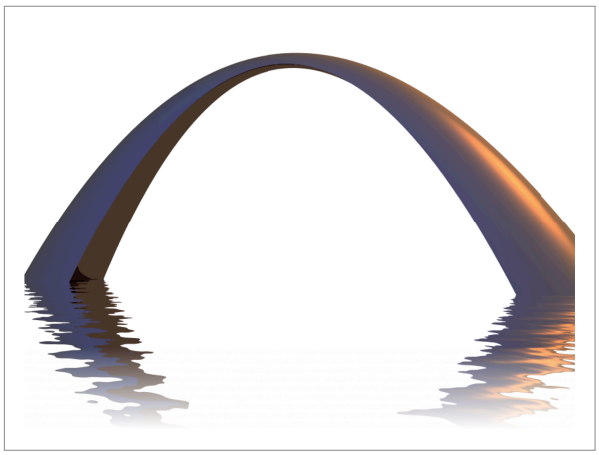


Crossover Communications Limited

Document Name: Final Report

Client: Ofqual

Project: Awarding Body Data Archive – GCSE Controlled Assessment Student Work Research Strand

Crossover Communications Limited,
Crown House, 231 Kings Road,
Reading, RG1 4LS, UK.
Tel: +44 118 909 8909 Fax: +44 118 909 8901
Email: enquiries@xover.co.uk Web: http://www.xover.co.uk

Contents

Executive Summary

This review has been carried out under the Awarding Body Data Archive (ABDA) programme, to consider the impact on GCSEs of the introduction of controlled assessment.

This report details the findings for GCSE Design and Technology: Food Technology, Geography, History and Music. It compares specifications, assessment materials and student work from 2009 (including coursework) and 2011 (including controlled assessment) collecting the views of a number of subject specialists.

Findings

Overall we found that internal assessment is necessary to enable students to demonstrate a wide range of subject related skills in a more practical context over a sustained period of time. In general controlled assessment sets out to assess an appropriate set of subject related skills, but is only fully successful in music. In each of the other subjects we identified specific skills or parts of skills which were not assessed in 2011.

However in each subject we considered that controlled assessment is appropriate to its purpose, and the improvements needed concerned the technical aspects of the assessment materials and the arrangements for assessment rather than the nature of the assessment.

Controlled assessment was considered to provide a more effective assessment tool than coursework in food technology. This related directly to the awarding organisation set tasks and an improved approach to marking, both of which are vital to ensuring all students can demonstrate their skills in this practical subject.

In geography, we found the approach to controlled assessment less successful, although we acknowledge the difficulties in providing an effective assessment of geographic fieldwork were also apparent in coursework. The concerns we found related to the relationship between fieldwork, skills and assessment rather than the nature of controlled assessment.

For history we found that coursework and controlled assessment both work well for different parts of the skill set. Controlled assessment does provide better opportunities for students to demonstrate their skills of historical interpretation, but these were not taken up in the student work we saw. There are also concerns about the comparability of the tasks set and the marking criteria used, although these were technical issues which can arise in every form of assessment.

We found that very little has changed for students in GCSE music, this is an intensely practical course and the approach to coursework and controlled assessment are very similar. However students who take the music technology option may be significantly more constrained by the time constraints on assessment than those who take the more instrumental route.

Section 1: Introduction

Context

The Awarding Body Data Archive (ABDA) programme was developed to collect evidence of student work in order to inform the evaluation of the Secondary Reform Programme. In particular, the evidence is collected to review particular components of qualifications and assessment.

This review considers the impact on GCSEs of the introduction of controlled assessment. In particular any changes in the skills students demonstrate, and how this differs from course work and external assessment.

The technical specification for this strand of the ABDA programme can be found at Appendix A.

Methodology

This review will examine two different specifications for each of four subjects within GCSE, the associated assessment instruments and samples of student work, by collating and analysing the views of a number of subject specialists.

The subjects to be reviewed are:

- Design and Technology: Food Technology
- Geography
- History
- Music

The following sections of this report detail the process of collecting and processing this information.

In this review we are considering the changes to GCSE qualifications when coursework was replaced by controlled assessment including:

- Controlled assessment and subject specific skills
- Comparison with external assessment

We may also generate evidence to inform discussions on the skills requirements for GCSE, and develop our approach to research.

The research and analysis framework for this review can be found at Appendix B.

Provision of assessment materials and student work

The ABDA Data Collection has a number of unique features that make it distinctive from other data/evidence collections:

- It provides a comparison over time for specifications that have been selected to be as similar as possible to previous years.
- Specifications across awarding organisations have been selected to be a similar as possible to each other.

- The cohort whose outcomes are being considered is similar between years.
- Student work is collected across all grades.
- The evidence is collected to review particular components of qualifications and assessment.

Full details of the materials drawn from ABDA for use in this review can be found at Appendix C and Appendix D.

The review team

Ofqual contracted a review team for each of the GCSE subjects included in this review.

Each team consisted of a seven experts in the GCSE subject, including a lead reviewer. All seven members of each team were contracted to review both the specification materials and student work.

A full list of reviewers can be found in Appendix K.

Analysis and comparison of the specifications and assessment materials

The review team members analysed the awarding organisations' materials using a set of forms, an example of which can be found in Appendix F.

The analysis using Form A is designed to:

- identify and evaluate the skills tested within each assessment component
- analyse the coverage and demand of skills within each assessment component
- evaluate the way in which the mark scheme and assessment criteria recognise and reward skills

Every team member analysed both sets of specification materials for their subject, and then compared the two using Form B.

The comparison using Form B is designed to compare:

- the skills tested within each assessment component with the skills specified within the GCSE controlled assessment regulations
- the structure of the assessment components
- the effectiveness of the marking criteria

The review team had an opportunity to discuss their overall findings before the lead reviewer produced their report. This report brought together the views of the whole subject team. These findings are presented in Section 2 of this report.

Analysis of student work

The review team members compared and analysed the student work using a set of forms, an example of which can be found in Appendix G.

To assess students' demonstration of subject related skills, all the reviewers were brought together for a two day meeting to analyse students' scripts (pieces of student work collected for ABDA). This process is referred to as a script review.

The meeting started with a briefing session to make sure all the reviewers had a common understanding of the methodology and the judgement criteria.

The scripts were organised into packs for consideration during the review. For this exercise each whole set of a student's work was separated into two scripts. The controlled assessment or coursework component was made up an internally assessed script and the written papers an externally assessed script.

As far as was possible, given the collection of scripts available, packs contained seven or eight scripts, two controlled assessments, two coursework and three or four written papers.

Reviewers were then asked to rank the scripts in each pack, from the best to the worst demonstration of subject related skills, using Form C as a data-entry sheet and to make comments on the scripts as necessary. Each reviewer completed a maximum of thirteen sessions during the two day meeting.

Reviewers also analysed sets containing all the work from individual students, and recorded their findings using Form D.

During the script review there were plenary sessions for the reviewers to discuss the script review process and the nature of the skills demonstrated in the scripts being analysed.

Data analysis

Ofqual use a software package called FACETS to analyse the results from data-entry sheets produced during the script review. FACETS uses a Rasch model (often classified under item response theory) to convert the qualitative ranking decisions made by reviewers into a single list for each subject which reflects the probable overall order of the student scripts, from the best to the worst demonstration of subject related skills.

We use these lists, alongside the qualitative comments made during the script review and findings from the specification review, to inform Section 3 of this report.

Section 2: Subject related skills

The most significant change to the structure of GCSE qualifications between 2009 and 2011 was the introduction of controlled assessment to replace coursework for assessing subject related skills.

The GCSE controlled assessment regulations (QCA 2008) define the subject related skills which must be assessed through controlled assessment for each GCSE subject.

In controlled assessment, the levels of control are designated at three key control points: task setting, task taking and task marking. The intention of the controls introduced was “to ensure reliability and authenticity” (QCA 2008, p4), in the internal assessment of student work.

The reviewers considered the way in which subject related skills were assessed within each specification for their subject. They did this by analysing the specification documents and the assessment materials used in 2009 and 2011. Details of the specifications included in the review are given in Appendix C.

2.1 Design and Technology: Food Technology

“The following skills must be assessed through controlled assessment:

- designing creatively
- making products
- applying systems and control, CAD/CAM, digital media and new technologies
- analysis and evaluation on processes and products.

In addition, elements of these skills may be assessed externally”

(QCA 2008, p32)

Overall, when comparing the two specifications there have been few changes in the skills students are expected to develop and demonstrate. Both specifications list all the subject content under either Designing Skills or Making Skills, and the lists are very similar. In the specification used in 2011 some skills are defined in more detail and the designing skills specifically emphasise creativity and innovation.

The schemes of assessment had the same weighting with 60% weighting for the internally assessed component and 40% for the external assessment. In 2009 the externally assessed component was tiered with students entered for either the foundation or higher tier written paper. By 2011 this was no longer the case and all students took the same written paper. In both years there was no separation within the internally assessed component.

Students had most opportunities to demonstrate designing and making skills in the internally assessed component in each year. However, the written papers also required some demonstrations of their designing skills and some of the skills linked to planning. The written papers also allowed students to apply practical knowledge developed through their designing and making experiences. For ex-

ample the 2009 question asking why electronic scales are important in the making of food products.

Internally assessed components

The controlled assessment in 2011 did give students opportunities to demonstrate their designing, although they did not have the opportunity to develop their own design brief for the practical work. However this could have been a hurdle for lower achieving students and can be addressed instead through the written paper.

The tasks allowed a wide range of making skills to be demonstrated and were accessible to students of all abilities. There were also opportunities for high achieving students to show their skills at a very high standard.

The applying of systems and control, CAD / CAM, digital media and new technologies is really limited by the facilities available to schools. Students can use IT including digital images in preparing their design folder, but the 2011 assessments did not address these skills in any depth.

The controlled assessment tasks did encourage students to carry out analysis and evaluation of their own products by including judgemental criteria in the question. For example: “sensory properties which appeal to teenagers”, “could be sold from a supermarket's 'Luxury' range”.

Task setting

In 2011 the level of control for task setting was high and the awarding organisation provided twelve design tasks for assessment in May 2011 and 2012. The tasks are expected to be reviewed each autumn and a revised list issued for the following academic years.

Teachers are permitted to contextualise the design tasks but only in consultation with the awarding organisation. Although the actual design tasks are still only a single line, each is presented within a context which makes clearer the expectations of the final product.

This is different from 2009 when the specification guidance on setting the centre assessed component (coursework) includes a list of twelve single line “project outlines” which teachers could use or adapt. Teachers could also prepare their own project outlines without approval from the awarding organisation.

Task taking

The level of control for task taking was medium, and the awarding organisation has provided guidance in the notes which accompany the design tasks. This includes expectations of the measures teachers should use to ensure only students' own work is included for assessment, together with examples of situations where students may be able to complete work or collect materials outside of the classroom.

There is also more detail about the way in which teachers should give feedback to students, and how this should be recorded.

In 2011 the specification recommends students spend approximately 45 hours on their controlled assessment task, an increase from 2009 when the coursework

project was expected to be “completed satisfactorily in 40 hours”. However the controlled assessment guidance does make it clear students are expected to produce a “focussed, concise and relevant” design folder. Placing the emphasis on quality rather than quantity of work.

Task marking

In principle the approach to marking of the internally assessed components has remained the same. Student work is marked by teachers and the results are subject to local and awarding organisation moderation to ensure performance is recognised and rewarded consistently. However, the specifications in this review show two different approaches to internal assessment.

In 2009 marking relied upon separate best-fit assessments of a student's Designing Skills, Making Skills, and Quality of Written Communication. The marking bands for Designing Skills and Making Skills were very wide, with very similar or identical criteria across some grades, increasing the risk that a large number of markers would fail to interpret them consistently.

For the 2011 assessments the awarding body provided five detailed sets of assessment criteria:

- investigating the design opportunity
- development of design proposals (including modelling)
- making
- testing and evaluation
- communication

These criteria rewarded the same achievements as the previous method but each set is numerically based within defined mark bands. The distinct differences between the bands support consistent marking and make clear the mark allocations within the controlled assessment ensuring students are rewarded for all aspects of their performance. It also gives useful guidance on the nature and quality of work students need to produce in order to succeed.

The type of marking is able to reward students who perform at different levels in the five areas of the assessment criteria.

Externally assessed components

The written paper in 2011 did assess some aspects of both “designing creatively” and “analysis and evaluation on processes and products”. The designing skills tested were closely tied to the pre-release material and this could limit students' opportunity to demonstrate creativity and innovation. There was some testing of designing in the written examination but the limited space on the papers constrained candidate’s ability to respond to the question.

There was also some analysis of product data and questions about devising test procedures. However these were mainly tests of student knowledge rather than an opportunity to demonstrate some skills.

2.2 Geography

“The following skills must be assessed through controlled assessment:

- identifying, analysing and evaluating geographical questions and issues
- establishing appropriate sequences of investigation incorporating geographical skills, including enquiry skills
- extracting and interpreting information from a range of different sources
- evaluating methods of collecting, presenting and analysing evidence, and the validity and limitations of evidence and conclusions.

In addition, elements of these skills may be assessed externally”

(QCA 2008, p44)

There are no significant differences between the two specifications.

The schemes of assessment had the same weighting with 25% weighting for the internally assessed component and 75% for the external assessment. In 2009 the externally assessed component was made up of two written papers weighted at 45% and 35%. By 2011 the externally assessed component was divided into three equally weighted written papers. In both years the written papers were tiered to allow students to take either the foundation or higher tier, however there was no separation within the internally assessed component.

The majority of the subject related skills were assessed via a fieldwork report. However the written papers continue to provide opportunities for students to use geographical communication skills in producing graphs, diagrams and sketch maps and also to assess students' extracting and interpreting skills.

Internally assessed components

The controlled assessment in 2011 did give students opportunities to demonstrate a wide range of geography related skills, although sometimes within a limited scope. In addition written paper 1 specifically addresses cartographic, graphical, geographical enquiry, ICT and Geographic information System (GIS) skills.

In 2011 the limited topic choice and direct reference to ICT and Geographical Information System (GIS) skills restricted the candidates development of their enquiry. To some degree this reflected the practical changes to fieldwork already made by schools in response to risk assessment.

Task setting

In 2011 the level of control for task setting was high and the awarding organisation provided eight tasks for assessment between June 2011 and May 2012. A new selection of tasks will be issued each year.

Each task consists of a single question, on a specific theme which is then given context by the choice of location or area to conduct the fieldwork.

This is very different from 2009 when the specification guidance describes the requirements for the students to undertake geographical investigations supported by fieldwork, but does not include any example tasks or require any awarding or-

ganisation approval. The teacher is specifically responsible for ensuring that the task is appropriate to the ability of individual candidates.

Task taking

The level of control for task taking was separated between:

- Research and data collection – limited level of control
- Analysis, conclusions and evaluation of findings – high level of control

The specification for 2011 suggests the students should spend one day on data collection in the field and about 20 hours in the classroom preparing for fieldwork, processing the data and preparing the final report. There is also clear guidance to teachers on the use of students' secondary research in the final report, which should be approximately 2000 words if produced in written format. The specification provides a list of alternative formats in which the report can be presented.

In 2009 the report should have a maximum length of approximately 2000 words and can include additional supporting material in an alternative format.

Task marking

In principle the approach to marking of the internally assessed components has remained the same. Student work is marked by teachers and the results are subject to local and awarding organisation moderation to ensure performance is recognised and rewarded consistently. The specifications in this review have very similar approaches to internal assessment.

For the 2009 assessments the awarding body provided five detailed sets of assessment criteria. Marking in 2011 was based on five very similar criteria, with the addition of a sixth, evaluation.

- purpose of investigation
- methods of collecting data
- methods of presenting data
- analysis and conclusions
- evaluation
- planning and organisation

Each set of criteria is numerically based within defined mark bands. The distinct differences between the bands support consistent marking and make clear the mark allocations within the controlled assessment ensuring students are rewarded for all aspects of their performance. It also gives useful guidance on the nature and quality of work students need to produce in order to succeed.

The type of marking is able to reward students who perform at different levels in the five or six areas of the assessment criteria.

Externally assessed components

The written papers in 2011 focus on extracting and interpreting information but with more opportunities than in 2009 to communicate graphically and through

adding to sketch maps. There is a considered effort to include many of the Geographical Skills listed within the specification except for those more appropriately assessed through the Controlled Assessment.

There was a notable increase in the use of multiple choice questions in the Foundation Tier written papers which focussed mainly on simple recall and basic skills. The Higher Tier written papers required extended writing to provide explanations and details of case studies, mainly testing recall rather then the application of knowledge or skills.

2.3 History

“The following skills must be assessed through controlled assessment:

- historical enquiry
- historical interpretation.

In addition, elements of these skills may be assessed externally”

(QCA 2008, p51)

There are no significant differences between the particular options within the two specifications chosen for consideration in this review..

The schemes of assessment had the same weighting with 25% weighting for the internally assessed component and 75% for the external assessment. In both years the externally assessed component was made up of two written papers weighted at 45% and 35%, with all students sitting the same paper for the topics they had studied; i.e.: no tiered papers in either year.

The 2009 internal assessment required students to produce two pieces of extended writing (approximate maximum 1250 words) which could be on the same or different topics. In 2011 students completed just one piece of work (approximate maximum 2000 words).

Internally assessed components

The controlled assessment in 2011 did offer a significant opportunity for students to demonstrate their ability to comprehend, analyse and evaluate, in relation to historical context, how and why historical events, people, situations and changes have been interpreted and represented in different ways. In addition the written papers test the skills of historical communication and enquiry to varying degrees.

There were changes in the presentation of the tasks from 2009 coursework to 2001 controlled assessment. The 2009 tasks focus on a particular aspect of history and ask questions which provide structure for the students response. The 2011 tasks are single questions which require adaptation by the teacher.

Task setting

In 2011 the awarding organisation offered four options for learning programmes leading to the controlled assessment:

- the role of the individual in history
- a thematic study in Twentieth Century history

- a modern world depth study
- a study in depth

the actual content of the learning programme was chosen by the teacher.

The level of control for task setting was set at high and the awarding organisation provided four generic tasks for assessment (one per option) which will change each year. However each task then had to be adapted by the teacher to fit their students' learning programme. It does not appear that these adaptations, or the learning programmes in which they were based had to be submitted for awarding organisation approval.

In 2009 the awarding organisation provided six coursework topics and a coursework guide which included tasks for assessment. Alternatively teachers could devise their own coursework assignments, and even their own coursework topics, but these required prior approval from the awarding organisation.

Task taking

The level of control for task taking was high and the awarding organisation provided specific guidance on the application of controls in the specification.

In 2011 the specification recommends students spend eight classroom hours to complete their controlled assessment task, four hours on preparation (research, note taking, planning and drafting) and four hours to write up their work. Students can be given the task one week before the controlled assessment starts in order to plan their approach,but there are restrictions on the material they can then take into the controlled assessment.

There is no guidance in the 2009 specification on the amount of time students should spend on their coursework, but there is clear guidance about teacher supervision of student work.

Task marking

The specifications in this review show two very similar approaches to internal assessment. Student work is marked by teachers and the results are subject to local and awarding organisation moderation to ensure performance is recognised and rewarded consistently. In both years the mark scheme is closely related to the Assessment Objectives defined within the specification, however the structure has changed considerably.

In 2011 the teacher had to consider all three Assessment Objectives together and award a mark within one of six possible bands. With the exception of Band 0 (Candidates submit no evidence or fail to address the question), each band covers aspects from each Assessment Objective within a range of up to twelve marks.

The mark scheme for Controlled Assessment, by failing to identify how the marks in each band related to the three assessment objectives, did not give centres and their candidates sufficient guidance on what was required in order to answer the questions. Furthermore the grouping of these assessment objectives into five very broad bands covering 50 marks in total gave insufficient guidance to centres and teachers to ensure consistency of marking, and hence reliability of standards.

Externally assessed components

The written papers were broadly similar in the two years considered, providing effective tests of knowledge and historical understanding. Both papers also provide students with some opportunities to demonstrate their ability to investigate historical questions and to use historical sources critically in context.

2.4 Music

"The following skills must be assessed through controlled assessment:

- performing/realising and composing music.

In addition, elements of these skills may be assessed externally"

(QCA 2008, p75)

There are no significant differences between the two specifications considered in this review.

The schemes of assessment had the same weighting with 60% weighting for the internally assessed component and 40% for the external assessment.

In both years the internal assessment required students to produce two performances and two compositions.

Internally assessed components

For 2011 the performing criteria had been altered so students no longer had to perform a piece from the same Area of Study as one of their own compositions. They also had an increased choice of performance options with the inclusion of, for example, DJ-ing. The mark scheme was modified slightly to take account of the choices available and to provide greater differentiation between work offered by lower ability candidates to higher.

In composition students were given more freedom to choose the Area of Study and/or style within a specific Area of Study which they used.

The levels of control which apply to the different elements are not specifically expressed in the specification.

Task setting

In 2011 the awarding organisation provided tasks for the students by describing the requirement for two performances (one solo and one ensemble) and two compositions or one composition and one arrangement. These requirements were less restrictive than in 2009 when students had to compose two pieces according to a brief, and one performance had to be a piece from the same Area of Study as one of the compositions.

The composition assessment in 2009 also included a short written response to show their understanding of the brief.

The level of control for task setting in Music is medium.

Task taking

In 2011 the specification states students must be supervised regularly in the preparation of their performances, and they have a maximum of 10 hours to complete the recording of each performance. It is not clear how the time limit relates to the production of a very short recording of either an individual or ensemble performance. More specific guidance is only given for the music technology option, where this amount of time may actually be required.

Composition is divided between research, which can be carried on outside the classroom, and write up and recording which must be completed under teacher supervision within a maximum of ten hours per composition.

There is no guidance in the 2009 specification on the amount of time students should spend on their coursework, or about teacher supervision of student work although there is an emphasis on teachers being able to authenticate the student's work.

The level of control for task taking in Music is medium.

Task marking

The specifications in this review show two very similar approaches to internal assessment. Student work is marked by teachers and the results are subject to local and awarding organisation moderation to ensure performance is recognised and rewarded consistently.

The wording of the performance mark schemes had been changed from 2009 to provide further clarity and to support consistency between the large numbers of markers.

For composition, the allocation of marks within the scheme was adjusted to reflect the removal of the brief.

The level of control for task marking in Music is medium.

Externally assessed components

In 2011, all the music extracts for the written paper were set works which had been analysed during the course. Therefore questions could be testing recall rather than aural skills, in particular for the extended writing questions.

Section 3: Findings

Each subject review team considered the student work drawn from ABDA for use in this review. Details of the materials used can be found at Appendix D.

Interpreting the charts

The findings for each subject are represented in two charts.

The chart showing the measure for each subject shows the spread of the student work as produced by the FACETS software. Each bold line indicates the measure related to the relevant ranked script. The difference between sequential measures demonstrates the strength of the difference in the ranking position. Large differences would illustrate that scripts were less close in terms of similarity of the student's demonstration of subject related skills than small differences. So there could be a larger difference in the demonstration of subject related skills between scripts ranked 1 and 2 than those ranked 2 and 3 (the difference in the demonstration of skills is not necessarily the same between ranked positions). These charts can be found at Appendix I.

The other charts, included in the text below, illustrate the standard error of measurement (SEM) to the corresponding range of measures. Because of the number of separate measures it would be very complex to illustrate the SEM for each individual measure. Therefore the range of measures are represented by the block and the largest extent of the SEM is represented by the whiskers.

The SEM illustrates the level of confidence that the measure is accurate: the greater the SEM, the smaller the confidence levels. Therefore large whiskers mean that there is less confidence that the measure was accurate. The whiskers illustrate the level of confidence, with upper and lower points at which the measure could lie.

The FACETS software will usually produce a rank order, even when there is little difference between the quality of the student work considered in the review. This is due to the natural slight variability between students who get the same mark. However, in this review the student work is spread across the full range of performance from A* to F for each year. This means we expect to see a distinct rank order with widely spread measures which demonstrate clear findings. However the wide range of student performance typically produces large SEM at the top and bottom of the rank order and therefore relying on the extreme measures may provide misinformation.

The student work has been separated by the type of assessment: controlled assessment (2011), coursework (2009) and written papers (both years).

3.1 Design and Technology: Food Technology

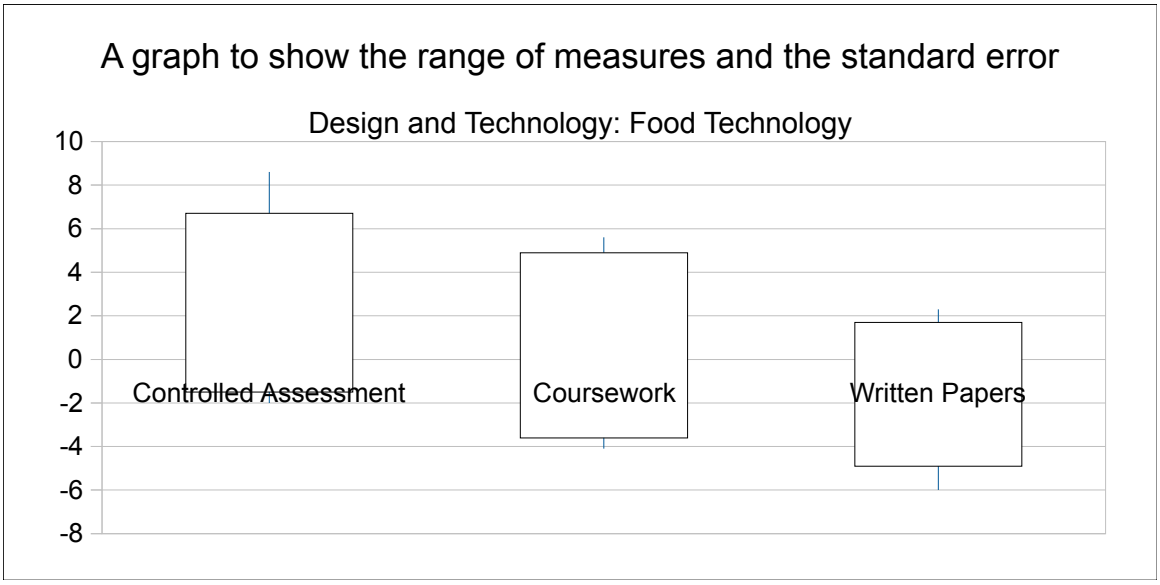
The FACETS rank order of student work within each type of assessment was very similar to the level of performance as measured by the original mark schemes.

The nineteen pieces of controlled assessment work were all judged to be above the lower quartile, and included five of the seven highest ranked pieces of work. All of the student work which had been awarded A* to B grades was placed in the

upper quartile. Only three pieces of controlled assessment work were placed below the median.

The student work which spreads most widely across the rank order is the coursework. However twenty of the twenty-six pieces of work were judged to be in the top half of the rank order.

The largest set of work is the written papers combining 2009 and 2011. None of the written papers were judged to be in the upper quartile, and thirty-seven of the forty-eight pieces of work were judged to be in the lower half of the rank order.



The rank order suggests that students demonstrate their designing and making skills better in response to controlled assessment, as compared to coursework and written papers.

Similarly, it suggests that students demonstrate the least designing and making skills in response to written papers.

These findings demonstrate that it is essential for students to have practical opportunities to develop and demonstrate a range of skills. Several of these skills cannot be demonstrated in a written paper, particularly those related to activities where students are required to use equipment, and work efficiently in terms of time, materials, ingredients and components.

Reviewers noted there was no evidence within the student work provided for this review that students were able to demonstrate their skills in applying systems and control or CAD/CAM. Similarly, the structure of the assessment tasks restricted opportunities for creativity in design.

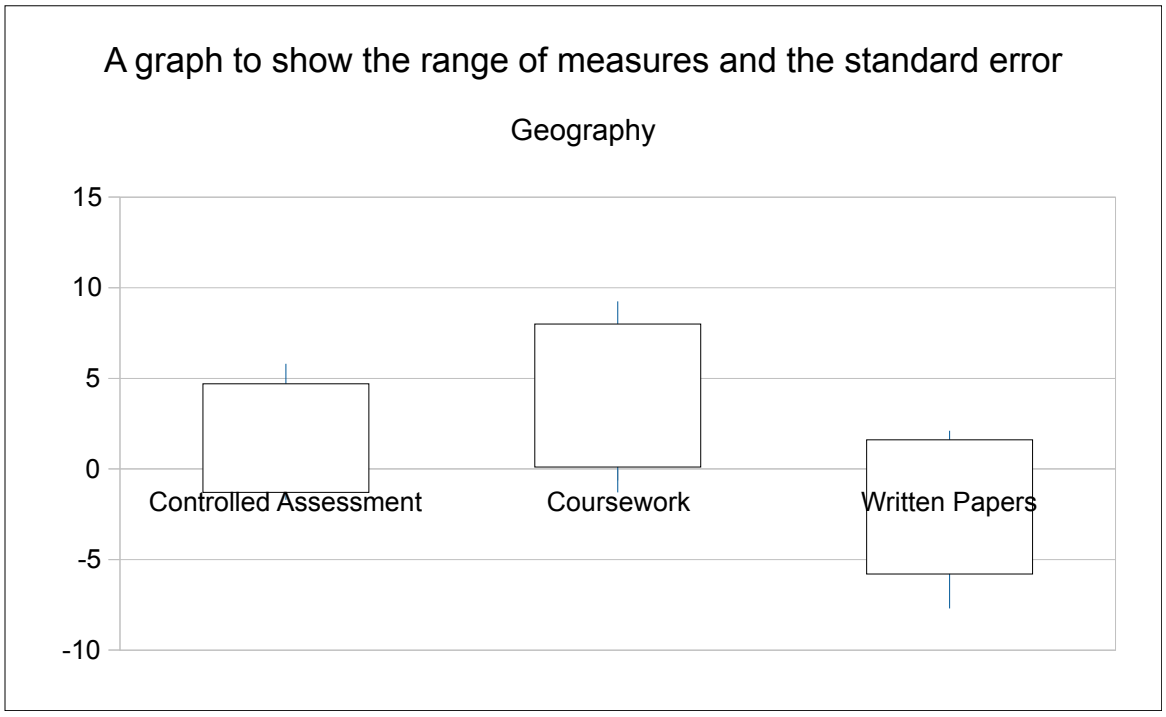
3.2 Geography

The FACETS rank order of student work within each type of assessment was similar to the order by level of performance as measured by the original mark schemes.

The twenty-seven pieces of controlled assessment work were all judged to be above the lower quartile, and included three of the top ten highest ranked pieces of work.

The student work which is most concentrated within the rank order is the coursework with only two pieces of work placed below the median, and none in the lowest quartile.

The largest set of work is the written papers combining 2009 and 2011. Only one of the written papers was judged to be in the upper quartile, with a further seven of the fifty-four pieces of work placed in the upper half of the rank order.



The rank order suggests that students demonstrate their geographic skills better in response to the coursework compared to the controlled assessment and written papers.

Similarly, it suggests that students demonstrate the least geographic skills in response to written papers.

This may reflect the changes to fieldwork brought about by the introduction of controlled assessment tasks. The tasks set by the awarding organisation reduces the opportunity for students to identify their own geographical questions and establish appropriate sequences of investigation. This restricts the range of skills students can develop and demonstrate. This is particularly restrictive for more capable students.

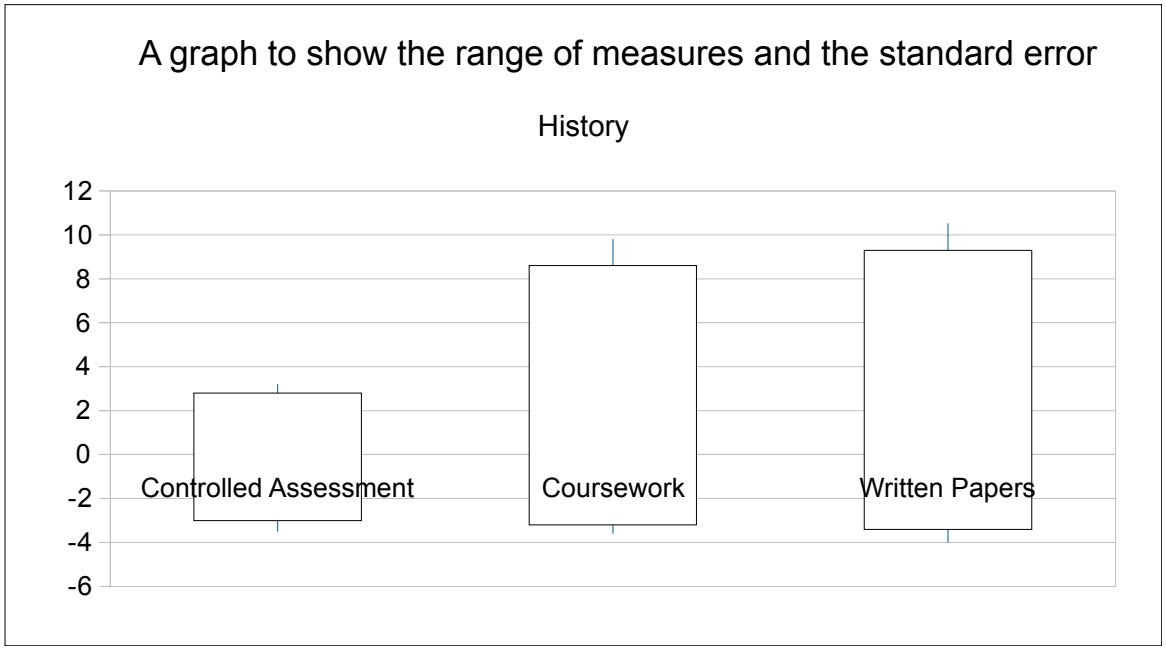
These findings confirm it is essential for students to have fieldwork opportunities in order to develop and demonstrate a range of skills. Most of these skills cannot be demonstrated in a written paper, particularly those related to activities where students are required to collect their own data for analysis, develop and evaluate their own conclusions and present their findings.

3.3 History

The FACETS rank order of student work within each type of assessment was similar to the level of performance as measured by the original mark schemes.

All three types of assessment work were distributed evenly through the rank order. A slightly greater proportion of the coursework pieces were placed above the upper quartile and a slightly greater proportion of the controlled assessment and written papers were placed below the lower quartile.

This lack of distinction between the assessment types reflects both the nature of the skills required for history, and the opportunities provided by the materials included in this review. Historical enquiry and interpretation require students to consider information in the light of their knowledge of its origin and context before selecting the most appropriate parts to use in support of their answer. However students opportunities to demonstrate the full range of skills were limited through all three types of assessment.



3.4 Music

The FACETS rank order of student work within each type of assessment was similar to the level of performance as measured by the original mark schemes.

All three types of assessment work were distributed throughout the rank order.

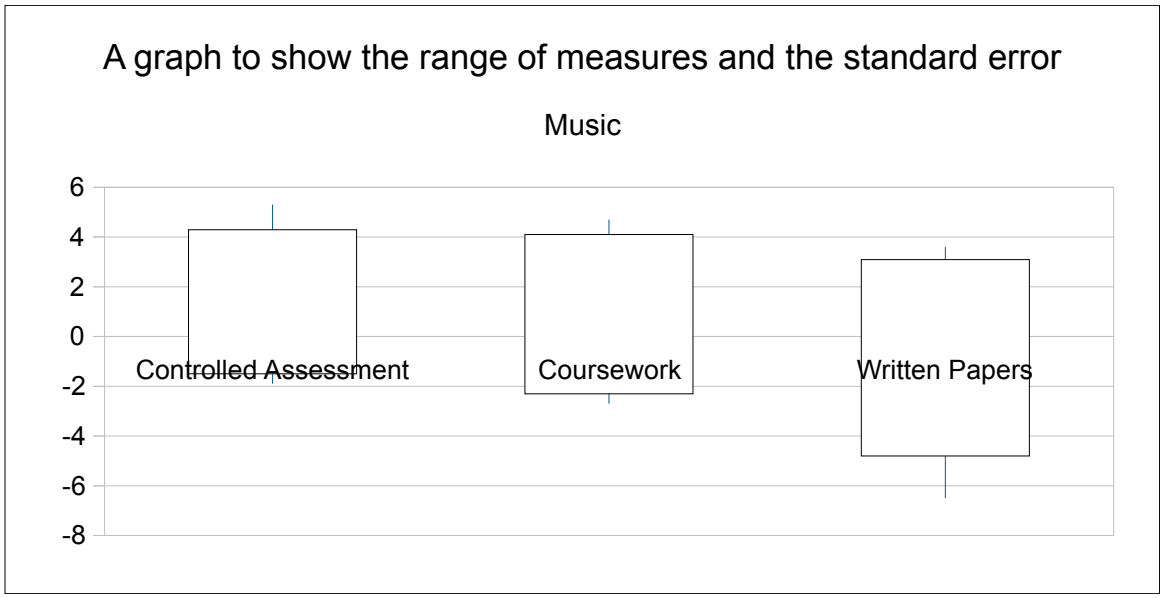
Fourteen of the sixteen controlled assessment pieces were judged to be above the lower quartile, along with all but one piece of coursework.

A significant proportion of the written papers were placed in the lowest quartile.

The outcomes probably reflect the nature of the skills required for music. Performance/realisation and composition are practical activities and these are directly tested by the controlled assessment and coursework.

The music written papers also require musical skills, as students must listen to pieces of music, recognise them and relate their knowledge of the piece to

provide an answer. However these skills are distinct from performance/realisation and composition.



The music review team were less confident of their judgements than the other subject teams because of the poor quality of the student materials available to them, with incomplete sets of documents and difficulties finding the correct recordings when required. The written paper sets were generally better presented. In addition the music review team had to consider three very distinct sets of subject related skills. All of these factors may have influenced the reliability of these judgements.

However, the similarities in the judgements for controlled assessment and coursework confirm the minimal changes between the 2009 and 2011 specifications.

Section 4: The research questions

4.1 Design and Technology: Food Technology

Controlled assessment and subject specific skills

1. In relation to skills controlled assessment is designed to test;

a. what subject specific and generic skills are demonstrated in students course work and controlled assessment component/ unit?

Candidates have the opportunity to demonstrate a wide range of the designing skills as listed in the specification. However the word creativity has been widely interpreted and in some cases design ideas were not creative.

In considering the skills controlled assessment should be designed to test there were some which were not demonstrated in either the written papers or the controlled assessment. In particular the use of systems and control, CAD / CAM, digital media and new technologies were not explored. The only opportunity was for candidates to make use of digital cameras for photographs.

Students had opportunities to demonstrate a wide range of making skills. Again, some were not demonstrated i.e. those related to industrial practices.

Opportunities to demonstrate evaluation, problem solving and the skills required to analyse processes and products were all provided in the controlled assessments.

b. is there variation in the types of skills that students are demonstrating in course work in comparison with controlled assessment?

c. what skills are seen in both assessment forms?

The skills were generally the same in both controlled and coursework assessment.

The use of the word creatively suggested a change in focus for 2011 in the controlled assessment, but this was not seen in the student work reviewed.

2. To what extent is controlled assessment restricting innovation or creativity in: task type, task setting and task taking?

The task type remains very similar to coursework, but tasks are more clearly defined by the awarding organisation. The tasks were certainly more relevant in 2011 compared to 2009, with a wide variety of contexts is provided.

Teachers are allowed to modify the context to fit with their local situation. This ensures the task is relevant for the students, but by presenting the tasks as a design brief this limits the opportunity for students to demonstrate their creativity by developing their own.

In comparison with external assessment

3. Can we see evidence of assessment of subject related skills in the controlled assessment unit which are not evident in the external assessment?

There are a number of designing and making skills which cannot be assessed in the external assessment, particularly those related to the production of a practical outcome or which require students to apply their knowledge; for example: working accurately, working efficiently, devising, selecting and applying test procedures.

a. are these skills essential for progression to A levels?

This aspect was not specifically considered during this review.

b. are we testing the right skills aspects of the syllabus for controlled assessment?

It is clear from the work done during this review that controlled assessment is assessing the correct skills. However, there may be a need to reconsider whether it is necessary to continue to include the application of systems and control, and use ofCAD/CAM, digital media and new technologies, as skills if these will not be tested.

4. Are the differences between controlled assessment and external assessment due to assessment form or other issues e.g. tasks type?

The main difference is that in the controlled assessment the students produce their food products supported by a portfolio that demonstrates how they have progressed through the design process. This is the design and make activity.

The external assessment tests knowledge, understanding and a very limited range of design skills. This balances the controlled assessment by covering a wider range of knowledge and ensuring that students have the knowledge and understanding to inform the application of their skills.

Skills requirement for GCSE

5. Is the controlled assessment component or unit appropriately meeting the overall subject aims/ grade descriptions and marking criteria required for GCSE level?

The specification used for 2011 allows candidates to show the wide range of skills which are appropriate for GCSE Food Technology. The improved marking criteria ensured that student's performance could be recognised and rewarded consistently. The design of the mark scheme clearly differentiates between different levels of performance.

The student work reviewed showed a wide range of performance across all grade descriptions.

Developing approaches to research

6. To what degree do subject experts agree with each other?

The reviewers had different approaches to the home-based review which meant the level of agreement was not always clear.

However the evidence from the script review event shows similar judgements were made, with all the reviewers recognising controlled assessment as the better method for assessing designing and making skills.

In discussions during the student work review event a wide variety of views were expressed. This was expected given the different backgrounds of the reviewers, only two are currently involved with teaching and examining GCSE Food Technology.

4.2 Geography

Controlled assessment and subject specific skills

1. In relation to skills controlled assessment is designed to test;

a. what subject specific and generic skills are demonstrated in students course work and controlled assessment component/ unit?

Subject specific and generic skills that are designed to be tested are essentially similar in coursework and controlled assessment. These include all the components of:

- methods of collecting data
- methods of presenting data
- analysis and conclusions
- evaluation
- planning and organisation

Demonstration of the skills within purpose of investigation is limited by the nature of controlled assessment.

b. is there variation in the types of skills that students are demonstrating in course work in comparison with controlled assessment?

c. what skills are seen in both assessment forms?

The skills were generally the same in both controlled and coursework assessment. However students who produced coursework in 2009 did have more opportunity to identify a geographical question and establish a sequence of enquiry.

2. To what extent is controlled assessment restricting innovation or creativity in: task type, task setting and task taking?

The tasks are provided by the awarding organisation and must be used as they are set. The teacher does provide context by deciding where the fieldwork should done..

The student work provided for this review suggests that fieldwork is arranged, and data collection organised and structured by the teacher for both coursework and controlled assessment. However, some more able students did show individuality in additional data collection or in choice of location.

Data presentation, analysis and conclusions is often where creativity is seen in coursework and for controlled assessment this work is subject to high control and time restrictions which may inhibit the students. However the controlled assessment in 2011 did allow the report to be presented in a variety of formats; e.g. written report, dvd, power-point, personalised GIS maps and web-pages which may

encourage creativity and innovation. However only written reports were available for this review.

In comparison with external assessment

3. Can we see evidence of assessment of subject related skills in the controlled assessment unit which are not evident in the external assessment?

The review found that all the skills listed in the specification for 2011 are demonstrated (or opportunities provided) and assessed through the Controlled Assessment, with the exception of 'identifying geographical questions and issues' and 'establishing appropriate sequences of investigation'.

However, only 'extracting and interpreting information from a range of different sources' is demonstrated (or opportunities provided) through the written papers.

a. are these skills essential for progression to A levels?

This aspect was not specifically considered during this review., however enquiry skills, in this case based on fieldwork and in a geographical context, are essential for progression to A levels and HE.

b. are we testing the right skills aspects of the syllabus for controlled assessment?

Fieldwork based enquiry skills are assessed most effectively through controlled assessment or a similar form of internal assessment. However the identification of questions and establishing appropriate sequences of investigation was not assessed effectively in the assessment materials and student work provided for this review.

There may be a need to consider how these important skills can be more effectively developed and demonstrated by students.

4. Are the differences between controlled assessment and external assessment due to assessment form or other issues e.g. tasks type?

Within Geography the different types of assessment are structured to assess particular areas of the specification content. This includes the balance of the Assessment Objectives targeted in the different components

Skills requirement for GCSE

5. Is the controlled assessment component or unit appropriately meeting the overall subject aims/ grade descriptions and marking criteria required for GCSE level?

There is a tension between the requirements for controlled assessment and the opportunities required by the subject criteria for students to undertake the full enquiry process individually.

Developing approaches to research

6. To what degree do subject experts agree with each other?

The reviewers took similar approaches to the home-based review and there was a strong agreement in their findings..

The evidence from the script review event shows similar judgements were made, with all the reviewers expressing concerns about the need to improve controlled assessment as a method for assessing geographical enquiry skills.

4.3 History

Controlled assessment and subject specific skills

1. In relation to skills controlled assessment is designed to test;

a. what subject specific and generic skills are demonstrated in students course work and controlled assessment component/ unit?

Historical communication and enquiry skills were demonstrated in both Coursework and Controlled Assessment. However historical interpretation was only demonstrated in a small number of instances for coursework, in response to explicit questioning. It was largely missing from Controlled Assessment.

b. is there variation in the types of skills that students are demonstrating in course work in comparison with controlled assessment?

c. what skills are seen in both assessment forms?

Coursework was more effective in enabling students to demonstrate their skills in evaluating historical sources whilst the controlled assessment was more effective in enabling students to demonstrate their skills in using sources to undertake an historical enquiry. The controlled assessment did allow students to demonstrate their skills in evaluating sources but evidence of this tended to be explicit only in the work of higher attaining students whereas in the coursework the precise questions targeting these skills ensured that students did demonstrate these skills.

The evidence from the 2011 students' work sample was that Controlled Assessment was not successful in testing historical interpretation. In just 2 of the 17 students was any evidence of this skill seen, and that at a low level.

2. To what extent is controlled assessment restricting innovation or creativity in: task type, task setting and task taking?

Disappointingly what could not be seen is the critical use of sources with explicit evaluation nor the consideration of historical interpretations although both were theoretically to be tested by Controlled Assessment.

In comparison with external assessment

3. Can we see evidence of assessment of subject related skills in the controlled assessment unit which are not evident in the external assessment?

The skills that can be seen in the Controlled Assessment unit that are not evident in the external assessment are the skills of enquiry in terms of using sources to investigate an historical question and the skill of writing at considerable length with time for reflection. Disappointingly what could not be seen is the critical use of sources with explicit evaluation nor the consideration of historical interpretations although both were theoretically to be tested by Controlled Assessment.

a. are these skills essential for progression to A levels?

This aspect was not specifically considered during this review.

b. are we testing the right skills aspects of the syllabus for controlled assessment?

The skills targeted of enquiry, communication and historical interpretations are the right skills to be tested by Controlled Assessment but from the evidence of the scrutiny of the students' work currently they are not being successfully assessed.

4. Are the differences between controlled assessment and external assessment due to assessment form or other issues e.g. tasks type?

They are due to the different type of tasks and to the differing conditions and time scales under which such tasks are undertaken.

Skills requirement for GCSE

5. Is the controlled assessment component or unit appropriately meeting the overall subject aims/ grade descriptions and marking criteria required for GCSE level?

The Controlled Assessment task does not adequately meet the subject aims. All reviewers agreed that the four questions that centres are required to adapt to their specific needs and circumstances were not of comparable demand and that this lack of comparability would be exacerbated as centres made their own amendments. From the scrutiny of the candidates' work is was also evident that some teachers had been more successful in tailoring one of the four questions to their needs.

Developing approaches to research

6. To what degree do subject experts agree with each other?

In terms of their interpretation of the A and B forms and their application to the specification materials there was a reasonable level of agreement across the review team.

The open definitions of the three historical skills on Form B did allow for individual interpretation which broadened the range of points to emerge.

For Form C the consideration of student work there was a pleasing consistency of judgement amongst reviewers at the top and bottom of the ability range but perhaps inevitably in the middle there was far greater variation.

4.4 Music

Controlled assessment and subject specific skills

1. In relation to skills controlled assessment is designed to test;

a. what subject specific and generic skills are demonstrated in students course work and controlled assessment component/ unit?

b. is there variation in the types of skills that students are demonstrating in course work in comparison with controlled assessment?

c. what skills are seen in both assessment forms?

In the main, it has been found that the types of skills that students demonstrate remain consistent between the coursework and controlled assessment.

The most significant change is the abandonment of compositional briefs together with the requirement to write a short explanation showing "understanding of the brief". This was an assessment point in 2009.

It was found that the skills tested in both assessment forms were very similar. Differences were that there was a greater choice of performance options in 2011 including multi- track recording and DJ-ing. The mark schemes were slightly rewritten. The actual wording of the assessment criteria hints at the subject specific skills acquisition rather than directly testing those skills. However, it was agreed that those skills that were tested were appropriate for a GCSE music based course.

In Performing, marks were awarded for Technical Control, Expression, Interpretation and Ensemble Playing, These were under the heading of "ACCURACY" and "INTERPRETATION" and were altered slightly to take account of the type of performance student was giving. These marks were adjusted for the difficulty of the piece using a grid provided by the awarding organisation.

In composing, the core skills were assessed under the banner of "use and development of ideas" "exploitation of the medium" and "structural intent" along with any two (2009) or three (2011) optional criteria from "melody" harmony and accompaniment", texture," "rhythm," "dynamics" and "use of technology".

2. To what extent is controlled assessment restricting innovation or creativity in: task type, task setting and task taking?

Because there was little difference between the coursework opportunities in 2009 and controlled assessment in 2011, there was no evidence that controlled assessment was restricting innovation or creativity.

However, students opting for a more technological route are disadvantaged by having to work to a strict time scale in comparison to students who took a more traditional route to, especially for performance.

There is more freedom for composition because students did not have to composed to set brief in 2011 and were given more freedom to express their skills within their chosen area of study.

In comparison with external assessment

3. Can we see evidence of assessment of subject related skills in the controlled assessment unit which are not evident in the external assessment?

The skills seen of performing and composing are essential skills for musicians. These skills are practical in nature and are the bedrock of music education from National Curriculum to life in the professional world. Skills of performing and composing test candidate's creativity, dexterity, an ability to communicate emotionally, culturally with precision and verve.

These skills cannot be assessed through a written paper, and have not been targeted in either of the written papers included in this review..

a. are these skills essential for progression to A levels?

This aspect was not specifically considered during this review, however at A-level, composing and performing are still a significant part of the course. However this review did not include a detailed consideration of this question.

Students taking the more technological pathway through GCSE music may be better prepared for A-Level Music Technology or BTEC rather than A-level Music.

b. are we testing the right skills aspects of the syllabus for controlled assessment?

The consensus was that the skills being assessed for the CA are entirely appropriate and required for success in the course.

4. Are the differences between controlled assessment and external assessment due to assessment form or other issues e.g. tasks type?

The different forms of assessment are necessary to allow the student to demonstrate the wide range of knowledge, understanding and skills required.

Skills requirement for GCSE

5. Is the controlled assessment component or unit appropriately meeting the overall subject aims/ grade descriptions and marking criteria required for GCSE level?

The mark schemes given for the controlled assessment modules were, by most respondents found to be meeting the overall subject aims. Respondents were happy that the mark schemes gave sufficient guidance to the assessor about the requirements to achieve the marks available, describing the difficulty levels required for certain performance tasks. However, as in the previous coursework assessment, there is still a level of subjectivity involved is assessing music and this was in evidence in the rank ordering tasks reviewers sometimes were at odds with some of the composition and performance grades of identical cohorts. The relation of subject aims to the mark schemes are clear although the detailed skills set as specified in the syllabus are more implicit within the successful completion of the performance or composition rather than explicitly assessed. The marking and assessment opportunities were deemed to be consistent between the 2009 examination and 2011 with the 2011 examination being found to have addressed some of the concerns about the lack of assessment of key music skills such as notation.

Developing approaches to research

6. To what degree do subject experts agree with each other?

The level of agreement between reviewers was high throughout the whole process. Many of the responses were consistent in the nature of the written comments made and in the presentation of those comments.

5 References

QCA 2008: GCSE controlled assessment regulations published by the Qualification and Curriculum Agency in February 2008. QCA/08/3512. http://www2.ofqual.gov.uk/downloads/category/186-controlled-assessment-regulations last accessed 12 October 2012

Appendix A: ABDA - GCSE Controlled Assessment Student Work Research Strand

Technical Specification – July 2012

Introduction

This technical research specification outlines the research objectives, methodology and timeline for GCSE Controlled Assessment Student Work Research Strand of the ABDA programme of work (Strand 3) and its fit with the controlled assessment review, which Ofqual launched in April 2012.

Context

The Awarding Body Data Archive (ABDA) project was developed to collect evidence of student work in order to inform the evaluation of the Secondary Reform Programme. The project has a number of unique features that make it distinct from other data/evidence collections:

- It provides a comparison over time for specifications that have been selected to be as similar as possible to previous years.
- Specifications across AOs have been selected to be a similar as possible to each other.
- The cohort whose outcomes are being considered is similar between years.
- Candidate work is collected across all grades.
- The evidence is collected to review particular components of qualifications and assessment.

The overarching purpose of the longitudinal ABDA research programme is to build the evidence base to:

- monitor and understand the impact of changes in the qualifications; researching the impact key modifications have had on GCSE and GCE assessment and qualifications forms.
- identify emerging intended and unintended consequences, behaviours, practices or strategies at the earliest opportunity to inform Ofqual on whether the market is acting efficiently or creating perverse incentives.
- establish the impact it has on learning and learning outcomes.

The move between course work and controlled assessment is one such change, and the ABDA project is well placed to provide a unique perspective on this.

As part of the 2007 review of GCSE qualification and subject criteria, the Qualifications and Curriculum Authority (QCA) commissioned a report on internal assessment in GCSEs. The report, published in June 2007[1], recommended that coursework be replaced, in the majority of subjects, with controlled assessment. Consequently controlled assessments were introduced with the first teaching of

1 Colwill, Dr I. (2007) Improving GCSE: internal and controlled assessment, (QCA)

the revised GCSEs in September 2009. They take place under supervised conditions and are either set by awarding organisations and marked by teachers, or set by teachers (under guidelines set out by awarding organisations) and marked by awarding organisations. These requirements have been set out in the subject criteria and controlled assessment regulations.

We have been monitoring the effectiveness of controlled assessment since it was introduced. We now believe that a more comprehensive review is required to further understand any issues related to controlled assessment and assess the usefulness of this assessment. As a result, as part of the review of controlled assessment we have commissioned three pieces of work, these are:

1. Subject review of criteria and specifications – to produce initial findings on whether controlled assessment in each of the subjects is valid and supportive of good teaching and learning, or the ways in which it could be improved.
2. Call for evidence - to secure the views of and collate evidence from a range of stakeholders on the effectiveness of controlled assessment.
3. Controlled assessment student work research – comparative analysis of student coursework, controlled assessment and external units work for 2009 and 2011 cohort.

The evidence from each of these strands will be brought together and we will report the outcome of the whole review in the autumn.

Research objectives

This strand of the research will review a sample of work from a number of students across four subjects comparing GCSE 2009 with 2011 cohort. The aim is to understand in detail student work/performance outcomes as a result of the introduction of controlled assessment in comparison to coursework and external assessment. This will inform the overall review of controlled assessment providing evidence on the effectiveness of controlled assessment in assessing subject related skills, the impact it has on student assessment and suggest ways in which it could be improved.

Therefore, research questions for this project are:

Controlled assessment and subject specific skills

1. In relation to skills controlled assessment is designed to test;

a) what subject specific and generic skills are demonstrated in students course work and controlled assessment component/ unit?

b) Is there variation in the types of skills that students are demonstrating in course work in comparison with controlled assessment?

c) What skills are seen in both assessment forms?

2. To what extent is controlled assessment restricting innovation or creativity in: task type, task setting and task taking?

In comparison with external assessment

3. Can we see evidence of assessment of subject related skills in the controlled assessment unit which are not evident in the external assessment?

a) Are these skills essential for progression to A levels?

b) Are we testing the right skills aspects of the syllabus for controlled assessment?

4. Are the differences between controlled assessment and external assessment due to assessment form or other issues e.g tasks type?

Skills requirement for GCSE

5. Is the controlled assessment component or unit appropriately meeting the overall subject aims/ grade descriptions and marking criteria required for GCSE level?

Developing approaches to research

6. To what degree do subject experts agree with each other?

Methodology and deliverables

Research method

The proposed methodology is (please note: the development of the research and analysis framework and the systematic review may result in the modification of the research method):

In broad terms the research will progress through 4 stages

Step 1	Literature Review to inform the development of research instruments	Research team (contractor working in research team)
Step 2	Expert review of specifications, question papers and mark schemes.	Panels of subject experts
Step 3	Expert review of candidate work and internal statistical analysis of cohort outcomes.	Panels of subject experts Research team
Step 4	Analysis of findings of steps 1-3 including analysis of expert rating, coding and interpretation of lead review reports and production of final report	Research team (contractor working in research team)

The sampling strategy

The table below lists the evidence collated for ABDA programme for this study we will be reviewing the subjects highlighted in yellow.

GCSE subjects	% controlled assessment & levels of control	Data collection year	Notes on the data
Geography	25%	2009-2011-2013	2009 - 13 candidates at F Tier and full set at H Tier 2011 – full set
History	25%		2009 – full set 2011 - full set
Food Technology	60%		2009 – full set 2011 - full set
Music	60%		2009 – full set 2011 - full set

A number of subjects collected for ABDA have been excluded for the following reasons:

- French[2] & Spanish - A detailed scrutiny of MFL has just been completed. The evidence from this study will be reviewed for this study.
- Religious Studies[3] & PE - No controlled assessment or nature of assessment means it is not collected for ABDA.
- English Literature, English, Mathematics, ICT (collected 2009-2012-2013) - These subjects will not be included in this study, as we need to collect the 2012 data. Plus Maths has no controlled assessment

Deliverables

The required outcomes of the work are as follows:

- Research and analysis framework
- Final Report - The final report will include subject specific chapters and synthesise of the findings and conclusions.
- Discussion document - outlining the key findings of this study and recommendations for further analysis and areas to explore longitudinally.
- Project data - A fully cleaned and validated NVivo file, containing all the evidence sourced and collected.

Management of the project and internal reporting

2 ibid
3 ibid

The work will be led by Somia Nasim and Mags Bexon and we will report to, Matt Glanville, Cath Jadhav and Dennis Opposs (Project sponsors). Project meetings will be set up accordingly - to discuss progress, any issues and initial findings.

Project Timeline – High level plan

Deadline	Stage	Activity	Assigned to
30th July	Project set – up	1. Confirm research questions, methodology and timetable. 2. Invite AOs and TSAR Group to comment on methodology with AOs 3. Recruit subject experts and education expert. 4. Recruit education expert (to support Steps 1-4).	Research Team Sign-off - Project team/ sponsors
22nd Aug	Step 1 - Evidence gathering	5. Research and analysis framework designed 6. Systematic evidence and / literature review 7. Recruit subject experts (to conduct the analysis -Steps 2-3).	Research team (contractor working in research team)
21st Sep	Step 2 and 3	8. Review of scripts and accompany assessment and specification documents. 9. Lead reviewer subject reports completed	Led by education expert and review conducted by subject experts
5th Oct	Step 4 Analysis and reporting	10. analysis of Steps 1-3 11. initial analysis report 12. project team review findings workshop.	Research team (contractor working in research team)
26th Oct	Final report	Final report for sign-off Discussion document for project team to review.	Research Team Sign-off - Project team/ sponsors

Budget

Standards and Research budget approved.

Appendix B: Research and Analysis Framework

1. Project overview

This research strand consists of a comparative analysis of student coursework, controlled assessment and external units work for 2009 and 2011 cohort.

The 2009 GCSE examination series included Coursework as an assessment component. By the 2011 series all Coursework had been replaced by Controlled Assessment.

The GCSE subjects chosen for this review reflect a range of approaches to internal assessment by both Coursework and Controlled Assessment. The subjects are:

- Design and Technology: Food Technology
- Geography
- History
- Music

The exact specifications, selected for 2009 and 2011 have been chosen to provide a close match of subject content and qualification structure.

The findings from this comparative analysis will contribute to the overall review of controlled assessment. In particular by:

- providing evidence on the effectiveness of controlled assessment in assessing subject related skills
- determining the impact controlled assessment has on student assessment
- identifying ways in which controlled assessment can be improved

2. The research questions

The research questions for this project, as detailed in the Technical Specification, are:

Controlled assessment and subject specific skills

In relation to skills controlled assessment is designed to test;

what subject specific and generic skills are demonstrated in students course work and controlled assessment component/ unit?

Is there variation in the types of skills that students are demonstrating in course work in comparison with controlled assessment?

What skills are seen in both assessment forms?

To what extent is controlled assessment restricting innovation or creativity in: task type, task setting and task taking?

In comparison with external assessment

Can we see evidence of assessment of subject related skills in the controlled assessment unit which are not evident in the external assessment?

Are these skills essential for progression to A levels?

Are we testing the right skills aspects of the syllabus for controlled assessment?

Are the differences between controlled assessment and external assessment due to assessment form or other issues e.g. tasks type?

Skills requirement for GCSE

Is the controlled assessment component or unit appropriately meeting the overall subject aims/ grade descriptions and marking criteria required for GCSE level?

Developing approaches to research

To what degree do subject experts agree with each other?

Research questions 1. (all parts), 2., 3. b) and 4. will be addressed within the scope of this methodology.

Question 3. a. addresses whether the subject related skills assessed through controlled assessment are essential for progression to A levels.

This is a significant question, and investigating it will require materials and resources outside the scope and timescale of this current research. The Ofqual Standards Over Time programme has been developing ways to consider progression from GCSE to A levels. The programme recruits subject experts with experience of qualifications at both levels and uses the same team of reviewers to consider both GCSE and A level specifications in the same subject. Consideration will therefore be given to incorporating relevant findings from the most recent Standards Review reports.

Question 5 asks whether the "*controlled assessment component or unit appropriately meeting the overall subject aims/ grade descriptions and marking criteria required for GCSE level?*"

Again this is a significant question which relates to the original accreditation of the specifications considered and is closely aligned with the work in the Ofqual Scrutiny and Standards Over Time programmes. Again the materials and resources required to conduct a thorough investigation are outside the scope and timescale of this project. Consideration will therefore be given to incorporating relevant findings from the most recent Scrutiny and Standards Review reports.

Research question 6 can only be answered by further analysis of the subject reviewers' judgement and comments after the comparative analysis has been completed. This further analysis will not be achieved within the timescale of this project and therefore is not addressed with the scope of this methodology.

3. The methodology

The methodology is based upon the approach used by Ofqual teams working on a number of projects across national qualifications, general and vocational, and international research.

It is designed to utilise the expertise of the seven reviewers on each subject panel to process a variety of information about the different forms of assessment. The judgements made by the reviewers will be collated and processed by the lead reviewer.

The methodology consists of:

- home based analysis of awarding organisation specifications and assessment materials[4] for each subject by a team of reviewers
- review of student work (A* to F grades) by the team of reviewers who completed the home based analysis
- completion of a summary report of the overall findings for each subject by the lead reviewer

3 a) Home based analysis

The comparative analysis will be carried out by a team of reviewers making independent, qualitative judgements, using two proformas. These judgements must be supported by clearly documented evidence drawn from the qualification specification and assessment materials.

The process consists of two activities: assessment analysis and assessment comparison.

Activity 1: Assessment analysis

The reviewers will complete an assessment analysis form (Form A) for each specification. This includes:

- identifying and evaluating the skills tested within each assessment component
- analysing the coverage and demand of skills within each assessment component
- evaluating the way in which the mark scheme and assessment criteria recognise and reward skills

The Form A is completed by each reviewer for each of their subject's specifications. Reviewers extract information from the specification and assessment materials in order to complete the form. This provides information about the 2009 and 2011 versions of the GCSE in a directly comparable format.

Activity 2: Assessment comparison

Each reviewer will then complete an assessment comparison form (Form B) for the two specifications they have analysed. This includes:

- comparing the skills tested within each assessment component with the skills specified within the GCSE controlled assessment regulations

4 Mark schemes, internally assessed tasks, question papers and supporting materials

- comparing the structure of the assessment components
- comparing the effectiveness of the marking criteria

The Form B is completed by each reviewer based upon the information they have extracted to the Form As.

The completed Form Bs will provide each individual reviewer's comparison judgements (with supporting evidence) on the following:

- which skills are assessed by each assessment component
- the structure of the assessment components
- the way in which the marking criteria recognises and rewards skills

This provides the reviewers with a sound understanding of the operation of the various assessment components prior to the student work review. The expert judgement and supporting evidence also begin to address research questions 1. a., b. and c.; 2.; 3.b. and 4. The information defines the opportunities provided to students to demonstrate their skills and will be collated by the lead reviewer for use in support of the findings from the student work review.

Work allocation

Design and Technology: Food Technology; Geography and Music – all reviewers process both specifications in full.

History – these are more complex specifications with significant options. The options were reviewed and the analysis of the student work found that all available work within each year group contains the same option choices for the externally assessed component. The focus of this research is the comparison between features of the controlled assessment and the other assessment component types (coursework and externally assessed). The review will only focus on the components the students have taken.

Therefore the history reviewers will concentrate on the following sections of the specification and assessment components for the home based analysis:

- 2009: Paper 1 Option A: Core Content with Germany, 1919-1945

  Paper 2

  Coursework - all options
- 2011: Unit 971 and Paper 1: Core Content and Germany, 1918-1945

  Unit 972 and Paper 2: How was British society changed, 1890-1918

  Unit 973 and Controlled Assessment: all options

Including all of the History Paper 1 and Paper 2 options from both specifications would only provide evidence to support a comparison between the external assessment options, but that is outside the scope of this research. In addition we do not have available the student work necessary to provide evidence in support of any findings from the analysis of these options.

3 b) Student work review event

The panel of 7 reviewers for each subject will consider student work during a number of individual and group sessions over two days. Reviewing the student work will provide evidence of the student performance in response to the different assessment components in 2009 and 2011 versions of the GCSE.

Each reviewer will consider the subject specific and generic skills demonstrated in each piece of student work. They will take into account the assessment set, how the assessment performed, and whether the evidence produced demonstrated subject related skills and appropriate generic skillsat the appropriate level.

This will include two separate activities:

- Activity 1: the comparison of controlled assessment with both coursework and external assessment
- Activity 2: analysis of whole candidate sets of work for 2011 cohorts

Materials available

The table below lists the student work that will be reviewed for this project.

	A*	A	B	C	D	E	F
Design & Tech: Food Technology 2009	3	3	3	6	6	5	3
Design & Tech: Food Technology 2011	3	3	3	3	3	3	3
Geography 2009	3	3	3	6	6	6	3
Geography 2011	3	2	3	6	6	6	3
History 2009	3	3	3	3	3	3	3
History 2011	1	1	3	3	3	3	3
Music 2009	3	3	3	3	3	3	3
Music 2011	2	1	4	4	5	3	3

Comparing controlled assessment with both coursework and external assessment

Student work will be separated into controlled assessment, coursework and external assessment items. These will be used to provide sets of materials for use in a comparative exercise. The basis for comparison will be the level of subject specific skills demonstrated by students. The exercise will compare:

- controlled assessment work with coursework, and
- controlled assessment work with externally assessed work

Each set of materials (controlled assessment, coursework and external assessment) will be used to make up packs of student work. Although the number of pieces of student work per pack and the number of packs needs to be determined, the following criteria are considered good practice within Ofqual:

- Each piece of student work should appear in at least two packs
- Repeating the same combinations of student work in different packs should be minimised, if possible avoided
-
- Reviewer allocations should be planned to ensure each script combination is considered by at least two different reviewers
- Reviewers will be provided with a ranking form on which to record the contents of a pack from the best to the worst demonstration of subject specific skills
- Reviewers will be encouraged to use the space on the form to comment upon their ranking decisions in terms of the subject specific and generic skills demonstrated by the students

The Lead reviewers can perform a basic analysis of these forms in preparation for their reports. The comments in particular will provide evidence to be considered in support of the findings of the home based analysis.

Ofqual has access to a software package called FACETS to analyse the results from the ranking forms produced during the script review. FACETS uses a Rasch model (often classified under item response theory) to convert the qualitative ranking decisions made by reviewers into a single list which reflects the probable overall order of the sets of student work. For this research the model could be used to order student work from the best to the worst demonstration of subject specific skills.

This analysis will then demonstrate the relative levels of skills demonstrated by students in response to the different assessment forms.

Analysis of whole student sets of work

Each reviewer will be allocated a number of whole candidate sets of work (2011 only) to analyse. The analysis will follow a pro-forma adapted from the Form B in consultation with the Lead Reviewer for each subject. This will ensure each reviewer will complete their individual analyses within a set framework.

There is also the opportunity to adapt the pro-forma prior to the exercise to reflect the interim findings from the home based analysis.

Given the number of reviewers available and the time constraints of the student work review event it will not be possible to complete this analysis for all sets of student work. Therefore, only the sets of student work from the 2011 cohort will be subjected to analysis.

Each set of student work will be analysed by at least three different reviewers.

This detailed analysis will be collated by the lead reviewers and will provide significant evidence to inform the findings of the research in all questions.

3 c) Reviewer meeting

During the student work review event there will be a meeting/meetings for each subject to discuss:

- interim findings from the home based analysis
- initial views of the impact controlled assessment has on student assessment
- ways in which controlled assessment can be improved

There will be a separate meeting for the four lead reviewers to discuss the structure for their subject reports and the expectations of their role during the weekend and in completing their summary report.

4. Summary report

Each lead reviewer will work with Ofqual to generate a findings report from both the home-based analysis work, the script review and the reviewer meeting.

A report structure will be provided to assist them in collating outputs from all parts of the review into a cohesive document which will inform and provide evidence in answer to research questions 1. a., b. and c.; 2.; 3.b. And 4.

The evidence will be generated as follows:

Research question	Assessment Analysis Form A	Assessment Comparison Form B	Comparison of assessment components	Analysis of student work Pro-forma	Reviewer meetings
1. a.	x	x	X	X	
1. b.	x	x	X	X	x
1. c.	x	x	X	X	
2		X			X
3. b.		x		X	X
4		x	x	x	X

The lead reviewer reports will be combined into a summary report to include recommendations for further research activities which might be considered for controlled assessment.

Appendix C: GCSE specifications reviewed

Subject	Year of assessment	Awarding Organisation	Specification Code
Design and Technology: Food Technology	2009	AQA	3542
	2011	AQA	4547
Geography	2009	Edexcel	1312
	2011	Edexcel	2GA01
History	2009	OCR	1937
	2011	OCR	J417
Music	2009	Edexcel	7010
	2011	Edexcel	2MU01

Appendix D: Student work reviewed

Subject		Design and Technology: Food Technology		Geography		History		Music	
Type of material	**Grade**	**2009**	**2011**	**2009**	**2011**	**2009**	**2011**	**2009**	**2011**
Controlled Assessment	A*		3		2		1		2
	A		3		2		1		1
	B		3		3		3		2
	C		3		6		3		3
	D		4		6		3		4
	E		2		5		3		2
	F		2		3		3		3
Coursework	A*	1		3		3		2	
	A	5		3		3		1	
	B	3		3		3		3	
	C	6		5		3		1	
	D	6		6		3		2	
	E	3		5		3		2	
	F	3		2		3		1	
Written Papers	A*	4	3	3	2	3	1	2	2
	A	2	3	3	2	3	1	1	1
	B	3	3	3	3	3	3	3	2
	C	4	3	5	6	3	3	1	3
	D	8	3	6	6	3	3	2	4
	E	3	3	5	5	3	3	2	1
	F	3	2	2	3	3	3	1	3

Appendix E: Availability of specification materials

Subject	Design and Technology: Food Technology		Geography		History		Music	
Materials	**2009**	**2011**	**2009**	**2011**	**2009**	**2011**	**2009**	**2011**
Specification	x	x	x	x	x	x	x	x
Question Paper and Mark Scheme	x	x	x	x	x	x	x	x
Coursework requirements and marking criteria	x	x	x	x	x	x	x	x
Controlled assessment tasks and marking criteria	x	x	x	x	x	x	x	x
Report from the examiners/moderators	x	x	x	x	x	x	x	x

Appendix F: Specification and assessment materials analysis forms

Appendix G: Student work comparison and analysis forms

Appendix H: Numbers of data pairs statistically analysed in the script review

Subject	Number of data pairs analysed	Number of blank lines	Number of missing/null observations	Number of valid responses used
Design and Technology: Food Technology	3584	0	0	3584
Geography	5040	0	0	5040
History	5822	0	0	5822
Music	4088	0	0	4088

Appendix I: Measure, standard error and infit values of the ranked scripts

The ‘measure’ value represents quality of subject related skills demonstrated by the student as judged by the reviewers. It is an estimate of where each script would be ranked if all the scripts were put in order from highest to lowest quality of subject related skills in a single list. Positive values represent the scripts in the top half of all those reviewed.

The SE is the standard error of the estimated measure value. It is likely to be an underestimate as the analysis changed the rankings (as completed by reviewers on the data-entry sheet for each session) into paired comparisons. The table below illustrates this. There are four rank positions. Each rank position is compared against every other position and not just in the order in which they appear.

Reviewer: number 1	
Ranking position	**Script number**
1	65
2	23
3	48
4	52

Paired comparisons made					
65, 23	23, 65				
65, 48	48, 65	23, 48	48, 23		
65, 52	52, 65	23, 52	52, 23	48, 52	52, 48

Each of the ranked scripts will be paired with each of the other ranked scripts twice for comparison. So, for example, rank 1 will be compared with rank 2 and rank 2 will be compared with rank 1 (hence the paired comparison).

The Infit Z value provides an indication of fit. The higher values indicate that there is more disagreement about the ranking of scripts. For example, scripts that were sometimes ranked above otherwise highly ranked scripts but at other times ranked below lowly ranked scripts (therefore, not consistently positioned within the rankings).

The separation reliability value (infit mean squared) provided is an estimate of the proportion of variance in the script measures attributable to ‘true’ variance as opposed to ‘error’ variance. This is likely to be overestimated, as the analysis changed the rankings into paired comparisons. The separation value, therefore, indicates how spread the group of measures of the scripts is. The higher the separation value, the better, as this indicates more confidence in the degree of separation between the scripts (that is to say that there is more certainty in the discrimination between them, as observed by the reviewers during the ranking exercise). So the order of the scripts (in terms of the quality of student performance) is more reliable for the sample of scripts reviewed. Note that the infit mean squared columns’ information will always be a positive number (as it has been squared).

The scripts are listed by student demonstration of subject related skills, with the highest first.

Material types: CA = Controlled Assessment, CW = Coursework, WP = Written Paper

Design and Technology: Food Technology									
Measure	**SE**	**Material type**	**Infit mean squared**	**Infit Zstd**	**Measure**	**SE**	**Material type**	**Infit mean squared**	**Infit Zstd**
6.71	1.84	CA	Maximum		-0.21	0.52	WP	1.22	0.7
6.14	1.86	CA	Maximum		-0.30	0.53	WP	1.06	0.3
5.05	0.83	CA	0.48	-1.0	-0.37	0.55	WP	0.89	-0.2
4.93	0.70	CW	1.32	0.8	-0.41	0.52	WP	1.30	0.9
4.48	1.07	CW	0.83	0.0	-0.44	0.55	WP	0.46	-1.8
4.47	0.68	CA	1.24	0.6	-0.86	0.48	WP	1.24	0.8
4.04	0.72	CA	1.23	0.7	-1.04	0.49	CA	1.14	0.5
4.03	0.72	CW	1.07	0.3	-1.18	0.57	WP	1.05	0.2
3.93	0.68	CW	0.50	-1.2	-1.38	0.52	WP	0.72	-0.9
3.75	0.65	CW	1.25	0.7	-1.39	0.53	WP	0.63	-1.3
3.67	0.83	CW	0.7	-0.3	-1.42	0.53	WP	2.17	2.6
3.57	0.93	CW	1.22	0.7	-1.43	0.43	WP	0.78	-1.1
3.57	0.66	CW	1.26	0.6	-1.47	0.59	CW	1.67	1.7
3.42	0.60	CW	0.85	-0.4	-1.48	0.56	WP	0.65	-1.3
3.14	0.50	CA	0.98	0.0	-1.49	0.67	CA	1.06	0.2
3.10	0.60	CA	0.7	-0.9	-1.53	0.46	CA	1.39	1.5
3.09	0.69	CW	1.10	0.4	-1.69	0.45	WP	1.37	1.2
3.01	0.58	CW	0.7	-0.8	-1.71	0.72	WP	0.77	-0.4
2.99	0.57	CA	0.85	-0.5	-1.74	0.46	WP	0.71	-1.2
2.76	0.48	CW	1.28	1.0	-1.76	0.42	WP	0.66	-1.7
2.61	0.55	CW	0.86	-0.4	-1.87	0.52	CW	1.12	0.5
2.57	0.48	CA	1.31	1.0	-1.98	0.58	WP	1.51	1.6

2.55	0.47	CA	0.59	-1.6
2.35	0.75	CW	0.63	-0.9
2.08	0.45	CA	1.06	0.3
2.06	0.51	CA	0.65	-1.2
2.05	0.54	CW	0.75	-0.7
1.93	0.58	CW	0.91	-0.1
1.74	0.55	WP	1.10	0.4
1.50	0.64	CW	1.06	0.2
1.36	0.48	CA	0.77	-0.7
1.34	0.50	CA	1.16	0.5
1.34	0.55	WP	1.15	0.5
1.27	0.58	CW	0.70	-1.0
1.06	0.46	WP	0.98	0.0
0.99	0.45	WP	1.29	1.1
0.83	0.42	WP	0.70	-1.2
0.80	0.51	CW	0.99	0.0
0.70	0.44	WP	1.01	0.1
0.39	0.51	CA	0.63	-1.3
0.32	0.44	CA	1.88	3.0
0.30	0.68	WP	0.82	-0.3
0.17	0.41	WP	0.56	-2.2
0.04	0.56	WP	0.67	-0.8
-0.07	0.45	CW	1.01	0.1
-0.11	0.65	WP	0.86	-0.2
-0.16	0.52	WP	0.84	-0.4

-2.12	0.49	WP	0.80	-0.7
-2.17	0.54	CW	0.87	-0.3
-2.26	0.57	WP	0.89	-0.2
-2.36	0.57	WP	0.72	-0.9
-2.42	0.66	WP	0.46	-1.7
-2.45	0.64	WP	0.78	-0.5
-2.47	0.68	WP	0.82	-0.4
-2.52	0.43	CW	0.97	0.0
-2.73	0.79	WP	1.17	0.4
-2.81	0.57	WP	0.66	-1.1
-2.94	0.68	CW	1.60	1.2
-2.94	0.54	WP	1.03	0.1
-3.03	0.70	WP	1.17	0.5
-3.18	0.65	WP	0.68	-1.0
-3.21	0.51	WP	0.77	-0.8
-3.25	0.63	WP	0.63	-1.1
-3.26	0.58	WP	1.00	0.1
-3.36	0.48	WP	1.11	0.4
-3.57	0.53	CW	1.38	1.3
-4.14	0.81	WP	0.80	-0.5
-4.26	0.84	WP	1.07	0.3
-4.31	0.70	WP	1.42	1.0
-4.59	1.06	WP	0.87	0.1
-4.91	1.06	WP	0.89	0.1

Geography									
Measure	**SE**	**Material type**	**Infit mean squared**	**Infit Zstd**	**Measure**	**SE**	**Material type**	**Infit mean squared**	**Infit Zstd**
5.05	1.84	CW	Maximum		-0.24	0.42	CA	0.98	0.0
5.01	1.85	CW	Maximum		-0.31	0.41	WP	1.06	0.3
8.03	1.22	CW	0.99	0.1	-0.38	0.42	WP	0.85	-0.6
7.34	1.21	CW	0.97	0.0	-0.41	0.40	CA	1.05	0.2
5.15	1.04	CW	0.96	0.2	-0.43	0.39	WP	1.38	1.7
4.71	1.05	CA	0.89	0.1	-0.43	0.42	WP	1.04	0.2
4.08	1.04	CW	0.89	0.1	-0.46	0.36	CA	1.02	0.2
3.97	0.84	CA	1.17	0.4	-0.49	0.37	WP	0.85	-1.0
3.70	0.63	CA	0.93	0.0	-0.49	0.31	WP	0.95	-0.3
3.29	0.62	CW	0.87	-0.2	-0.50	0.38	WP	0.93	-0.3
2.90	0.62	CA	1.06	0.2	-0.51	0.37	WP	1.03	0.2
2.85	0.57	CW	1.19	0.5	-0.63	0.34	CA	1.07	0.5
2.22	0.47	CW	1.26	0.9	-0.66	0.47	WP	1.49	1.5
2.13	0.46	CW	0.92	-0.2	-0.70	0.36	WP	0.91	-0.5
2.17	0.52	CA	1.15	0.5	-0.70	0.37	WP	0.79	-1.1
1.96	0.40	CW	0.73	-1.0	-0.73	0.39	WP	0.77	-1.1
1.87	0.52	CW	1.13	0.5	-0.80	0.28	WP	0.86	-1.3
1.75	0.41	CA	0.90	-0.3	-.090	0.37	WP	1.03	0.2
1.59	0.55	WP	0.99	0.0	-0.97	0.41	CW	1.23	1.1
1.54	0.41	CA	0.88	-0.4	-0.97	0.37	CW	1.20	1.1
1.44	0.49	CA	1.00	0.1	-0.99	0.40	WP	0.68	-1.8
1.43	0.37	CW	1.01	0.0	-1.06	0.36	CA	0.69	-2.0

1.31	0.46	CW	0.74	-0.9
1.27	0.47	CA	1.27	1.0
2.26	0.36	CW	1.11	0.7
1.05	0.47	CW	0.80	-0.7
1.03	0.43	CW	1.13	0.6
1.03	0.46	CW	1.62	2.4
1.00	0.38	WP	0.82	-0.8
0.99	0.44	CA	0.97	0.0
0.82	0.39	CA	1.22	1.1
0.82	0.43	CW	0.82	-0.6
0.80	0.47	CW	0.89	-0.3
0.77	0.44	CA	0.58	-1.7
0.75	0.46	CA	1.40	1.2
0.67	0.34	WP	1.10	0.8
0.62	0.35	CA	0.88	-0.8
0.60	0.36	WP	1.23	1.4
0.53	0.36	CW	1.10	0.0
0.44	0.34	CW	0.97	-0.1
0.31	0.40	CW	1.20	1.0
0.31	0.39	CA	1.04	0.2
0.29	0.38	CA	0.88	-0.6
0.16	0.36	WP	1.18	1.0
0.06	0.35	CW	1.18	1.2
0.04	0.40	CA	0.99	0.0
0.04	0.37	WP	1.20	1.0

-1.11	0.41	WP	0.72	-1.3
-1.22	0.36	WP	0.97	-0.1
-1.29	0.42	CA	1.21	0.9
-1.36	0.40	WP	0.97	-0.1
-1.41	0.39	WP	1.03	0.2
-1.43	0.42	WP	1.09	0.4
-1.49	0.40	WP	0.88	-0.6
-1.53	0.36	WP	0.85	-0.8
-1.54	0.41	WP	1.10	0.5
-1.56	0.43	WP	0.86	-0.5
-1.57	0.43	WP	1.14	0.7
-1.85	0.41	WP	1.08	0.3
-1.94	0.44	WP	0.90	-0.3
-2.03	0.45	WP	0.68	-1.4
-2.04	0.48	WP	0.79	-0.8
-2.04	0.47	WP	0.96	0.0
-2.06	0.46	WP	1.01	0.1
-2.54	0.54	WP	1.06	0.3
-2.54	0.55	WP	1.07	0.3
-2.56	0.47	WP	0.83	-0.6
-2.57	0.56	WP	1.36	0.8
-2.59	0.50	WP	0.99	0.0
-2.74	0.47	WP	1.06	0.2
-2.74	0.50	WP	0.87	-0.3
-2.91	0.60	WP	0.63	-1.0

0.00	0.38	CA	0.86	-0.8
-0.07	0.41	CW	0.83	-0.9
-0.09	0.32	CA	1.43	2.3
-0.10	0.42	CA	0.83	-1.0
-0.12	0.41	WP	1.34	1.8
-0.16	0.37	WP	1.20	1.2
-0.20	0.39	CA	1.13	0.8

-2.92	0.48	WP	0.95	-0.1
-2.97	0.54	WP	1.22	0.7
-3.22	0.65	WP	1.11	0.3
-3.73	0.66	WP	1.27	0.7
-5.01	1.03	WP	0.96	0.2
-5.14	1.84	WP	Minimum	
-5.80	1.85	WP	Minimum	

History									
Measure	**SE**	**Material type**	**Infit mean squared**	**Infit Zstd**	**Measure**	**SE**	**Material type**	**Infit mean squared**	**Infit Zstd**
5.20	1.86	WP	Maximum		-0.41	0.30	CA	0.84	-1.0
9.30	1.22	WP	0.99	0.1	-0.45	0.32	CW	1.17	1.2
8.61	1.20	CW	0.97	0.0	-0.45	0.39	CW	1.02	0.1
4.28	0.53	CW	0.97	0.0	-0.49	0.33	CW	1.32	1.7
3.99	0.63	CW	0.95	0.0	-0.66	0.42	WP	0.92	-0.3
2.80	0.40	CA	1.31	1.1	-0.72	0.35	CA	1.29	1.5
2.75	0.36	WP	0.90	-0.4	-0.89	0.44	WP	0.84	-0.6
2.68	0.43	CW	0.76	-0.9	-0.93	0.30	WP	1.03	0.2
2.45	0.37	WP	0.94	-0.2	-1.01	0.38	CW	0.83	-1.0
2.41	0.33	WP	1.02	0.1	-1.02	0.38	WP	0.96	-0.2
2.33	0.33	CA	1.15	0.8	-1.05	0.31	CA	0.94	-0.3
1.97	0.42	WP	0.79	-0.7	-1.10	0.39	WP	0.92	-0.5
1.69	0.54	CW	0.64	-1.0	-1.10	0.33	CW	1.00	0.0
1.61	0.40	CW	1.21	0.8	-1.32	0.39	WP	1.39	1.6
1.54	0.48	WP	0.87	-0.4	-1.34	0.37	CA	0.92	-0.3
1.47	0.34	CW	1.10	0.5	-1.35	0.36	WP	0.80	-1.0
1.30	0.34	CW	0.90	-0.5	-1.38	0.37	CA	0.84	-0.8
1.14	0.36	CA	1.30	1.5	-1.38	0.37	WP	0.73	-1.6
1.09	0.37	WP	1.22	1.0	-1.53	0.46	WP	0.93	-0.2
1.04	0.40	WP	1.01	0.1	-1.55	0.36	CA	0.91	-0.5
1.03	0.36	CA	0.90	-0.5	-1.70	0.41	CW	1.01	0.1
0.99	0.39	CA	1.29	1.2	-1.72	0.36	CW	1.44	2.5

0.92	0.39	CW	0.89	-0.5
0.77	0.39	WP	0.85	-0.6
0.71	0.34	WP	0.95	-0.2
0.53	0.45	WP	1.15	0.6
0.24	0.34	CW	0.96	-0.1
0.19	0.47	WP	0.80	-0.7
0.17	0.35	WP	0.90	-0.4
0.10	0.33	WP	1.34	1.7
0.01	0.41	WP	1.01	0.1
-0.01	0.36	CW	0.89	-0.5
-.10	0.36	CW	1.03	0.2
-0.17	0.40	CW	0.95	-0.1
-0.28	0.32	CA	1.11	0.7
-0.37	0.34	CA	0.94	-0.3

-1.73	0.29	WP	0.97	-0.2
-1.79	0.36	CA	0.96	-0.1
-2.22	0.46	WP	0.91	-0.3
-2.28	0.35	WP	1.06	0.4
-2.31	0.49	WP	1.07	0.3
-2.39	0.45	WP	0.95	-0.1
-2.44	0.35	CA	1.22	1.3
-2.45	0.30	CA	1.12	0.8
-2.61	0.44	WP	0.99	0.0
-2.85	0.44	WP	0.95	0.0
-2.92	0.48	WP	0.72	-1.0
-3.06	0.47	CA	0.85	-0.4
-3.17	0.47	CW	0.91	-0.3
-3.41	0.60	WP	1.14	0.5

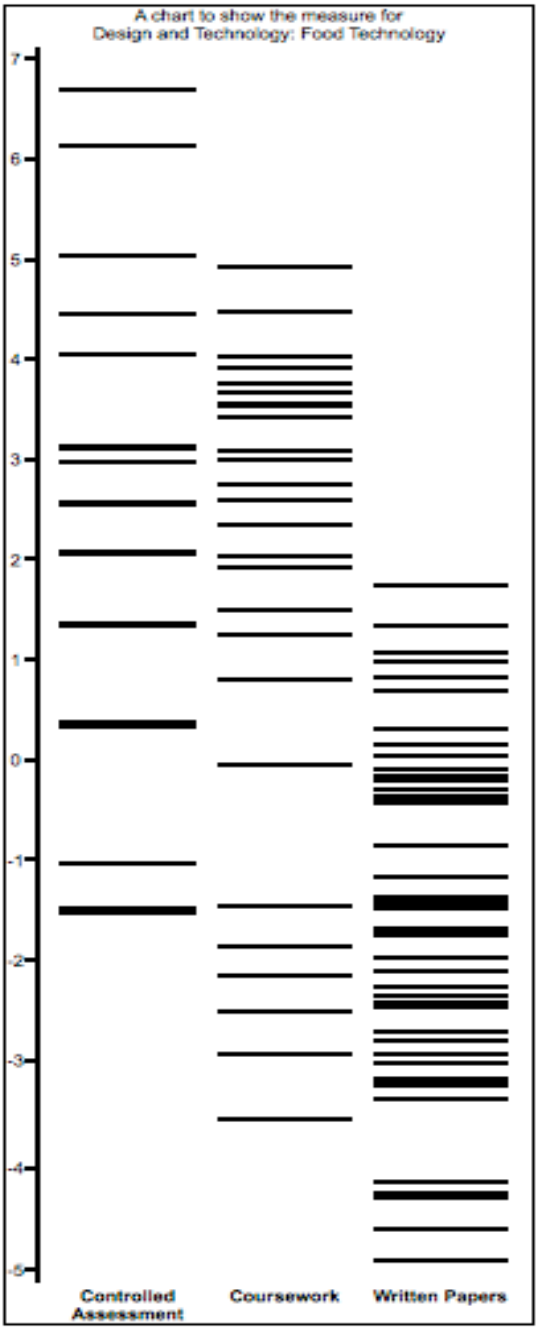
Music									
Measure	**SE**	**Material type**	**Infit mean squared**	**Infit Zstd**	**Measure**	**SE**	**Material type**	**Infit mean squared**	**Infit Zstd**
4.26	1.02	CA	0.97	0.2	-0.04	0.34	CA	1.24	1.2
4.11	0.54	CW	1.06	0.2	-0.19	0.26	CW	0.94	-0.4
3.11	0.45	WP	1.10	0.5	-0.30	0.33	CA	0.89	-0.7
3.03	0.41	CW	0.88	-0.3	-0.33	0.31	CA	1.12	0.8
2.84	0.43	CA	1.00	0.0	-0.59	0.27	CW	1.18	1.2
2.64	0.36	CW	1.02	0.1	-0.68	0.34	CA	1.56	2.8
2.28	0.37	CA	0.75	-1.1	-0.74	0.29	WP	0.87	-0.8
2.25	0.41	WP	1.15	0.6	-0.89	0.35	WP	0.68	-2.1
1.94	0.39	CA	0.88	-0.5	-0.98	0.32	WP	0.88	-0.7
1.68	0.31	WP	1.22	1.1	-1.21	0.29	CW	0.93	-0.4
1.55	0.37	CA	0.81	-0.9	-1.23	0.34	CA	1.45	2.2
1.32	0.39	CW	0.95	-0.1	-1.31	0.29	WP	0.80	-1.4
1.30	0.29	WP	0.95	-0.2	-1.39	0.27	CA	1.17	1.3
1.14	0.28	WP	1.34	2.0	-1.42	0.34	WP	0.90	-0.6
1.09	0.27	CW	1.24	1.8	-1.48	0.31	CA	1.19	1.2
1.03	0.26	WP	0.81	-1.7	-1.53	0.39	CA	1.25	1.2
0.94	0.31	CA	0.95	-0.2	-1.54	0.34	WP	0.64	-2.2
0.72	0.27	CW	0.90	-0.6	-1.67	0.36	WP	0.89	-0.6
0.64	0.35	WP	1.06	0.3	-1.72	0.36	WP	0.86	-0.6
0.57	0.29	CW	0.86	-1.0	-2.00	0.40	WP	0.98	0.0
0.31	0.31	WP	0.73	-1.5	-2.20	0.27	WP	1.24	1.6
0.30	0.29	CW	0.85	-0.9	-2.33	0.38	CW	1.04	0.2

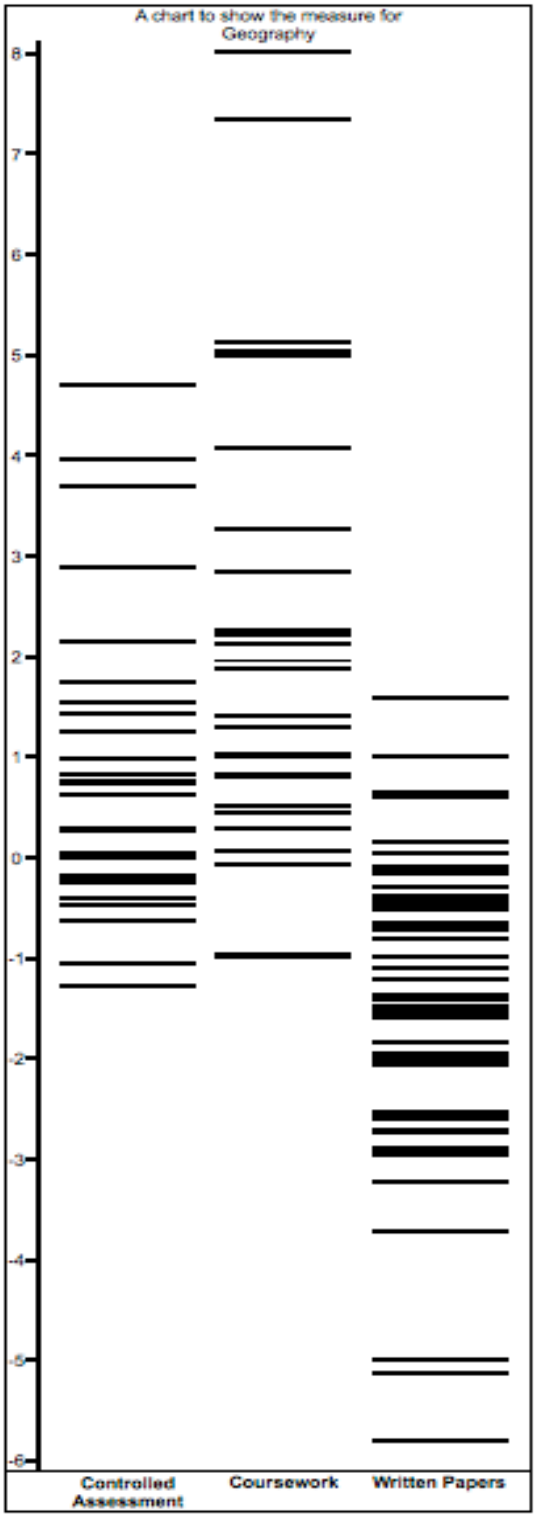
0.27	0.37	WP	0.99	0.0
0.25	0.32	CA	0.83	-1.0
0.23	0.29	WP	1.02	0.2
0.06	0.35	WP	0.58	-2.4
0.05	0.28	WP	0.9	-0.6
0.01	0.30	CA	1.12	0.8

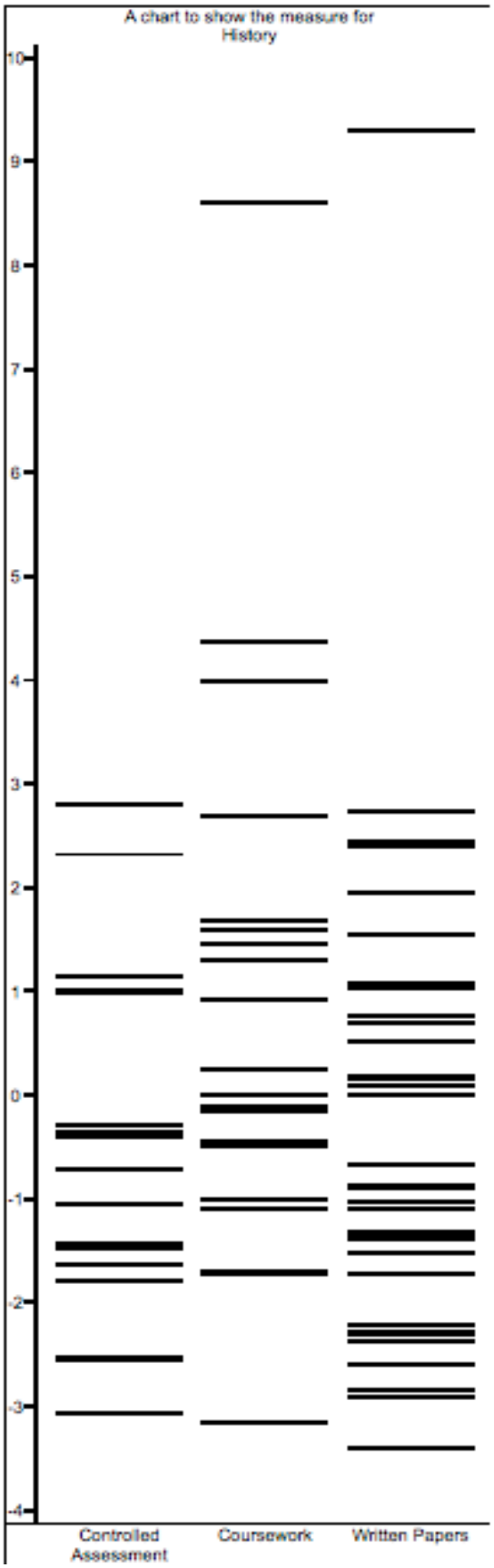
-2.40	0.32	WP	0.84	-0.9
-2.44	0.35	WP	0.75	-1.2
-2.52	0.40	WP	0.87	-0.4
-3.17	0.54	WP	1.00	1.0
-3.62	0.54	WP	1.18	0.5
-4.65	1.84	WP	Minimum	

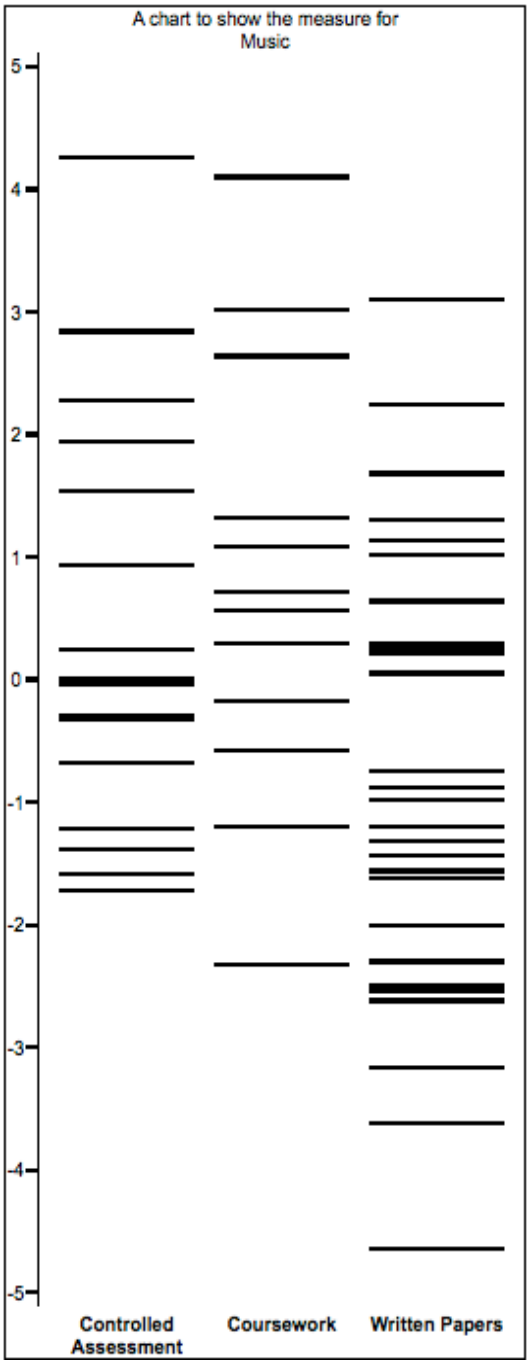
Appendix J: Charts to show the measure

The chart showing the measure for each subject shows the spread of the student work as produced by the FACETS software. Each bold line indicates the measure related to the relevant ranked script. The difference between sequential measures demonstrates the strength of the difference in the ranking position. Large differences would illustrate that scripts were less close in terms of similarity of the student's demonstration of subject related skills than small differences. So there could be a larger difference in the demonstration of subject related skills between scripts ranked 1 and 2 than those ranked 2 and 3 (the difference in the demonstration of skills is not necessarily the same between ranked positions).






A chart to show the measure for History
10
9
8
7
6
5
4
3
2
1
0
-1
-2
-3
-4
Controlled Assessment
Coursework
Written Papers


A chart to show the measure for Music
5
4
3
2
1
0
-1
-2
-3
-4
-5
Controlled Assessment
Coursework
Written Papers

Appendix K: Review panel members

Design and Technology: Food Technology

Valerie Fehners	Lead Reviewer
Susan Blanch	Reviewer
Rachael Croft	Reviewer
Jeremy Curtis	Reviewer
Paul McAndrew	Reviewer
Carol Slinger	Reviewer

Geography

Jan Bond	Lead Reviewer
Rachel Atkins	Reviewer
David Croot	Reviewer
Sue Driver	Reviewer
Tom Miller	Reviewer
Derek Trueman	Reviewer
John Vernon	Reviewer

History

David Martin	Lead Reviewer
Andrew Ashwin	Reviewer
Barbara Hibbert	Reviewer
Julia Mole	Reviewer
Vanessa Musgrove	Reviewer
Amanda Sexton	Reviewer
Alex Woollard	Reviewer

Music

Andy Plank	Lead Reviewer
Marjorie Ayling	Reviewer
Sue Cottrell	Reviewer
Rebecca Lawton	Reviewer
Steve Lewis	Reviewer
Sarah McClure	Reviewer
Sarah Nicholson	Reviewer