b) The convenience of the cleaner of the waiting area.

i) _____

_____

ii) _____

_____

c) Describe **two** important design features for a chair which would meet the needs of both the people sitting on the chairs and the cleaner of the waiting area.

i) _____

_____

_____

ii) _____

_____

_____

6 marks

1/6a

4. The fabric used on chairs in waiting areas must be chosen carefully. What **two** physical properties of fabrics would be important for new chairs in a waiting area? Explain why you think each is important.

First physical property _____

Why is it important? _____

_____

_____

Second physical property _____

Why is it important? _____

_____

_____

2 marks

1/6c

5. **Look at picture C**. Many of these people would like something to eat or drink, while they wait. Imagine that the person responsible for the waiting area has asked you to look into the possibility of starting up a small refreshment area, but is concerned about the following:

a) the cost of starting up
b) litter
c) health and safety
d) what food to provide

Describe **two** things you would need to investigate about each of these to provide enough information for the owner to decide whether to have a refreshment area.

a) The cost of starting up:

i) _____

_____

ii) _____

_____

b) Litter:

i) _____

_____

ii) _____

_____

c) Health and safety:

i) _____

_____

ii) _____

_____

d) What food to provide:

i) _____

_____

ii) _____

_____

8 marks

1/7a
1/7c

6. a) **Look at picture D**. The man making the telephone call is having a problem because he is finding it difficult to hear. You have been asked to evaluate a number of different solutions to this problem. Explain why you would need to consider the following:

• access • cost • privacy • safety

Access _____

_____

_____

_____

_____

Cost _____

_____

_____

_____

_____

Privacy _____

_____

_____

_____

_____

Safety _____

_____

_____

_____

_____

b) In deciding if a possible design solution was suitable, you would need to investigate, in detail, access, cost, privacy and safety. Use a diagram or chart to show how you would investigate each of these.

Access

Cost

Privacy

Safety

8 marks

1/8a
1/8b

INSET - 1991 Pilot

The introduction of the National Curriculum in design and technology posed the majority of schools with a new curriculum challenge. Consequently INSET provision assumed a much wider significance than preparation just for the pilot of an assessment procedure, involving delivery, assessment and standardisation. The piloting proved to be a potent means of exemplifying the subject to the majority of schools taking part. Especially as the materials which had been developed fully embraced the philosophy which underpins the statutory Order.

Several models of INSET were employed for this pilot to determine the efficiency of possible approaches. In addition, every school taking part was provided with an INSET kit see (a) below. The six INSET models used for design and technology are as follows:

a. Standard
A one day INSET session provided by the central team for two teachers from each participating school. These teachers were provided with a distance learning pack - 'School Kit' - which provided them with the materials to organise and run similar sessions for all teachers involved in delivering and assessing the SAT. The kit was designed in a flexible fashion to allow teachers and schools to modify it to suit their specific needs.

b. Enriched
One day INSET session provided by the central team for two teachers from each participating school. This was followed by a session in each school conducted by the central team. In this model each teacher taking part had direct contact with a member of the central team.

c. Distant
A representative from the selected LEAs received training from the central team so that they could assume the role of INSET Deliverer and Moderator. This person was then responsible for all aspects of the SAT within that LEA.

d. Exam Board
An exam board (WJEC) was responsible for all aspects of SAT production, INSET, and moderation for Welsh medium schools.

e. INSET relating to choice of SAT and assessment device
The pattern of INSET was different for the schools offered a choice of both the SAT they trialled and the assessment device they used. These schools received an additional half day INSET to provide them with information to enable them to make the choice.

f. Standard with pre-SAT information
Some schools were be sent copies of the SAT stages leaflet prior to the INSET day.

All teachers attending INSET sessions, regardless of model, completed an evaluation form covering all aspects of the INSET. The analysis of this evaluation is shown below. This evaluation is based on the views of 128 teachers, 15 LEA representatives plus HMI and SEAC representatives who attended the INSET days. The general level of satisfaction was extremely high and consistent across all the clusters.

Analysis of evaluation for the D&T SATs INSET courses

Tick boxes as appropriate

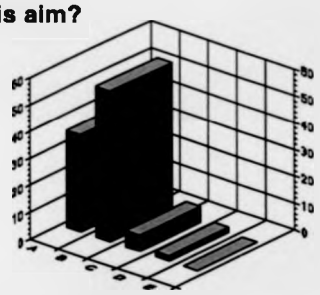
A	B	C	D	E

Excellent Satisfactory Poor

1. The aim of this course was:

To provide appropriate training for the teachers responsible for in-school INSET and SAT organisation and delivery in schools taking part in the pilot.

• How well do you feel that the day went meeting this aim?



Comments:

Clear aim which was successfully delivered in a calm atmosphere. More information before the day might have helped but most felt reassured about how to approach the SAT.

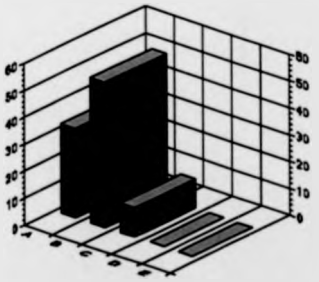
2. The specific objectives of this course were:

1. *To appreciate the objectives of the pilot.*
2. *To familiarise a member of staff in each school with the SAT materials.*
3. *To give colleagues an opportunity to explore the kinds of outcomes the pupils could be expected to produce in the activity.*
4. *To discuss issues relating to the management of the materials in school.*
5. *To consider the mechanism for assessment and the issues this raises.*
6. *To promote the standardisation of assessment.*

• How well do you feel that these objectives were met during the course?

Comments:

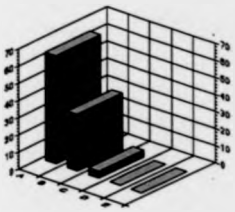
All objectives were addressed but perhaps more time for standardisation of assessment. Materials for promoting discussion were excellent and questions on pupil's work were very appropriate.



N.b Vertical percentage axis limit varies from graph to graph

3. Please give comments or reactions to any of the following aspects of the course:

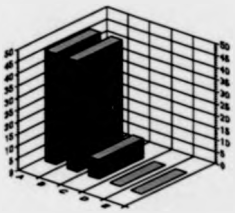
- **Quality of the inputs**



Comments:

Clearly delivered and well presented. Held the attention of course members. Perhaps a little more time for session 3.

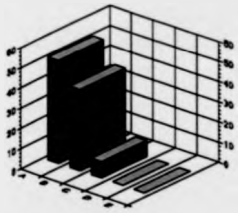
- **Form of the inputs**



Comments:

Interesting and illuminating. A good variety. Excellent supportive material. Perhaps more time for looking at projects.

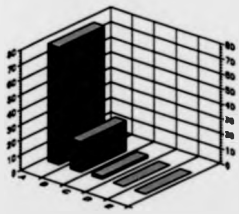
- **Order in which material was presented**



Comments:

Articulate and well presented. Would have appreciated an activity in the morning. Logical progression. Use of a flow chart to work through procedure may have been useful.

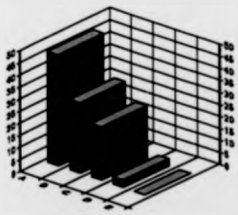
- **Quality and quantity of documentation**



Comments:

Excellent. Clear and easy to read. Generally very extensive and appropriate. Superb quality of support material.

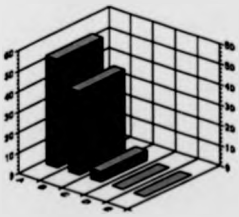
- **Interactions and discussion with others**



Comments:

Stimulating and very valuable. "We always need more time to do this!" Idea of small groups was liked. Anxiety started to surface. Longer on card game.

- **Enjoyment**



Other comments:

"You took away the fear!" Tiring but enjoyable. A very well organised course. Excellent INSET session leading to motivation of staff. Would like to have seen more examples of pupils' work. Very informative and extremely useful.

Model B - The enriched model was provided in Ealing and Hounslow. Some schools however decided that they preferred to be responsible for their own INSET. Of the schools which did receive INSET at a school level no significantly enhanced response in either pupil performance or standardising of assessments was detected. The sample was undoubtedly too small to overcome teacher/centre effects. Indeed it is possible that taking the responsibility of cascading INSET away from teachers had a detrimental affect. As a result, the coordinators may have failed to become totally familiar with all aspects of the material and provide the essential expertise required in each school.

Model C - The distant model of INSET resulted in a level of teacher satisfaction indistinguishable from the standard model. The LEA nominees were advisers and consequently experienced and highly competent INSET providers. It is unlikely that in a national trial individuals of a similar calibre will assume this responsibility. Efficient distance learning packs could be produced for the training of LEA based SAT coordinators - who would have responsibility for INSET, standardisation and moderation.

Model D - The WJEC provided an exemplary model in relation to the exam board model. The Board also had responsibility for the translation of Welsh medium material. It is quite apparent that all exam boards have the expertise and infra-structure to undertake this role if required.

Model E - There appeared to be no inherent problem in providing INSET for schools if a choice of SAT is offered. Teacher satisfaction was similar to that produced as a result of the standard model. INSET of this nature required more deliverers but the additional cost which this entails could be offset by increasing the numbers attending.

Model F - There was some value in providing pre-INSET material if it is used by all participants. If this is not the case, ill founded assumptions might be made or alternatively repetition could result in participant dissatisfaction. A practical task prior to INSET rather than a passive task such as reading information is preferable. If teachers perceive that the task addresses a real need - to gather information or answer some questions for example - they will attempt to do it.

SAT evaluation - teachers' response -Summer 1991

This appendix gives full statistical details of the teacher evaluation. The results are produced in the same style as the evaluation form. In completing this form, teachers did not answer every question. Percentages were, however, calculated on the basis of all teachers returning a form. Consequently, there is a percentage of missing data. This percentage has not been shown to improve clarity.

In the main report, wherever possible, 'd' responses have been coded into the three other categories. This is not the case in this evaluation. Two typical 'd' responses have been included in this evaluation to illustrate the kind of comments made by teachers. In addition teachers were invited to make comments on a separate sheet. The views expressed on these sheets forms the basis of chapter 7 - Findings on Implementations. Responses were received from 78% of all teachers involved in any aspect of the SAT; this includes support teachers, non-teaching coordinators etc. Of those responding we cannot be totally accurate as to teachers subject area as some claimed expertise in several categories.

Art & Design	10%
Business Studies	1%
Craft, Design & Technology	44%
Information Technology	3%
Home Economics	24%
Textiles	14%
Others	4%

Of those in the others category 40% (4 teachers) were scientists, 15% child development and the remainder English, engineering, needlework, PE, RE and woodwork.

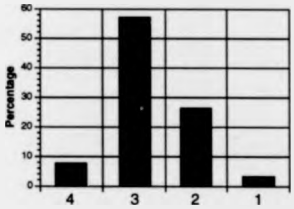
Teacher materials		**%**
A1	Did you find SAT Stages provided you with	
	a *insufficient information*	12.2
	b ***sufficient information***	**76.0**
	c *too much information*	6.9
	d "Insufficient info. for main project - first part fine More central resources need to be made avaiable."	2.0
A2	Did the format of the teacher kit make it	
	a *difficult to find what you required*	7.9
	b ***easy to locate information***	**61.5**
	c *less useful than a ring binder*	26.6
	d "format provided valuable resource ref. for future development/planning. Once all info out of section, difficult to put back quickly."	2.0
A3	In relation to delivering the SAT did the teacher materials provide you with	
	a *insufficient information*	20.1
	b ***sufficient information***	**71.4**
	c *too much information*	5.9
	d "Insufficient provision for SEN pupils."	1.6

A4 In relation to assessing the SAT, did the teacher kit provide you with

		%
a	*insufficient information*	13.5
b	***sufficient information***	**68.7**
c	*too much information*	9.5
d	"Confusing, diff. to understand/use. Easy to misinterpret. Too vague for valid interpretation."	5.6

A5 Please circle a rating for the overall effectiveness of the teacher kit

excellent			poor
4	3	2	1



The pupil materials
Video and quest

B1 In relation to the theme, did the video provide the pupils with

a	*insufficient information*	29.9
b	***sufficient information***	**48.4**
c	*too much information*	7.2
d	"Info did not focus them in useful direction. Too much for less able, too little for more able."	11.8

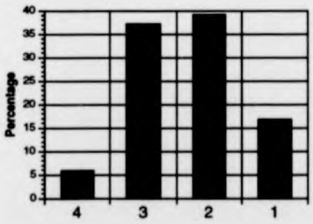
B2 When the pupils were investigating the theme, did the quest provide

a	*a poor focus*	37.8
b	***a satisfactory focus***	**45.7**
c	*a good focus*	8.9
d	"Not necessary, tended to use up time. Complicated matters.... Many "lower ability" pupils found it confusing. They needed to almost be told the answers."	5.6

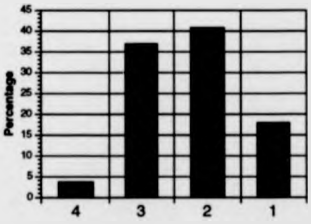
B3 Please circle a rating for the overall effectiveness of

	excellent			poor
a Video	4	3	2	1
b Quest	4	3	2	1

a



b



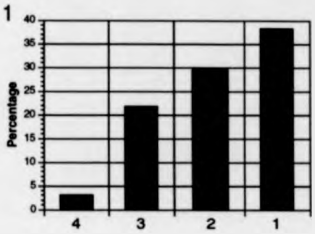
Short task %

C1 Did the short task

a	***serve no valuable purpose***	**40.8**
b	*help pupils get a feel for the theme*	31.3
c	*involve them to such a degree that they lost sight of the main task*	14.5
d	"Seemed quite worthwhile as exercise but not all pupils could relate this to later work. Not valuable but gave breathing space."	7.9

C2 Please circle a rating for the overall effectiveness of the short task

excellent 4 3 2 1 poor



Labels

D1 Are the labels in relation to Y9 pupils (levels 3 to 7)

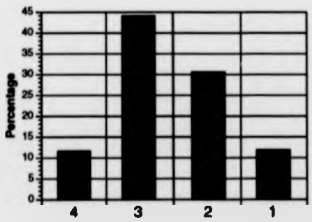
a	*too advanced*	30.9
b	***about right***	**53.9**
c	*rather simplistic*	1.0
d	"Pupils need more experience in using them Need further explanation to avoid ambiguity."	11.2

D2 Do the labels

a	*confuse pupils*	26.6
b	***help pupils structure their projects***	**36.5**
c	*help pupils structure their projects and provide good assessment evidence*	19.1
d	"With more experience, our students would have found them useful. Needed framework for design process first."	10.5

D3 Please circle a rating for the overall effectiveness of the labels

excellent 4 3 2 1 poor



Manageability %

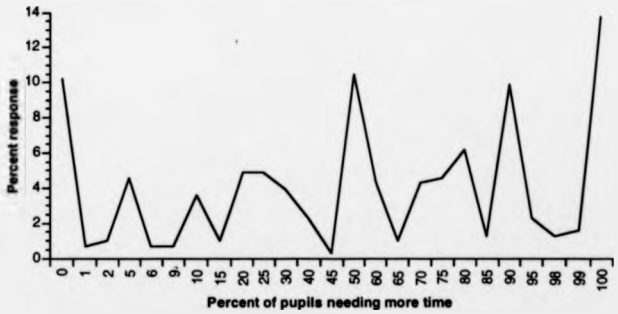
E1 In the context of NC design and technology, does this model of a SAT offer pupils

a	*little opportunity to demonstrate their capability*	23.7
b	***a reasonable opportunity to demonstrate their capability***	**54.9**
c	*too demanding a challenge*	9.5
d	"To do correctly needed skills/experience they did not have at this age... more practical skills needed. Demands skills over wide range yet diminishes time over when they can learn those skills."	9.5

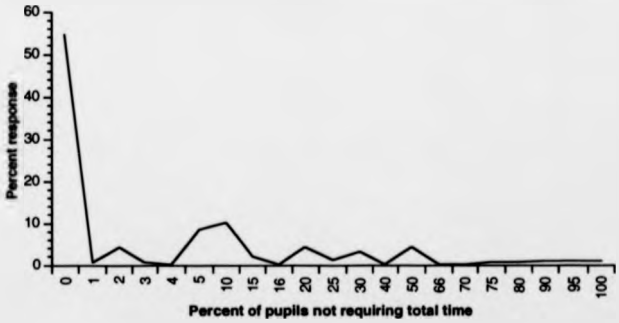
E2 The time the SAT allowed pupils to provide evidence of their capability was

a	***insufficient***	**48.7**
b	*reasonable*	30.9
c	*excessive*	8.6
d	"Insufficient particularly in resistive materials. Too much time on written work, insufficient time for making artefacts."	9.9

E3 Roughly what percentage of the pupils you taught required more time?



E4 Roughly what percentage of pupils you taught did not need all the time available?

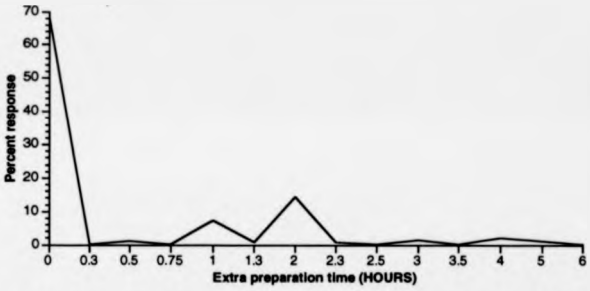


E5 Did the short task at the start of the SAT %

a	***serve no useful purpose***	**34.9**
b	*create time to negotiate briefs with individuals*	24.3
c	*prevent you from negotiating briefs as it took all the time*	26.6
d	"Useful as an intro but time taken imposed itself upon main task. Distanced pupils further from marking stages which caused impatience."	8.6

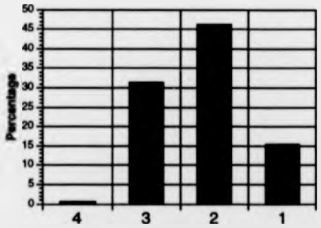
E6 Was the preparation time you required for the SAT

a	*less than normally required*	13.8
b	***about the same***	**40.1**
c	*more than normally required*	35.5
d	"Obviously more time as we hadn't done it before Difficult to know what preparation required and what involved for a test."	4.9



E7 Please circle a rating for the overall manageability of the SAT

excellent			poor
4	3	2	1



Assessment

F1 How much time would you normally devote to assessing a Y9 pupil's achievements during a term

a	*up to ten minutes*	11.8
b	*about fifteen minutes*	37.5
c	***more than twenty minutes***	**39.1**
d	"Difficult to say-assessment is a continuous process. I would be continually assessing a pupil's achievements."	8.6

F2	Did you manage to record pupils' achievement and progress	%
	a *during the lesson*	6.6
	b *after the lesson*	29.6
	c ***not at all***	**36.2**
	d "Impossible to service pupils with materials and help and keep any records during lessons-I was run off my feet!...During odd lessons when getting a 'feel' of what they had done."	17.4
F3	Was the pupil's SAT assessment based on	
	a *the SAT outcome (project folder and solution)*	27.6
	b ***the outcome and teacher recall of SAT performance***	**46.1**
	c *the outcome and teacher records made during the SAT*	13.2
	d "Combination of a, b and c...frequently assessing work of pupils we had not seen during SATs."	6.2
F4	Did you find the SAT sample assessments	
	a *of little help in making standardised assessments*	40.8
	b ***helpful in making standardised assessments***	**45.1**
	c *very helpful in making standardised assessments*	7.2
	d "Thought it helpful until standard I set was changed at moderation....Only knew of their existence during moderation."	5.3
F5	When making a pre-SAT assessment was the level guide	
	a *not very helpful*	22.0
	b ***helpful***	**56.3**
	c *very helpful*	15.1
	d "Language actually caused problems."	3.0
F6	Were the labels when assessing the work	
	a *of no use*	13.2
	b ***helpful***	**61.2**
	c *very helpful*	13.5
	d "Help where pupils had used them correctly. Of some help but often areas to be assessed were still under many labels and everything had to be re-read."	7.9
F7	Was the assessment instrument you used to determine the level secured by each pupil	
	a *too difficult to use*	11.8
	b ***initially difficult but OK with experience***	**61.8**
	c *straight forward and sensible*	11.8
	d "Difficult to understand- plainer English needed. Not always relevant to work in hand."	8.9

F8 Given the national curriculum Order for design and technology, was the assessment procedure

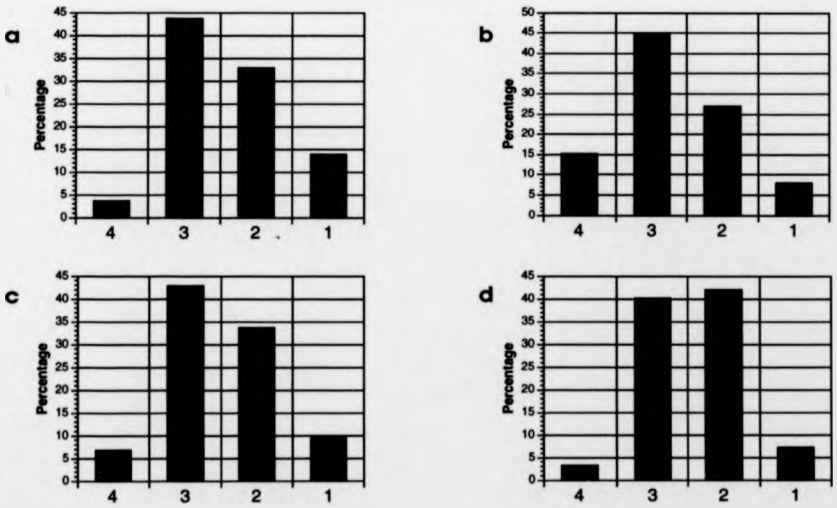
		%
a	*rather superficial*	13.8
b	***fair and objective***	**39.8**
c	*over complex*	36.2
d	"Complex- if results given to students, would they or their parents understand it... Open to misinterpretation or different interpretation by individual teachers."	3.9

F9 How long did the assessment of each pupil's SAT take

a	*up to ten minutes*	21.1
b	***about fifteen minutes***	**38.5**
c	*more than twenty minutes*	29.3
d	"Variable - depending on content, stucture of pupil's work."	6.9

F10 Please circle a rating for the overall effectiveness of

	excellent			poor
a SAT sample assessments	4	3	2	1
b level guide	4	3	2	1
c assessment instrument	4	3	2	1
d procedure	4	3	2	1



SAT issues **%**

G1 If your pupils had followed the KS3 programme of studies would they have found the theme

a	*narrow and restrictive*	16.4
b	***a fair challenge***	**57.2**
c	*too difficult*	5.3
d	"Unrealistic problem for Y9 ending in unreleted work. Too broad, pupils found it difficult to pick out one area."	9.2

G2 Suggest some themes which you think would be appropriate for a SAT

(See 3.15.5)

G3 Do you think the structure of the SAT was

a	*too loose*	32.6
b	***about right***	**40.5**
c	*too tight*	15.5
d	"Too biased towards information gathering not enough actual making...As with most new things, it could be both too loose or too tight for individual children."	5.9

G4 For pupils at the end of key stage 3 is this style of SAT

a	*too simple*	2.3
b	***appropriate***	**51.6**
c	*too complex*	27.3
d	"Depended on child's ability. More able - about right. Less able - incomprehensible...We have some able pupils - they tended to find it too simple."	12.5

G5 In relation to the normal level of resources (consumable materials etc.) did the pupils

a	*need less*	4.3
b	*need about the same*	25.3
c	***need more***	**56.6**
d	"Much wider/unpredictable range. Same but of different and varied materials."	8.9

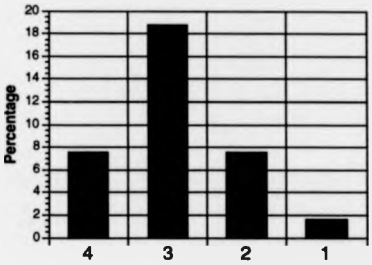
G6 Do you think that your normal approaches to organisation and teaching style changed during the course of the pilot

a	*not at all*	7.9
b	***in some aspects***	**59.9**
c	*considerably*	24.3
d	"Less organised, less structured, less actual teaching, more provider of resources/equipment. 'A' level teaching techniques are similar, is it appropriate at this level."	3.0

G7 Did you receive INSET for the SAT **No: 24.4% Yes: 75.6%**

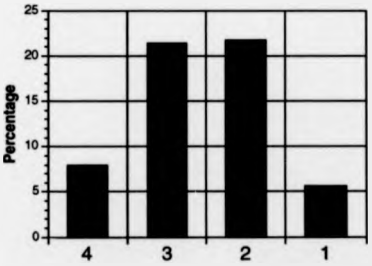
a If you received INSET from MEGNAP, could you rate its usefulness now the SAT is completed

excellent poor
4 3 2 1



b If you received INSET from colleague (cascade) could you rate its usefulness now the SAT is completed

excellent poor
4 3 2 1



To whom it may concern

Research and Development of Statutory Assessment in design and technology at key stage three - 1989 to 1994

As the other principal academic in the Middlesex University team contracted by the Schools Examination and Assessment Council to undertake the research and development necessary for statutory assessment, I wish to confirm the status of the research contained in this thesis. The research undertaken for SEAC was extremely wide ranging and dealt with many other aspects not covered in this thesis such as, information technology, special needs, culture fairness and issues relating to language. Richard Tufnell has drawn only on the aspects of the research for which he was responsible and which are rightly his intellectual property.

John Cave
Professor of Technology Education
Middlesex University

Bibliograpy

Primary sources

Blackburne, L. (12th February 1993). 'Probe into pass kits for technology tests'. *The Times Educational Supplement*

Blackburne, L. (20th March 1992). 'Technology test leak raises fears'. *The Times Educational Supplement*

Blackburne, L. (26th June 1992). 'Technology U turn starts'. *The Times Educational Supplement*

Boswell, T. MP, (13th May 1993). Letter to the Head Teacher Daventry William Parker School, Northamptonshire

Clarke, K., Secretary of State for Education, (November 1990). Interview in *The Times*

Clarke, K., Secretary of State for Education, (January 1991). Reply to tabled question, House of Commons, London: *Hansard*

Clarke, K., Secretary of State for Education (January 1991). Speech to the Society of Education Officers: Harrogate

Clarke, K., Secretary of State for Education (undated, received at SEAC 12th December 1991). Letter to Lord Griffiths of Forestfach, Chairman of SEAC

The Daily Telegraph Education Correspondent (7th May 1992). 'There's a spanner in the Curriculum' London: *The Daily Telegraph*

Duffy, M. Headteacher, King Edward VI School, Morpeth (February 1993). Reported in *The Times Educational Supplement*

Eggleston, J. (22nd November 1991). 'A fading flower'. *The Times Educational Supplement*

Graham, D. (May 1989). 'The Introduction of the National Curriculum', lecture, Cambridge Institute of Education

The Guardian (21st October 1991). 'Blue Peter lessons cost pupils dear'. *The Guardian*

The Guardian (7th January 1992). 'Tests for pupils of 14 to start in June'.*The Guardian*

Hardcastle, J. (12th March 1993). Correspondence Page. *The Times Educational Supplement*

Hutchinson, P. Deputy Head, Moulton School, Northampton (26th February 1993). Correspondence Page. *The Times Educational Supplement*

Ives, M HMI (1995). Keynote address - TEP Summer School 18th August, Middlesex University - unpublished

Kimbell, R., (16th. October 1992). 'A problem of success', *The Times Educational Supplement*

Lord Griffiths of Forestfach (2nd December 1991). Letter to the Rt. Hon. Kenneth Clarke QC MP, Secretary of State for Education

MEGNAP/Middlesex Polytechnic (Aug. 1990a). '1990 Trial - Report, Key Stage 3 design and technology', Middlesex Polytechnic: Technology Education Centre

MEGNAP/Middlesex Polytechnic (Nov. 1990b). '1990 Trial - Report: Key Stage 3 design and technology - Supplementary Report', Middlesex Polytechnic: Technology Education Centre

MEGNAP/Middlesex Polytechnic (Aug. 1991). 'Pilot 1991 - Report: Key Stage 3 design and

technology', Middlesex Polytechnic: Technology Education Centre

MUTEC, (Aug. 1992a). 'Pilot 1992 - Report: Key Stage 3 design and technology', Middlesex University: Technology Education Centre

MUTEC, (Dec. 1992a). 'Pre-Statutory Assessment Trial', Middlesex University: Technology Education Centre

MUTEC, National Curriculum Assessment at Key Stage 3 in Design and Technology, 1989 - 1994 - archive, Middlesex University: Technology Education Centre

Melkie, J. (3rd June 1993). 'Heads heap derision on Patten'. *The Guardian*

Melkie, J. (5th June 1993). '£35m cost of tests 'a waste of paper''. *The Guardian*

Midland Examining Group National Assessment Programme (1989). 'Proposals for the Development of Standard Assessment Tasks for pupils at the end of the third key stage of the National Curriculum', Cambridge: Midland Examining Group

NFER/Brunel University, (Dec. 1993). 'Confidential Report to SCAA on the 1993 National Assessment of English and Technology', Slough: NFER

Partington, P. President, Association of Head Teachers (2nd June 1993), Presidential address of the Association of Head Teachers Annual Conference, Newcastle

Parish, T. (26th February 1993). Correspondence Page. *The Times Educational Supplement*

Patten, J. The Secretary of State for Education (2nd June 1993). address to the Association of Head Teachers Annual Conference, Newcastle

Patten, J. (2nd June 1992), 'Statement to The House of Commons', London: *Hansard*

Pawsey, J. (January 1991), Question to the Secretary of State for Educatrion, House of Commons, London: *Hansard*

Schools Examination and Assessment Council. (1990). Arrangements for the 1990 Trial Exercise

Schools Examination and Assessment Council. (1991). 'PR 91/7 - SEAC awards contracts for fourteen-year-olds tests'.

Schools Examination and Assessment Council. (1989). Specification for the Development of Standard Assessment Tasks for pupils at the end of the third key stage of the National Curriculum.

Schools Examination and Assessment Council. (September 1989 July 1991). Minutes of the key stage 3 technology - subject steering committee meetings and correspondence

Schools Examination and Assessment Council, (September 1989 July 1991). Minutes of the key stage 3 - subject steering committee and correspondence

Shaw, B., Chief Adviser for Bolton. (5th February 1993). Reported in the *Times Educational Supplement*

Skidelsky, Lord, *The Sunday Telegrah*, (20th March 1993)

Tate,N. (September 1991 to November 1994-). All Technology SAT review meeting: letters, serving as minutes, recording main outcomes, Schools Examination and Assessment Council

The Times Educational Supplement. (September 1989 to November 1994)

Tufnell, R., (September 1989 to November 1994). All correspondence with the Schools Examination And Assessment Council and the Schools Curriculum and Assessment Authority

Published sources

Angoff, W.H. (1974). 'Criterion-referencing, norm-referencing and the SAT', *The College Board Review*, Volume 92

Apple, M.W. (1978). 'Ideology, reproduction and educational reform', *Comparative Education Review*, No. 22

Assessment of Performance Unit Project in design and technology(1985 -1990), Goldsmiths College: University of London

Board of Education (1926). *The Education of the Adolescent*. (The Hadow Report) London: Board of Education

Beadle, E., Prais, S.J., (1991). 'Pre-Vocational Schooling in Europe Today', London: *National Institute of Economic and Social Research, Report Series, Number 1*

Bernbaum, G & Mathieson, M. (1988). 'The British Disease: A British Tradition?' *British Journal of Educational Studies* Vol. XXVI No. 2

Blenkin,G.M., Edwards,G. & Kelly,A.V. (1992). *Change and the Curriculum*. London: Paul Chapman

Broadfoot, P., Dockrell, B., Harlen, W., & Nuttall, D., (Eds) (1993). 'Policy Issues in National Assessment', *BERA Dialogues 7*

Bruner, J. (1960). *Towards a theory of instruction*. Cambridge, Mass: Harvard University Press

Cox,C.B. & Dyson,A.E. (Eds) (1969/70). *Fight for Education: A Black Paper*. London: Critical Quarterly Society

Cross,A. & McCormick,B. (Eds) (1986). *Technology in Schools*. Milton Keynes: Open University

Daugherty, R. (1995). *National Curriculum Assessment - A Review of Policy 1987 - 1994*. London: Falmer Press

Dearing, R. (1993a). *The National Curriculum and its assessment: Interim Report*. London: National Curriculum Council/Schools Examination and Assessment Council

Dearing, R. (1993b). *The National Curriculum and its assessment: Final Report*. London: National Curriculum Council/Schools Examination and Assessment Council

Department for Education/The Welsh Office (December 1992). *Technology for ages 5 to 16*. London: HMSO

Department of Education and Science (1989). Circular No. 5/89, London: HMSO

Department of Education and Science, (1977). *Curriculum 11-16,' Working Papers by H.M. Inspectorate: a contribution to the current debate*. London: HMSO

Department of Education and Science (1988). *Education Reform Act*. London: HMSO

Department for Education and Science (1985). *GCSE General Criteria*. London: Department for Education and Science/Welsh Office

Department of Education and Science (1989). *Policy into Practice*. London: HMSO

Department of Education and Science, (1988a). *Task Group on Assessment and Testing - Report*. London: Department for Education and Science/Welsh Office

Department of Education and Science, (1988b). *Task Group on Assessment and Testing: Three Supplementary Reports*. London: Department for Education and Science/Welsh Office

Department of Education and Science (January 1991). Press release 22/91, London: Department of Education and Science

Department of Education and Science and the Welsh Office (1988). *The Interim Report of the National Curriculum design and technology Working Group.* London: HMSO

Department of Education and Science, Assessment of Performance Unit (1988). *Design and technology activity.* London: HMSO

Department of Education and Science/The Welsh Office (1989). *Design and Technology for ages 5 to 16.* (The Parkes Report), London: HMSO

Department of Education and Science/Welsh Office (1990). *Technology in the National Curriculum.* London: HMSO

Department of Education and Science - Assessment of Performance Unit (1982). *Understanding design and technology.* London: Department of Education and Science

Department of Science and Art: Form 813. *Suggestions for Manual Instruction - Circular 44* (5th June 1890). London: Department of Science and Art

Design Council The (1982). 'Schools Design Prize' The Competition Brochure. London: The Design Council

Design Council The (1980). *Design Education at Secondary Level.* London: Design Council

Despatches (1993). *All our Futures.* London: Channel 4 Television

Dodd, T. (1978). *Design and Technology in the School Curriculum.* Sevenoaks: Hodder & Stoughton

Eggleston, J. (1992). *Teaching Design and Technology,* Buckingham: Open University

Filer, D. (3rd June 1992) Press Release, London: The Engineering Council

Gifford, B.R. & O'Connor, M.C. (Eds) (1992). *Changing Assessments: Alternative Views of Aptitude, Achievement and Instruction.* Kluwer, Boston and Dordrecht

Gipps, C (Ed), (1992). *Developing Assessment for the National Curriculum.* London: Kogan Page

Gipps, C. (1992). *Assessment - A teacher's guide to the issues.* London: Hodder & Stoughton

Gipps, C.V. (1994). *Beyond Testing - 'Towards a Theory of Educational Assessment.* London: Falmer

Glaser, R. (1963). 'Instructional Technology and the Measurement of Learning Outcomes: Some Questions'. *American Psychologist,* Volume 18

Goldstein, H. (1994). 'Recontextualising Mental Measurement'. London: *Educational Measurement: Issues and Practice,* Volume 13

Goodson, I.F. (1988). *The making of Curriculum.* London: Falmer Press

Grady,D. (1976). 'Are we teaching bad design', in Proceedings - *Conference of the Institute of Craft Education.* London: Batsford

Graham, D. and Tytler, D (1993.) *A lesson for us all.* London: Routledge

Handwork in secondary schools' in *The Book of School Handwork.* (Ed). Holman, H. (1913) London: Caxton

Henessey, S.., McCormick, R., Murphy, P. (1993). 'The truth about problem solving in design and technology', *Design and Technology Times.* Autumn. Manchester: University of Salford

Hilsum,J. (1976). 'Assessment of design based project work' in *Studies in Design Education and Craft.* Volume 8,2

HMI Discussion Document (1978). *The Curriculum 11-16.* London: Department of Education and Science

HMI Report (1992). Technology - *Key Stages 1, 2 and 3; a report by HMI on the First Year 1990-91.* London: Department of Education and Science

Jenkins,D. & Shipman,M.P. (1976). *Curriculum: an Introduction.* London: Open Books

Jessop, G. (1991). *Outcomes: NVQs and the emerging model of education and training.* London: Falmer Press

Kelly, A. (1989). *The National Curriculum: A Critical Review.* London: Paul Chapman

Kimbell, R. (1982). *Design Education: The Foundation Years.* London: Routledge & Kegan Paul

Lawton, D. (Editor) (1989). *The Education Reform Act: Choice and Control.* Sevenoaks: Hodder & Stoughton

Lawton, D (1992). *Education and Politics in the 1990s: Conflict and Census?* London: Falmer Press

Layton, D. (1991). *Aspects of National Curriculum Design & Technology.* York: National Curriculum Council

Layton,D. (1972). 'Science as General Education'. *Trends in Education*

Layton, D. (1995). 'Constructing and Reconstructing School Technology in England and Wales' *International Journal of Technology and Design Education,* Volume 5

Marks, J. (1991). *Standards in Schools: Assessment, Accountability and the Purposes of Education.* London: Social Market Foundation

Marks, J. (1993). *Why there is no time to teach: What is wrong with the National Curriculum 10-level scale.* London: Social Market Foundation

Ministry of Education 15-18. (1959). *The Crowther Report.* London: HMSO

Ministry of Education. (1963). *'Half our Future'* (The Newsom Report,). London: HMSO

Ministry of Education. (1945.) *The Nation's Schools, Pamphlet No 1.* London: HMSO

Ministry of Education. (1945). *The New Secondary Education, Pamphlet No 9.* London: HMSO

National Centre for Schools *Technology (1983). Understanding school technology.* Loughborough: NCST

National Curriculum Council. (May 1992). *National Curriculum technology: The case for revising the Order.* York: National Curriculum Council

National Curriculum Council. (1989). *Consultation Report on Statutory Proposals for Technology.* York: NCC

National Council for Vocational Qualifications. (1991). *Guide to National Vocational Qualifications.* London: NCVQ

NFER/Brunel University. (1992). *Teacher Assessment in Mathematics and Science at key stage 3.* London: Schools Examination and Assessment Council

O'Hear, P. and White, J. (Editors) (1993). *Assessing the National Curriculum.* London: Paul Chapman

Penfold, J. (1988). *Craft, Design and Technology: Past, Present and Future.* Stoke-on-Trent: Trentham

Platt, J. (1972). *Report of Research into the Operation of the new pattern of BA Final Assessments.* University of Sussex

Porter, D.J. (1967). *A School Approach to Technology.* Curriculum Bulletin No. 2, London: HMSO

Pring, R. (1989). *The New Curriculum,.* London: Cassell

Rees, D. (1987). *Teaching GCSE Craft, Design and Technology.* London: Hodder and Stoughton

Ripper, W., *Technical Training as a main feature in National Education.* Report of the fifty sixth meeting of the British Association for the Advancement of Science

Robinson, P. & Smithers, A. (1992). *Technology in the National Curriculum - Getting it Right.* London: The Engineering Council

Rowntree,D. (1977). *Assessing Students: How shall we know them?* London: Harper & Row

Royal College of Art (1979). *The Three R' s' Design in General Education.* London: Royal College of Art

Saloman, O. (1892). *The Theory of Educational Sloyd*

Schools Council.(1975). *Education through Design and Craft.* Eggleston, J. London: Arnold

Schools Council. Working Paper 26, (1969). *Education through the use of Materials.* London

Schools Council. (1974). *Materials and Design: a fresh approach.* Eggleston, J. *et. al.* London: Arnold

Schools Council Field Report 3. (1966). T*echnology in Schools.* London: Schools Council

Schools Council Working Paper 18. (1968). *Technology and the Schools.* London: Schools Council

Schools Council - Examination Bulletin 23A. (1971). *Common Systems of Examining at 16+.* London: Evans/Methuen

Schools Curriculum and Assessment Authority. (1994). *GCSE Examination Results 1993.* London: Schools Curriculum and Assessment Authority

Schools Examination and Assessment Council. (1991). *Teacher Assessment at Key Stage 3.* London: Schools Examination and Assessment Council

Schools Examination and Assessment Council (1991). *Key Stage 3 Teacher Assessment in Practice,.* London: Schools Examination and Assessment Council

Schools Examination and Assessment Council. (1991). *National Curriculum Assessment at Key Stage 3: A review of the 1991 Pilots with implications for 1992.* London: EMU - Schools Examination and Assessment Council

Schools Examination and Assessment Council. (1992/3). *School Assessment Folder,* London: Schools Examination and Assessment Council

Schools Examination and Assessment Unit. (1991). *The Assessment of performance in design and technology.* (The Goldsmiths' Report) London: HMSO

Schools Examination Council and The Open University. (1986). *Craft, Design and Technology - GCSE A Guide for teachers.* Milton Keynes: OPen University Press.

Secondary Examinations Council. (1985). *Course work assessment in GCSE.* Working Paper 2

London: SEC

Schools Examination Council. (1986). *General Criteria for the General Certificate of Secondary Education.* London: HMSO

Shaw, K.E. (1991). *Teaching Design Technology -Perspectives 43,*. Exeter: Exeter School of Education

Slatterly, D. (1989). *Assessments in Schools.* Oxford: Basil Blackwell Ltd.

Technical and Vocational Initiative (1982). The Manpower Services Commission, London: Department of Employment

Thorndike, R.L. & Hagen, E. (1951, 2nd edition 1972). *Measurement. and Evaluation in Psychology and Education*, New York: John Wiley & Sons, page 193

Editorial, *View* (1981). London, Industry Education Unit, Department of Industry

Wood, R. (1993). *Assessment and Testing: A Survey of Research,*. Cambridge: UCLES

Young, M.F.D. (Ed) (1971). *Knowledge and Control: New Directions for the Sociology of Education.* London: Collier-MacMillan