TWENTY FIRST CENTURY SCIENCE PILOT

EVALUATION REPORT

February 2007

CONTENTS

INTRODUCTION

Jenifer Burden, Peter Campbell, Andrew Hunt and Robin Millar
Twenty First Century Science project team
(University of York and Nuffield Curriculum Centre)

Background

The central aim of the *Twenty First Century Science* project pilot study was to evaluate a more flexible model of the science curriculum at Key Stage 4 (students aged 15-16) and the courses that comprise this. The steps leading to the *Twenty First Century Science* pilot are summarised in Table 1.

Table 1 *Twenty First Century Science* timetable

Date	Agency	Action
1998	Seminar series on the future of the school science curriculum, funded by Nuffield Foundation, led by Robin Millar and Jonathan Osborne	*Beyond 2000* report published
2000	Qualifications and Curriculum Authority (QCA) working group Keeping National Curriculum Science in step with the changing world of the Twenty First Century	Three studies commissioned to explore aspects of Key Stage 4 science: the meaning of the term scientific literacy, the possibilities for a more flexible curriculum, assessment of ideas and evidence (How Science Works)
2001	University of York Science Education Group (UYSEG) in collaboration with the Nuffield Curriculum Centre and ASE	Report to QCA on a more flexible model for Key Stage 4 based on *Beyond 2000* principles with Core Science and alternative versions of Additional Science.
2002	UYSEG in collaboration with the Nuffield Curriculum Centre	Second report to QCA with draft specifications for the proposed new courses
2002	UYSEG and the Nuffield Curriculum Centre working together as the *Twenty First Century Science* project	Apply for and gain grants from the Nuffield Foundation, Salters' Institute and Wellcome Trust for a development project.
2002	QCA	Commissions the Awarding Body OCR to set up and run a pilot of the new courses.
2003	OCR and the *Twenty First Century Science* project	Rework the course specifications for the pilot. Invite schools to apply to join the pilot. 78 schools apply.
2003	OCR and the *Twenty First Century Science* project	Run a three-year pilot in 78 schools.

The curriculum model and courses piloted

The curriculum model that underpins the *Twenty First Century Science* project divides the 20% of curriculum time allocated to science at Key Stage 4 in most schools into two equal components. One is a single GCSE course for all students, designed specifically to enhance their 'scientific literacy', that is, to provide the kinds of knowledge and understanding that might be useful to anyone in interpreting and reaching an informed view on matters concerning science and technology that they might encounter in everyday life. In the pilot this course was called Core Science. Alongside this, two Additional Science GCSE courses were offered, one with an emphasis on scientific models and concepts and one with an emphasis on the application of science in real workplace settings. In the pilot these were called Additional Science (General) and Additional Science (Applied). The relationship of the pilot courses to the revised courses, available to all schools from September 2006, is summarised in Table 2.

Table 2 The *Twenty First Century Science* courses

Pilot courses	Published GCSE courses from September 2006	Characteristics
Core science (the only course to be evaluated)	GCSE Science	Science in context, to develop all students' scientific literacy
Additional Science (General)	GCSE Additional Science	Concept-led science, as a sound basis for more advanced study of science
Additional Science (Applied)	GCSE Additional Applied Science	Work-related science in vocational contexts, providing a foundation for post-16 applied science courses
	GCSE Biology, Chemistry and Physics courses	Specialist, separate science courses
	Entry level Science	Science for those working at GCSE grade G or below

Evaluation studies commissioned

A proportion of the funding obtained by the *Twenty First Century Science* project team from the Nuffield Foundation, the Salters' Institute and the Wellcome Trust (around 10% of the total) was allocated to evaluation of the pilot as it proceeded. The aim was to gain information and insights that might improve the courses and the support materials, and which might be useful to others considering similar initiatives, particularly in relation to the growing aspiration, in many countries, to focus school science education more explicitly on the goal of improving students' 'scientific literacy'.

Although the sum that was earmarked for evaluation of the pilot was a significant percentage of the total project budget, it was relatively small in absolute terms – barely above the ESRC threshold for a 'small grant'.

The *Twenty First Century Science* project team, following a discussion with staff of the Nuffield Foundation, decided:

• to focus the evaluation on the Core Science course, on the grounds that teaching for 'scientific literacy' is currently of wide international interest, that the Core course was the element of the initiative that all students would experience, and that it was more significantly different from previous science courses than some other elements of the project;

• to invite several researchers to carry out specific pieces of work, rather than inviting one research group to carry out a more general overall evaluation study;

• to invite an experienced science educator to act as 'co-ordinator' of the evaluation studies, to reduce the risks of fragmentation inherent in involving several groups in the work, to help identify links between the studies as the progressed, and to write an overall report on the evaluation studies at the end.

The specific aspects of science teaching and learning on which it was decided to focus were as follows.

1 The understanding which students acquire of the two 'pillars' of the Core Science course: *Science Explanations* and *Ideas about science*.

2 The attitudes towards science and school science of students following the *Twenty First Century Science* course.

3 The classroom teaching of the Core Science course in *Twenty First Century Science*, in particular the aspects of this which might be considered novel by many science teachers.

More specifically, the researchers who undertook Study 1 were asked to evaluate the understanding of some chosen *Science Explanations* and *Ideas about science* before and after the course, and to compare this with similar students following other GCSE Science courses. Similarly, the Study 2 research team was asked to use an attitude assessment instrument to assess any changes in the attitudes towards science and/or school science of pilot school students, and to compare this with similar students following other GCSE Science courses. The researchers who carried out Study 3 were asked to report on how pilot school teachers handled the more novel requirements of Core Science, in particular the teaching of *Ideas about Science* and the management of classroom discussions of science-related issues which may involve a range of social, economic, political and ethical ideas, as well as scientific ones.

From the perspective of the *Twenty First Century Science* project team, these studies were primarily formative evaluations, collecting data that would augment that being collected by the project team and would feed directly into the process of revision and improvement of the teaching materials and teacher support package.

All three groups used the first year of the pilot to develop and test the instruments and procedures they wished to use for data collection, and collected their data on the second cohort of students to take the *Twenty First Century Science* courses, in school years 2004-6.

Twenty First Century Science schools and students

The schools who responded to the invitation from QCA to take part in this pilot were widely distributed across England. The school sample was not in any way selected or constructed, but was not in any obvious way atypical of maintained schools in England. It contained a mix of urban, suburban and rural schools, of differing sizes and student age range (11-16 and 11-18).

In many pilot schools, only part of the year group followed the *Twenty First Century Science* courses, the remainder following the GCSE Science course that the school had previously offered. As a result the sample of students following *Twenty First Century Science* may not accurately represent the student population in the pilot schools. Data collected by the project team at the end of Year 2 of the pilot (July 2005) (table 3) suggested that the student sample following *Twenty First Century Science* may be skewed towards the lower end of the attainment range.

Table 3 Characteristics of students studying *Twenty First Century Science*

Students studying *Twenty First Century Science*		Number of schools
Whole of year group		17
Part of year group:	Lower band classes only	14
	Middle band, or full ability range, classes	11
	All classes except top band	4
	Higher band classes only	5

The nature of the student sample needs, therefore, to be borne in mind in interpreting the findings of all three projects, particularly those parts of the projects which involve comparisons between students involved in the pilot and students following other GCSE Science courses.

STUDY 1 KNOWLEDGE AND UNDERSTANDING

Students' understanding of *Science Explanations* and *Ideas about Science*

Phil Scott, Jaume Ametller, Katie Hall, John Leach, Jenny Lewis and Jim Ryder (Centre for Studies in Science and Mathematics Education, University of Leeds)

EXECUTIVE SUMMARY

Purpose of the study

To provide evidence of the nature and extent of students' understanding of the two main strands of the Core Science course in the *Twenty First Century Science* pilot: *Science Explanations* and *Ideas about Science*.

Research questions

How does students' understanding of major *Science Explanations* and of important *Ideas about Science* change as a result of following the Core Science course in the *Twenty First Century Science* pilot?

How does the understanding of students involved in the *Twenty First Century Science* pilot compare with that of comparable students following other GCSE Science courses?

Methods used

Diagnostic questions on several *Science Explanations* and *Ideas about Science* taught in the *Twenty First Century Science* Core Science course were developed, and pre-tested on a sample of students. The *Science Explanations* selected were the gene theory of inheritance, chemical change and radiation; the *Ideas about Science* chosen were: data and its limitations, correlation and cause, scientific explanations, and risk.

A diagnostic test consisting of a sub-set of these questions was used to collect data from a sample of students following the Core Science course in the *Twenty First Century Science* pilot, and from a comparable sample of students following other GCSE Science courses, at the start Year 10. A similar test, which included some of the questions used in the previous one, was used to collect data from the same students, towards the end of Year 11.

Samples

A stratified random sample of *Twenty First Century Science* pilot schools was first identified. The 14 schools selected were then asked to administer the diagnostic test to all Year 10 students taking the *Twenty First Century Science* Core Science course. A sample of 14 non-*Twenty First Century Science* schools with matching characteristics was then identified. Each of these comparison schools was asked to administer the test to one class in each of the ability bands represented in the corresponding *Twenty First Century Science* school.

In the Baseline Survey, 14/14 of the *Twenty First Century Science* schools (1069 students) and 13/14 of the comparison schools (471 students) returned completed questionnaires. In the Final survey 10/14 of the *Twenty First Century Science* schools (440 students) and 6/13 of the comparison schools (262 students) returned completed questionnaires. Matched samples of students who participated in both Baseline and Final Surveys were then generated, consisting of 336 *Twenty First Century Science* students and 199 comparison school students.

Findings

Science Explanations

Demonstrating an understanding of basic scientific ideas/processes

Overall, the findings showed *Twenty First Century Science* students making progress in understanding over the course.

By the end of the course, their performance was, in broad terms, sound for the physics topic selected, somewhat patchy in the chemistry topic selected, and poor in the biology topic selected. So, for example, the students knew about the greenhouse effect and radiation; they were successful in identifying chemical changes in some contexts but not others; they were generally poor in demonstrating basic knowledge of genetics (which has been shown in other research studies to be a difficult topic for many students of this age).

Overall, the performance of the *Twenty First Century Science* students in demonstrating an understanding of basic scientific ideas/processes was not significantly different from the comparison group.

Applying scientific ideas to develop an explanation

Although they also demonstrated progress in most contexts, the *Twenty First Century Science* students did not perform particularly strongly in questions that required them to develop explanations in familiar and novel contexts. For example, about a quarter were able to offer a basic explanation for a TV remote control working by way of reflection from a wall; fewer than 10% were able to explain why the mass of the combustion products of a fuel is greater than the mass of the fuel itself; about 20% recognised that gene modification is not the same as selective breeding but most were unable to explain this.

Overall, however, the performance of the *Twenty First Century Science* students in applying scientific ideas to develop explanations was not significantly different from the comparison group.

Ideas about Science

Overall, the findings showed that *Twenty First Century Science* students made progress in understanding in most contexts (though not on the topic of risk).

The performance of the *Twenty First Century Science* students in responding to questions about *Ideas in Science* was significantly better than the comparison group on questions about data and its limitations; on the other topics tested, there was no statistically significant difference between the *Twenty First Century Science* sample and the comparison sample.

STUDY 2 ATTITUDES TO SCIENCE

Students' attitudes to science and school science

Judith Bennett and Sylvia Hogarth
(Department of Educational Studies, University of York)

EXECUTIVE SUMMARY

Purpose of the study

To gather data on attitudes to science and to school science held by students involved in the *Twenty First Century Science* pilot study.

Research questions

Are there differences in students' attitudes to science before and after taking the *Core Science* course in the *Twenty First Century Science* pilot study?

Are there differences in students' attitudes to science and school science between those taking the *Core Science* course in the *Twenty First Century Science* pilot study and those taking other GCSE Science courses?

Methods used

An existing instrument, the *ATS*[3] (*Attitudes to School Science and Science)* was used to gather data on attitudes to science and school science. This presents 25 statements covering aspects of science in school and outside school. Students are required to 'Agree', 'Disagree' or 'Neither agree nor disagree' with each statement and either choose reasons for their view from pre-set options or give their own reason.

The instrument was used twice with *Twenty First Century Science* students, once at the start of the *Core Science* course and again towards the end. Comparisons were also made with findings from an earlier study, of students' responses to the same instrument at the end of conventional GCSE science courses.

Samples

Five *Twenty First Century Science* schools in the North of England were chosen, reflecting the variation of school types in the whole set of pilot schools (location; catchment; size; gender mix). Schools were asked to administer the ATS[3] attitude assessment instrument to Y10 classes starting *Twenty First Century Science*, and again to the same students towards the end of their study of the Core Science course. Data from 197 students who completed the instrument on two occasions were analysed. This sample was quite strongly skewed towards students judged by their teachers to be of below average attainment in science.

The group used for post-course comparison consisted of 118 students from three schools. This sample was, however, significantly skewed towards students judged by their teachers to be of above average attainment in science.

Main findings

1 Attitudes to school science

The data show relatively few statistically significant differences in response between students before and after their experiences of *Twenty First Century Science*, or between students following *Twenty First Century Science* and students following more conventional GCSE Science courses.

For the *Twenty First Century Science* sample as a whole, three statements relating to attitudes to school science produced statistically significant differences:

Statement *A04: The things we do in science make me more interested in science*. More students agreed with this statement before taking *Twenty First Century Science* than after. Looking at the reasons given, fewer expressed interest in activities where they could put forward their point of view or which involved working in groups on activities such as posters or presentations. Fewer students were also interested in finding out how living things work and in doing experiments.

Statement *A05: When I have a choice after GCSE, I will choose at least one science subject*. The number of students agreeing with this statement before following the *Twenty First Century Science* course was significantly greater than after the course.

Statement *A08: Everybody should study science up to age 16*. After the *Twenty First Century Science* course, 70.2% of the student sample selected 'Agree' responses, compared with 55.4% beforehand, a statistically significant increase.

Twenty First Century Science students and students on more conventional courses gave broadly similar responses to the majority of statements relating to school science. The one highly significant difference that did emerge was in responses to statement A06: *I enjoy reading science textbooks*. Here 'Agree' options were selected by nearly twice as many *Twenty First Century Science* students (47.4%) students as students taking more conventional courses (23.7%). Fostering interest, developing understanding and helping with revision were the most frequently-cited reasons.

2 Attitudes to science outside school

The data showed no significant differences in responses to school science between students before and after their experiences of *Twenty First Century Science*. What is apparent in the response pattern is that the less personal the statement, the more positive the response. Thus, for example, more than three times as many students thought it was important for this country to have wellqualified scientists than thought it would be good for them personally to have a job as a scientist.

Though there was little difference between responses from students before and after their experiences of *Twenty First Century Science*, a number of significant differences emerged between students following *Twenty First Century Science* and students following more conventional courses. *Twenty First Century Science* students were significantly more likely than students following conventional courses to find news items on science interesting (statement B03), to report liking reading a science books other than school science text books (statement B04), to believe that the Government should spend more money on scientific research (statement B12) and to feel that it is important to promote this country as a scientific nation (statement B14). Fewer *Twenty First Century Science* students indicated agreement with the statement 'I would trust something a scientist said' (statement B05).

3 Gender differences

The data revealed some gender differences in the responses of students following *Twenty First Century Science*. In particular, significantly more girls than boys reported they enjoyed reading science textbooks (statement A06), and reading about science in newspapers and magazines (statement B02). Girls were also significantly more likely to agree that everybody should study science up to age 16 (statement A08). None of these gender differences was found in the comparison data from students undertaking conventional courses.

Study 3 Changes in classroom practice

The teaching of Twenty First Century Science GCSE, and teachers' and students' views of the course

Mary Ratcliffe, Pam Hanley (School of Education, University of Southampton) and Jonathan Osborne (Department of Education & Professional Studies, King's College London)

Executive Summary

Purpose of the study

To report on the teaching of the Core Science course in the *Twenty First Century Science pilot*, and on teachers' and pupils' views of this course.

Research questions

To what extent do teachers recognise the two central elements of the *Twenty First Century Science* Core Science course, *Science explanations* and *Ideas about science*, and how is the emphasis on these, and their interrelationship, interpreted in the teaching of the course?

How successful are teachers in handling science-related issues which involve a range of social, economic, political and ethical ideas, and in using these to develop their students' understanding of *Ideas about science*?

What are students' views of the Core Science course, in particular the emphasis on *Ideas about science* and the inclusion of open-ended issues? How do they see this impacting on their learning and motivation, and on the teaching they have experienced?

Methods used

A questionnaire was administered to a sample of teachers in all *Twenty First Century Science* pilot schools (achieving 121 questionnaires from 84% of schools). This asked about teachers' classroom practices and views of the Core Science course. A sample of lessons in a range of schools in the South of England was also observed. This included observations of several teachers on more than one occasion over the two years of the pilot. These were supplemented by up to three interviews with the teacher concerned, and a focus group discussion with a group of students from their class.

To collect further data on the third research question above, a questionnaire was sent to all *Twenty First Century Science* pilot schools with a request to administer this to a sample of 10 students, of a range of abilities, and where possible, a similar sample of 10 students following a 'traditional' science GCSE. The extent to which these samples are well-matched depends therefore on the judgment of the teachers in each school, and may be influenced by the criteria on which students were allocated to (or chose) the *Twenty First Century Science* course or another GCSE Science course. Also schools offering another Science GCSE alongside *Twenty First Century Science* may not be entirely representative of pilot schools as a whole. The questionnaire asked students about the lesson activities they had experienced and their views of the Core Science course.

Samples

Completed questionnaires were returned by 121 teachers in 68 schools (84% of pilot schools). The original intention had been to observe one teacher in each of 8 schools on five occasions, but staff changes made this impossible. A full set of observations, interviews and focus group discussion were completed in 5 schools. In total, 28 lessons were observed in nine schools. Twenty-two teacher interviews and eight student focus group discussions (in six schools) were conducted.

A total of 381 completed questionnaires were collected from *Twenty First Century Science* students in 39 (or 48% of) pilot schools, and a further 225 from students following a GCSE Science course other than *Twenty First Century Science* in 26 pilot schools which also offered another GCSE science course.

Main findings

Aims of the course and their realisation

The evidence collected suggests that many teachers do not recognise the two elements of the course, *Science explanations* and *Ideas about science*, as having equal emphasis in teaching and assessment, as implied in the specification. Instead *Science explanations* and *Ideas about science* are often seen as interrelated, rather than distinct aspects of the course. This is not unexpected as it is almost impossible to teach an *Idea about science* other than in the context of a *Science explanation*. This intertwining might not matter if both components were given equal prominence in teaching. However, teachers' perceptions and actions indicate that the more familiar *Science explanations* frequently dominate over *Ideas about science*.

In their questionnaire responses, teachers declared confidence both in their understanding of *Ideas about science* and of how they should be taught. Evidence from observations and interviews, however, suggested that many teachers' understanding is superficial. A lack of familiarity with *Ideas about science* may limit teachers' ability to prioritise these in their lesson planning and lead to *Ideas about science* being addressed implicitly rather than explicitly. As a result, students may not appreciate that understanding of a specific *Idea about science* is a core aim of a lesson. Classroom observations suggested that very explicit treatment of *Ideas about science* was unusual. More positively, however, there was some evidence that, after one complete cycle of teaching the course, teachers were developing a clearer and fuller understanding of the objectives of the course and of how they might be realised.

Pedagogic strategies

The evidence collected suggests that teachers are developing and extending their range of pedagogic strategies through the experience of teaching *Twenty First Century Science*. Many reported more use being made of activities that rely on student contributions, e.g. voicing and sharing ideas through discussion. The use of more interactive teaching methods was, however, still low compared with the 'knowledge transmission' approach that characterises much science teaching – and may be lower than is necessary for the aims of the course to be realised.

Teachers sympathetic to the aims of the Core Science course often perceived a need to adapt their practice, but do not find it easy or quick to do so. For example, much of the classroom discourse observed was based on closed questions which offered little opportunity for extended teacher-student dialogue, nor did it encourage student-student dialogue in small groups. Nonetheless, students *were* encouraged to reason and think critically, using topical contexts. While their active participation was encouraged, this, however, was mainly in the form of individual contributions to structured written tasks and whole class question and answer sessions. Few collaborative activities were observed – a situation which was regretted by many students. There were, however, some positive indications that teachers may be slowly modifying their teaching strategies to align more closely with the intentions of the Core Science course, particularly when they had opportunities for collegial professional development. Interview and questionnaire data indicated that teachers believed they learnt most about teaching the course from sharing their views and experiences with other colleagues piloting the course.

Value of the resources

Twenty First Century Science students viewed their textbook as more interesting than their counterparts following other GCSE Science courses. A range of views were expressed in student questionnaires and focus groups about the merits of the textbook for supporting learning, with some students seeing it as an important resource. A consistent complaint from students was that there were too many worksheets. A course which emphasises 'scientific literacy' inevitably makes demands on students' ability to read and interpret information. Perhaps more effort is needed to explain to students why the skills required to analyse and interpret texts of different types (explanatory, expository, narrative, etc.) are of value to them.

Teachers were very appreciative of the high quality ICT resources provided, particularly where these were also easy to introduce into lessons (such as PowerPoint presentations and videoclips). There were indications in the data that use of these ICT resources helped teachers to make the connections between school science and contemporary science. However, there is also a danger that such resources unwittingly reinforce or support an approach to pedagogy which sees learning as a process of information transmission. In the longer term, a more interactive pedagogy might be encouraged and supported by ICT resources which focus on particular *Ideas about science* and require students to collaborate to reach the required endpoint.

Student and teacher satisfaction

The data indicate that these teachers considered the *Twenty First Century Science* Core course more interesting to teach than other GCSE science courses they had taught. In one sense this is not surprising as the schools were all volunteers. It should, however, be set against the fact that teaching any new course is very demanding of teachers' time, and requires a lot of reading, planning and trialling of new materials, often with a less than entirely clear sense of the overall goals or of how each lesson contributes to these. The fact that many teachers, despite these demands, found it more interesting than other science GCSE courses is significant.

A higher proportion of *Twenty First Century Science* students rated their GCSE science course 'topical' and 'easy' than students taking the standard GCSE. The differences were statistically significant. More *Twenty First Century Science* students rated their GCSE science course 'enjoyable' and 'relevant' than students following other GCSEs, though the differences here were not statistically significant. A marginally higher proportion of the students taking *Twenty First Century Science* claimed they would study science post-16 than those taking other GCSE Science courses – though this difference was not statistically significant. The student comments in focus group discussions suggested that the course was achieving one of its primary goals, of relating school science more clearly to students' everyday life. Many remarked that the relevance and value of school science were now more evident. This contrasts markedly with many comments reported in previous research on students' views of school science.

Twenty First Century Science Pilot

Evaluation Co-ordinator's final report

Jim Donnelly
Centre for Studies in Science and Mathematics Education, University of Leeds

Contents

INTRODUCTION

The *Twenty First Century Science* (C21) pilot is one of the most ambitious and comprehensive of post-War science curriculum reforms. It is also one of the most extensively evaluated. There have been seven studies of the pilot project, or aspects of it, excluding examiners' and moderators' reports. The sequence began with two evaluations undertaken by QCA within one year of the start of the pilot project. There are also three studies commissioned by the project itself (on behalf of the project sponsors, the Nuffield Foundation, the Wellcome Trust and the Salters' Institute) which focused on specific aspects of the pilot, and a technical evaluation of the examination, undertaken by the awarding body (OCR). Finally, there is the present report, which the project also commissioned. It should be recalled that all of these evaluations were concerned with the pilot stage of the C21 project, and that they were therefore intended principally to have a formative role in the further development of the project.

The overarching aim of the pilot project was to create a set of courses and specifications which would cater for students with a range of interests and intentions, and thus provide a more differentiated science curriculum. These courses were: a single-award GCSE, followed by all students and usually called the 'Core'; and two 'additional' single award GCSEs (Additional General and Additional Applied). (See the table on page 3 for what these courses are now called.) Each of the courses and specifications has distinct learning aims. The commissioned evaluation studies were targeted on the Core course. This emphasis seems realistic and appropriate, but the existence of the Additional courses must be borne in mind, and may at times introduce uncertainties into the interpretation of the data. In particular, where one of the commissioned studies (Study 3) asked teachers and students for their views about the Core it is possible that the responses were influenced, to a degree that cannot be determined, by experiences of other courses within the pilot project.

The foci of the three commissioned studies were:

for Study 1: 'the understanding which students acquire of the two "pillars" of scientific literacy in the Core Science course: *Science Explanations* and Ideas about Science.'

for Study 2: 'the attitudes of students to science and school science, and how these change during Key Stage 4.'

for Study 3: 'how teachers are dealing with the more novel requirements of Core Science, and … teachers' and students' views of the *Twenty First Century Science* Core course.'

(Statements taken from C21 website in 2006, during the pilot.)

My principal role as Evaluation Co-ordinator has been to offer an overview of these three commissioned evaluation studies, and to integrate, so far as possible, the findings of the other studies. The Brief for the Evaluation Co-ordinator role is given in Appendix B. After the initial review of the Briefs given to the other teams, I have undertaken relatively little in the way of co-ordination while the teams have been working. Their work was done under considerable pressure:

each group provided regular reports to an Evaluation Committee, and these reports demonstrated that the studies were properly focused on their stated topics.

In this report I turn to my final task: to provide an overview of the findings of the evaluations, with particular emphasis on the three commissioned studies. In undertaking this task I have also paid attention to the QCA and OCR evaluations, and the documentation surrounding the project: the examination papers which have been used to assess it, the moderators' and examiner's reports produced by the examining body, and the web-based discussion forum maintained by the project for teachers and others. I have had conversations with members of the development team, and with the Principal Moderator for the Core. However, my principal empirical data sources are the findings of the teams which carried out the three commissioned studies and the reports of the other studies of the pilot.

I have divided my report into five sections, with the following headings:

- Aims of the project
- Judgements the evaluations might make, and their inter-relationship
- Findings of the evaluation studies, and related material
- Conclusion
- Discussion

AIMS OF THE PROJECT

I feel that it is necessary to devote some attention to the aims of the project, for three reasons. First, the project should be evaluated principally against the aims which it has set itself, rather than those which have been ascribed to it, notably within the wider policy-making process. Second, in my view the aims themselves are of several kinds: I will particularly distinguish 'direct' and 'indirect' aims. Third, it can be argued that the achievement of some of these aims is dependent on the fulfilment of others. Some account of these issues is therefore needed before the findings of the three commissioned studies are addressed.

By *direct* aims I mean those aims and associated learning outcomes which are stated in materials produced from within the project. The aims of the Core course are centred on the promotion of scientific literacy. This term is used in various senses: the C21 project team's understanding of it is stated on the project website. The examining body specification also contains a formal statement of aims, though this needed to be in the style and format required by the regulatory body (QCA). These two principal sources are reproduced in Appendix C. The development team's interpretation of the term scientific literacy is of course further articulated across a wide range of locations: in the statements of learning outcomes given for *Ideas about Science* (IaS) and *Science Explanations* (SE); in the course materials which were devised for the use of teachers and students; and in the assessment instruments which were developed by the awarding body. It is clear that, in a complex project such as this, several versions of the aims can be found, each conditioned by their particular purposes: there can be no single and definitive source.

There are also *indirect* aims associated with the project: the word 'indirect' needs some explanation. C21 appears to have influenced important aspects of the structure and substantive emphasis of the reform of the revised KS4 Programme of Study and GCSE criteria which was promulgated in 2004. The project is referred to more or less explicitly in some government documents relating to 14-19 reform and in the Science and Innovation Investment Strategy (e.g., HM Treasury *et al*. 2004: 89). In these settings the focus tends to be on a set of aims defined by the priorities of politicians and policymakers, and it is these which I particularly characterize as 'indirect':

- promoting more positive views of the science curriculum among students;
- promoting a recognition among students of the importance of science and scientists in social, political and economic life;
- generating increased uptake of the sciences post-16 and post-18.

Such aims, which are in part attitudinal, are not identified explicitly in the project documentation. This puts it at some risk of being judged on its 'delivery' of outcomes which it was not set up to 'deliver'. The course was not developed so as to make students' judgements of science more positive, but more informed, analytical and systematic, particularly in connection with its 'applications'. The attitudinal impact of such learning is uncertain, but clearly distinct from the learning itself. On the other hand the expectation of a positive impact on students' judgements of the science curriculum is perhaps not unreasonable. We might anticipate that a more up-to-date, lively and relevant curriculum would carry broader appeal. The project team has not sought to evade these indirect aims, but has acknowledged them within the briefs given to evaluation Studies 2 and 3.

Other indirect aims can be identified. It might be argued that a proper emphasis on promoting scientific literacy within the course would require alterations in the teaching methods commonly employed by science teachers. Again, the project development team has faced up to this issue, both by providing extensive training and support for teachers, and by including an emphasis on teachers' knowledge and teaching methods, especially in relation to *Ideas about Science*, within the remit of Study 3.

What I have sought to highlight in this section is the complexity and heterogeneity of the aims of the project, and the way in which the perception of them is influenced by the broader policy agenda. I will argue below that these aims are also interdependent. In my view, these points reinforce the sense in which these evaluations must be treated as formative and part of an ongoing developmental process, rather than leading to strong summative judgements on the project.

JUDGEMENTS THE EVALUATIONS MIGHT MAKE, AND THEIR INTER-RELATIONSHIP

In this section I will address methodological issues within the commissioned studies, and the relation between their target domains. I will not discuss the QCA or other evaluation studies from these perspectives.

Study 1 'Knowledge and understanding' must be at the core of the evaluation of any project which is principally concerned with learning outcomes. The Study examined such outcomes in relation both to *Ideas about Science* and *Science Explanations*, employing a pre- and post-intervention model, with a non-C21 sample as a comparator. Benchmark performance was assessed across a range of domains early in the course, and with a further assessment towards the end of the pilot, using overlapping but not identical instruments.

The impact of issues such as domain sampling (that is to say the extent to which the questions chosen to represent the domain of *Science Explanations* and *Ideas about Science* adequately reflected the universe of questions which might have been employed), pupil motivation to complete the assessment instruments and, especially, comparability across the samples is significant. It would have been unrealistic within the resources available for the team undertaking Study 1 to address the issue of domain sampling on any scale. Thus their instrument displays validity in the sense that the questions used fit the domain, but cannot claim validity in the sense of systematically sampling the domain. There is some evidence from the pattern of national examination results that C21 schools did not enter their highest attaining students for the pilot, so that there may be not only random differences between the samples, but also systematic differences between the populations from which they are drawn. The main effect of these difficulties is to increase uncertainty, and to suggest caution in drawing conclusions from the quantitative student data.

Study 2 'Attitudes to science' was concerned with attitudinal outcomes and it thus addressed particularly some of the indirect aims to which I referred earlier. It employed a pre-existing research instrument, not tailored directly to C21, though with some slight modifications. The instrument included a strong emphasis on students' views on the place of science in society, but also examined their judgements of science in school and their likely future choice of subjects. Study 2 generated pre- and post-intervention datasets for C21 students. It employed a pre-existing dataset for non-C21 students, treating it as if it were comparable with the C21 post-intervention data. The Study had some difficulty in obtaining data from schools, more especially in the post-intervention phase, as a result of resource limitations. In addition, the attainment distribution of the C21 and non-C21 students diverged. The team made some effort to compensate for this difference, but these circumstances again reduce the confidence which can be placed on comparisons between C21 and non-C21 students.

Study 3 'Changes in classroom practice' examined teachers' pedagogy and attitudes to the course, and the attitude of students. Questionnaires were deployed with both teachers and students (the latter including both C21 and non-C21 respondents). Interviews with teachers and focus group discussions with students were undertaken. Finally, Study 3 involved a significant amount of classroom observation, using an instrument derived from earlier research into teaching methods for *Ideas about Science* (Bartholomew *et al*. 2004). The observational element within Study 3 also had a significant longitudinal aspect, though there was some attrition across schools during the period of the study.

It will be apparent that all of the studies experienced methodological difficulties in relation to access, the maintenance of samples and comparability across C21 and non-C21 settings. These comments are not intended as a criticism of the teams. Each undertook a major study under difficult circumstances, working with schools which had already accepted a large commitment, and for whom the involvement of the evaluation projects could only be an added complication. The comments should however lead us to be cautious about drawing strong conclusions from the data.

The three studies were set up independently, and essentially conducted their work independently, with some attention to the sample of schools so as to avoid multiple approaches to the same school. There is some overlap between Study 2 and Study 3 in relation to the attitudes of students to the course. Broadly, Study 2 focused on students' general views of science education while Study 3 focused more narrowly on the course itself. The sharpest point of contact between the two studies is perhaps at question A05 in Study 2 and question 13 in the students' questionnaire in Study 3. The former asked students to indicate their agreement or disagreement with the statement 'When I have a choice after GCSE, I will choose at least one science subject.' The latter asked 'How interested are you in studying science further or getting a job which uses science?', with students choosing a response on a five-point scale from 'very' to 'not at all'. Just under 30% of the C21 students from Study 2 indicated an intention to follow a science course after GCSE, compared with 37% of those in Study 3. Given the difference in the two questions these data might be thought to show a fair degree of comparability.

Despite their different emphases there is a linkage across the foci of the three studies. In particular, it can be argued that those direct and indirect outcomes which are the focus of Studies 1 and 2 are dependent on successful teaching of the course, and perhaps also on the attitude of teachers to the course and its aims. It is often suggested that the teaching of *Ideas about Science*, and the teaching of *Science Explanations* in a context where *Ideas about Science* are prominent, requires a significant shift in teaching methods, away from those judged to be the norm in science teaching. That shift is thought to involve a greater emphasis on classroom interaction and discussion, and the acceptance of learning outcomes which are less precisely defined than those usually sought by science teachers.[1] Study 3 was charged with identifying the extent to which teachers were successful in meeting these challenges when teaching the course, as well as their attitudes to the course and its aims. For these reasons I judge that Study 3 occupies a fundamental position in the evaluation, and that the proper interpretation of the findings from the others is dependent on it.

---

[1] I am setting aside the argument that the teaching approaches identified here are equally necessary if greater success is to be achieved in teaching more traditional science 'content'.

To summarize:

Study 3 might tell us whether the pedagogic and attitudinal conditions, among teachers, for the course to achieve its aims have been met. It might also say something about the personal response of students to the course.

Study 1 might tell us about direct learning outcomes, the most explicit of the project's aims: but its findings are likely to be conditional on the findings of 3.

Study 2 (and perhaps also Study 3) might tell us whether the course has generated any indirect effects on students' attitudes towards science at a broad level, rather than specifically in relation to this course. Whether these changes in attitude are conditional on effective learning in the narrower sense employed in Study 1, or can be achieved independently, remains open to question.

FINDINGS FROM THE EVALUATION STUDIES

QCA evaluations

QCA undertook two evaluations in the first year of the project, based on both questionnaire and observational data in schools, and reported them together (Qualifications and Curriculum Authority 2005). The relevance of these early QCA studies to judgements about the impact and long-term future of the project appears limited.

The pattern of findings which emerged was influenced to a significant extent by the early stage at which the work was undertaken, and there was some emphasis on issues which were contingent on this (such as late arrival of materials and technical difficulties with ICT equipment). Five aspects of the findings which were reported at this early stage seem still to have relevance at the end of the pilot.

• Both teachers and students (as judged by their teachers) found the elements of the pilot as a whole livelier, more relevant to citizens' lives and more interesting than traditional Double Award (QCA: 29, para. 56).

• Most schools found the learning outcomes of the Core more demanding than traditional Double Award (QCA: 18, para. 27-30).

• Many schools commented on the absence of practical work in the Core (QCA: 10, para. 33; 15 para. 10).

• Many schools commented on the absence of differentiation in the pilot as a whole (to the disadvantage of less able students) (QCA: 30, para.56).

• Teachers were judged to recognize a need for adjustments in their teaching methods (QCA: 6; 9, para. 27 et seq.; 22, para. 39).

OCR technical evaluation

The other formal evaluation of the pilot was a statistical analysis undertaken by OCR. This study focused on the relationship between the specifications (and other science specifications) and other subjects, and the relationship with KS3 performance. There was also a study of the relationship between the different elements of the pilot. In general the conclusion of the evaluation was that the C21 specifications were comparable with other science specifications in their relation to other subjects. In addition the study suggested that consistent standards were being applied across the different courses within C21.

The commissioned studies

I turn now to the commissioned studies, and, in the light of the discussion above, I will begin with Study 3**.**

Study 3 Changes in classroom practice

Teachers' views and practices

In discussing this aspect of the Study 3 findings, I will first reflect what I see as the researchers' strong emphasis on the challenges which most teachers experienced in addressing the aims that the Core course embodies. I will however go on to pay attention to those findings which demonstrate that the majority of teachers found this a more rewarding course to teach than 'traditional' Double Award Science.

This team was asked first to examine the recognition and interrelationship of the two components, *Science Explanations* and *Ideas about Science*, in the teaching of the course. They identify an imbalance in the emphasis given to the two within the pilot schools: teachers apparently placed a greater stress on *Science Explanations* than on *Ideas about Science*. This was detectable both within teachers' judgements of the weighting of the assessment instruments and in the teaching that was observed, though they are more circumspect on the latter. They suggest that, with multiple goals, there is a danger that the 'Idea about science' emphasis can be lost as an explicit outcome. The researchers also lay some stress on teachers' tendency to rate the aims of the project (as given in the specification and identified in Appendix CII) as equally important, and interpret this as a lack of discrimination. They also note teachers' apparent unfamiliarity with the overarching aims given in the specification and on the project website. The teachers are judged likely to have focused instead on schemes of work and teaching materials. One might add 'and on assessment instruments as these became available'. There is an important underlying issue here, about the place of teachers' judgements in interpreting and perhaps influencing the future development of the course, to which I will return in my final discussion.

The study identifies ambivalences in teachers' judgements about their own knowledge of *Ideas about Science*. When being interviewed they often saw themselves as lacking in expertise in this area. Elsewhere, and particularly in the questionnaire responses, they were judged to report quite positively on their knowledge. It may be that teachers are referring here to different domains: how to teach *Ideas about Science* (pedagogic content knowledge, as it is often called nowadays) and their own knowledge of *Ideas about Science*, respectively. Alternatively it may be that the context (the former an interactive interview, the latter a more impersonal questionnaire) influenced the view they offered.

The second area on which Study 3 was asked to focus was the teaching of *Ideas about Science* and science-related issues. The researchers developed a perspective on teaching generally, and the effective teaching of *Ideas about Science* in particular, derived from earlier work in this area (Bartholomew *et al*. 2004). It is not necessary to subscribe wholly to this analysis to gain a sense from the report of the demands teachers experienced in seeking to understand and teach *Ideas about Science*. This sense is reflected quantitatively in the finding that 72% of teachers found the course harder to teach than other courses. There are several inter-penetrating issues here; the explicit knowledge discussed above is only one. The researchers explored teachers' confidence, their use of discussion, their willingness to alter well-established practices, and their general view of how learning and teaching is to be understood. In all of these areas, Study 3 paints a picture of teachers who are conscious of the challenges that teaching *Ideas about Science* presents, but who are in broad agreement that teaching activities should encourage active (i.e. physical, verbal and social) participation. This commitment appeared fairly uniform across both teachers and topics and was not simply directed towards the teaching of *Ideas about Science*. There were only occasional signs of differing views, such as the questionnaire finding on teaching methods actually employed, where 'Teacher explaining ideas' headed the list, and the observation of one teacher interviewed that '*I feel I'm experienced at explaining (scientific ideas) and can help (students) get the ideas quicker and more efficiently than by other methods*').

Study 3 emphasizes one quite specific aspect of the teaching methods involved in teaching *Ideas about Science*, the use of pupil discussion, although the list of possible teaching strategies whose frequency of use teachers and students were asked to rate implies that discussion is merely one of a number of means by which critical analysis of issues might be promoted. The project resources and the listing of activities in the specification (section 10.4 onwards) imply something similar.

Study 3 is cautious in its handling of the issue of teacher progression in relation to teaching methods. From its interview data it found some evidence of development, particularly through the experience of teaching the course. However the findings from the observational study did not suggest major changes in practice amongst teachers as a whole.

Teachers reported positively on the extensive training and support which the project provided for them. However they found working with supportive colleagues to be the most useful form of support, while the Internet, perhaps experienced as both more impersonal and more inflexible, was judged least positively. The supportive colleagues involved had access presumably to similar external ideas and support, and this invites questions about the character and impact of this collegial activity. It perhaps suggests forms of collective creativity in meeting these new challenges.

So far, in this account of Study 3, I have focused on what I see as the researchers' central theme: the challenge to teachers, and their response. I have also placed this issue in the foreground because of its significance in interpreting the other evaluation studies. Yet, running in parallel with this theme, the study tells a positive story about teachers' attitudes and experiences. A clear majority of teachers indicated that they enjoyed teaching Core Science more than other courses. Given that nearly three-quarters also found it more demanding to teach, this clearly represents a major achievement. This finding echoes others within the QCA reports (see above), and more informal findings reported from within the project itself (Millar 2006: 1513 et seq.).

I doubt that there is a Hawthorne effect in play here, i.e. that teachers would have responded positively to *any* change. Teachers' professional lives are pressured, and they do not judge change positively unless the benefits to their core classroom activity are very clear. Furthermore, while it seems reasonable to assume that the degree of sympathy for the course's aims would influence teachers' effectiveness, or at least their commitment to developing their teaching so as to address the course's aims, it cannot be assumed that the teachers involved were all volunteers. Data from Study 3 suggest that up to 45% of teachers may have been required to teach the course rather than teaching it from choice. Of course this does not necessarily imply that they were out of sympathy with its aims: but neither does it suggest that the work was done by a group of committed enthusiasts.

Students

Some of the ambivalences between questionnaire and interview findings which the Study 3 team found amongst teachers are echoed amongst pupils. Within both focus group and open-ended responses one can find enthusiasm for the distinctive emphasis on contemporary issues within the course. The researchers sum this up with the statement that, for pupils, 'the relevance of the course is now transparent and self-evident'. This finding does not mean that the course offers a panacea for the challenges which face science in the curriculum. They identify a number of issues of which the course developers need to take account: for example, ensuring that the language used in resources is accessible across the full range of pupils; and responding to students' (and teachers') expectation of experimental work. Some of these issues were identified early in the pilot, through both the QCA evaluations and the development team's own monitoring, and were addressed through additions to, and modifications of, the teaching materials as the pilot proceeded. In the case of experimental work the issue is not a simple one, and I will discuss it further below.

As with the teachers, the quantitative data, which included responses from both C21 and non-C21 students, were not entirely consistent with those from the interviews. In relation to the issues of *Topicality* and *Relevance,* the focus group findings appeared to be largely confirmed, though tests of significance of several observed differences produced ambivalent outcomes. One possible explanation is that aggregate data are concealing differences in response from different sub-groups of pupils. Judgements of *Enjoyment* of their science course, and of the likelihood that it would promote *Further study of science,* are similar for C21 and non-C21 students. In responses to a question on the *Ease* of the course the total number of students rating their course as Easy or Very Easy was significantly higher for C21 than non-C21 students. This finding, as the researchers note, is open to varying interpretations.

There are several more nuanced findings within the detail of Study 3. There was, for example, evidence of differences of interpretation of the role of textbooks between the team developers and some students. C21 students expressed moderately greater enthusiasm for coursework than did non-C21 students. The well-established pattern of relative enthusiasm and antipathy across the different science disciplines appeared to be maintained, to the disadvantage of the physical sciences. Topics such as *Air Quality* were judged among the least enjoyable and the most difficult, whilst *Keeping Healthy* showed the reverse pattern.

In judging all of these student questionnaire findings where a comparison between C21 and non-C21 students is offered, it needs to be recalled that the comparison is based on two different groups. None of the students, so far as is known, had experienced both C21 and non-C21 science courses at KS4. I have noted earlier the critical issue of comparability across both the samples and the populations.

Study 1 Knowledge and understanding

In an earlier section I suggested that the findings from Study 3 provide the conditions against which Studies 1 and 2 need be interpreted. My reading of the Study 3 findings is that we should be cautious in drawing conclusions about the impact C21 *might* have, when properly 'embedded' in teachers' practice, from the impact it appears to have had during this pilot.

If this judgement is a fair one, then it is unsurprising that Study 1 detects no statistically significant differences across nearly all outcomes, in relation both to *Ideas about Science* and *Science Explanations*. In *Ideas about Science* there is a slight difference in the area of 'Data and its implications', with C21 students performing rather better. It is worth noting that this area of the curriculum involves the sorts of difficulties that teachers have traditionally dealt with in practical classes. During experimental work students commonly produce a wide range of 'data' of varying quality, which teachers need to manage in some way. It is possible that teachers might already have a collective understanding of learning aims and classroom practices in this field, which they were able to deploy in a systematic way, under the stimulus of the Core. In *Science Explanations* the fact that the data show little difference in understanding between C21 and non-C21 students can be construed positively, given the extensive attention paid in the Core to *Ideas about Science*. However the very moderate attainment displayed by both groups of students in *Science Explanations*, particularly when asked to use their knowledge in novel contexts, reflects what has been known since the mid-1980s, when the Assessment of Performance Unit deployed instruments across random samples of students on a similar basis (Archenhold *et al.* 1988: chapter 9). The C21 pilot cannot claim to have broken this pattern.

Study 2 Attitudes to science

Study 2 looked for impacts on students' substantive judgements of science and school science, and their declared future intentions. They relate to what I earlier called indirect aims (and outcomes). As with Study 1 it seems reasonable to assume that these impacts will be dependent on the quality of the teaching which students experienced. The findings support some of those associated with Study 3. Students' interest in textbooks and issues relating to science in the news, as well as their judgements of the general social and political significance of science, appear more positive than those to be found amongst non-C21 students. Subject to my earlier comments on the dangers of generalizing from these data, they offer the possibility that, at an affective level at least, the course has had an impact on students, and that that impact has been positive, in the sense that it has made them more ready to acknowledge the importance of science. However, in a specific question, C21 students were found to be significantly less likely to trust what a scientist said. This is perhaps unsurprising in relation to a course which, in my view, gives more attention to provisionality in scientific knowledge and to the social processes by which that knowledge is obtained, than is to be found in traditional Double Award science. The circumstances in which the judgement of the scientist's statement is made are not clarified in the question, and might be thought important in framing the response, as some students indicated in their answers.

Again, there are ambivalences in the findings of this Study. A narrower focus on students' own judgements of schooling and their own declared future intentions reveals a different emphasis from that just outlined. Responses to a sequence of questions suggest that the attitudes of C21 students to science in school did not alter significantly while following the course, and their declared likelihood of pursuing science post-16 declined significantly. Unfortunately there are not comparable longitudinal data for non-C21 students, so it is not possible to say how the decline compared with the latter group. The Study 2 team draw our attention to the similarity between their findings and those in other work, referring particularly to that of Jenkins and Nelson. The latter authors sum up their study of student views about science in the title of their report: for most students science is 'important, but not for me' (Jenkins and Nelson 2005).

Related material from the awarding body

The awarding body had a parallel responsibility (independently of the project team that developed the teaching resources and teacher support package) for developing the assessment material for the project. It also produced examiners' and moderators' reports. These reports are not evaluations in the usual sense of this term. Nevertheless they do give insights into the project and its outcomes. No part of the commissioned studies has examined the institutional relationships between the development team and the assessment team within OCR: I will suggest in the Discussion section below that this is an important aspect of the present and future realization of the project.

Assessment of the Core has successfully broken away from the emphasis on investigations which has dominated GCSE coursework in science for the last decade, and is generally acknowledged not to have achieved its aims. By contrast the C21 Core requires a more focused Data Analysis activity, which matches well on to aspects of *Ideas about Science*, and a Case Study. The former may have contributed to some of the positive findings which Study 1 identified, and has allowed teachers some flexibility in how the dataset to be analysed was created. In sum, it appears well-fitted for its purpose. I have included as Appendix D an excerpt from the 2006 examiner's report on these elements of the assessment process, written by the Principal Moderator. This material seems to me to be an important source of evidence. It comments on the ways in which schools were able to adapt the Data Analysis activity, with specific reference to lower-attaining candidates:

"…in these cases data collection activities involving whole class participation were generally the most successful. These included pollution surveys, fitness studies and habitat surveys. In these activities, the whole class can be involved in the planning stage. Each candidate takes some measurements, so that they are familiar with the practical difficulties involved. The total body of data collected can be very large, so that there is plenty for students to say about accuracy, validity and reliability." *(See Appendix D for source.)*

The introduction of the Case Study is the most radical shift in assessment methods which has occurred in the Core. The Principal Moderator comments that these Studies have 'drawn a most positive and enthusiastic response from candidates of all abilities'. Two key themes run through the full set of comments on the Case Study in Appendix D: the positive engagement which the activity generated, and the importance of the relationship with substantive science. The Principal Moderator's report draws attention to the need for Case Study reports to display a proper engagement with the underlying science in whatever issue is chosen as the focus of the report, and he argues that such engagement is an important discriminator between more and less highly attaining students.

CONCLUSION

For those who have followed the development of C21 from its inception it is not difficult to identify its achievements. The creation and piloting of a multi-faceted and ambitious course is a major achievement in itself. The quality of the pilot materials and the range of support which has been provided to teachers are also noteworthy, as is the liaison which has been achieved across a range of institutions. Making an overall judgement on the achievements of the project would however be difficult since there has not been, so far as I am aware, a systematic evaluation of the Additional General (as it was originally called) and Applied components. Their contribution to the overall outcomes of the pilot is likely to have been significant. There is a good deal of anecdotal evidence to suggest that the latter in particular may have offered a distinctive and appealing science curriculum to students who might have been unimpressed by the established curriculum at Key Stage 4. There is also a question about the extent to which teachers' judgements as documented in Study 3 relate to the Core only, or include (though without making this explicit) views arising from involvement with the Additional courses. Though my subsequent comments here relate to the Core course, and to its distinctive concern with scientific literacy, these points should be borne in mind.

Perhaps the single most positive feature in the findings of these studies is that the project has succeeded in persuading a majority of the pilot teachers that it is more professionally rewarding than standard Double Award science. It has thus, in my view, engaged to a measurable degree with the desire of science teachers to offer a curriculum, particularly at Key Stage 4, which is livelier and more engaging than what has traditionally been available. Moreover it cannot be assumed that the teachers involved are all volunteers, or sympathetic to the aims of the project. This achievement also needs to be set against the well-established view that producing educational change of any kind is difficult: the institutional, professional and political barriers are large. All this is a significant achievement.

Given the balance within participating teachers' views it is disappointing that Studies 1 and 2 have been able to detect limited measurable impact on the judgements and achievements of students who have followed the course. On the other hand Study 1 data suggest that, despite its broadened aims and idiom, the Core has not diminished students' learning in those more traditional aspects of the curriculum (*Science explanations*) to which it has inevitably given less attention. The ethical and professional responsibilities of a venture such as this to the students and teachers who participate in it are critical. The evidence (from the evaluations, the examiners' reports and the views of teachers) is that students have not been disadvantaged by following the pilot. Studies 2 and 3 also indicate some broad positive impact on students' judgements of the importance of science, and, to a lesser extent, of the relevance of the science curriculum.

All three teams have applied statistical tests of significance to their data, in some cases converting ordinal to numerical data for this purpose. Such tests are of course sensitive to the variance within the datasets which are being compared. Given the variation which seems likely to exist across classrooms, and the critical role of teachers in most educational settings, it is important to underline that findings of 'no statistically significant difference', of which there are many in the evaluation studies, refer to the aggregate level. We simply do not know whether differences exist at a finer grained level and might be detectable if it were possible to take systematic account of some of the variations in teacher practice towards which Study 3 points.

I have repeatedly urged caution in drawing strong conclusions from the data reported by Studies 1 and 2, partly for the reasons just given, but also in consequence of the finding of Study 3 that the knowledge and teaching methods which might be thought necessary to teach *Ideas about Science* effectively are not strongly embedded in teachers' understandings and practices. Overall, the studies' empirical findings are probably best treated as having set a benchmark from the pilot project. Their outcomes are on balance neutral: they do not undermine the project's aims, and they demonstrate a substantial measure of teacher enthusiasm and goodwill. They evidently invite, indeed demand, replication and development during the years after the completion of the pilot.

DISCUSSION

In the previous sections of this report I have focused on the findings of the empirical evaluation studies. In this Discussion I will address wider themes and offer a more personal response. The significance of these themes derives from the importance to science education of the demanding and ambitious aims which the project has set itself (particularly in respect of the Core), and towards which the pilot project has made such an energetic and comprehensive start. The themes that I will discuss are:

- the interpretation and realization of the aims of the Core;
- the role of teachers in these processes;
- the place and representation of substantive science.

The notion of scientific literacy, to which the Core is addressed, does not support a single interpretation (Laugksch 2000). Developing a scheme of teaching and assessment to promote scientific literacy has therefore been not merely a major practical task for the project, but also an intellectual and professional interpretive challenge. As recent debates in the news media illustrate, it can also court political controversy.[2] The process has involved a wide division of responsibility, incorporating what might be called 'chains of interpretation' across individuals and institutions. The outcomes have been expressed in various locations – principally through the specification, the teacher and pupil materials, and the assessment instruments, each created under specific constraints, priorities and timescales. Maintaining a clear intellectual unity across this process requires ongoing attention. Several comparisons could be attempted across the network of locations, but I will illustrate my point by reference to the two broad statements of the aims of the Core (reproduced in Appendix C): one from the examining body (CI) and one from the project team (CII). My purpose is not particularly to highlight differences here, though I believe that there is some difference in emphasis across the two texts. CII identifies media reports and personal decision-making as key contexts in which science learning might be used. This emphasis is less visible in CI. Nor do I suggest that these differences are problematic, though they do invite the question of whether the manner in which media reports are read and evaluated, and personal decisions are taken, by citizens draws wholly or mainly on the knowledge cited in CI. I use this comparison simply to stress the importance to the project of continuing to maintain an overview of its aims and their meaning across the several locations I identified above. This process has both institutional and intellectual aspects.

A critical location at which the meaning of scientific literacy is given expression is in the assessment instruments developed in connection with the course. In particular, how is the validity of the instruments to be judged? Traditional formulations of content validity are grounded in expert judgement. In a developing, and occasionally controversial, field such as this the key locations of such expertise are not always easy to identify. It is important that the forms of assessment continue to be scrutinized for consistency, fitness for purpose, and

---

[2] http://news.bbc.co.uk/1/hi/education/6038638.stm. (Accessed 16 November 2006.)

for the meaning they offer for scientific literacy, across the range of perspectives and settings which exist within the project. It is also important positively to seek accommodation across these settings. The Case Study which students undertake as part of the coursework received very positive comments within the Principal Moderator's report from the examining body. This is in my view a very valuable outcome from the pilot, particularly given the experience of GCSE coursework in science in the past. It is an achievement that could easily be overlooked. But it is important to monitor how the realization of the learning outcomes of the project within this assessment mode compare with others in the assessment regime. The Principal Moderator identifies the controversy surrounding the MMR vaccine as a popular theme for students ('Is the MMR jab safe?'). A question in more traditional form from the 2006 examination round (2006/7797/4933/02, q.3) addresses the same theme. It seems to me that it is the Case Study that is more likely to encourage that holistic engagement with the MMR issue to which we might aspire for the scientifically literate citizen. I am of course aware that assessment instruments operate under policy constraints, and that recent policy changes by government seem likely to reduce the flexibility available to awarding bodies.[3] Furthermore, I thoroughly endorse the Principal Moderator's emphasis on engagement with substantive science as an important element in judgements of attainment in the Case Study (PMR 2006: 4, 6), and will return to this point in the final part of this Discussion.

The preceding comments have focused on the project development team and the awarding body, but there is another critical professional grouping in the project: teachers themselves. A leading member of the team has argued that the project as a whole is akin to a hypothesis-testing activity: the hypothesis in question is concerned in part simply with whether it is possible to create a workable curriculum in scientific literacy. He has also stressed that the development process is essential to clarifying the 'meaning and implications for the science curriculum of a scientific literacy emphasis' (Millar 2006: 1500). I agree with this claim, which is similar to Pawson's argument that programme development is akin to theory testing (Pawson 2003). But I particularly wish to comment on how it invokes the views and actions of teachers. Unless teachers are to be understood, in the common English formulation, simply as 'delivering' a curriculum and a pedagogy devised by others, their views and actions, must be understood not merely as relevant, but to a significant degree as *constitutive* of the curricular and pedagogic meaning of scientific literacy.

Study 3 argues that many teachers' engagement with the aims of the project, again both in terms of their interpretation of the specification and their teaching actions in the classroom, is 'superficial' and 'lower than might be necessary to fulfil the aims of the course effectively'.[4] Putting this and the previous points together, it appears that to a significant degree we do not yet know what a curriculum and pedagogy dedicated to scientific literacy will eventually look

---

[3] http://news.bbc.co.uk/1/hi/education/5385556.stm (Accessed 16 November 2006.)
[4] I understand that this type of language, and other examples ('all the research would indicate that teachers of science generally have a somewhat simplistic understanding of the nature of their own subject') might be interpreted by some as within the 'delivery' paradigm of curriculum change. I am choosing not to interpret it in this way.

like. Its meaning will only be realized when the course becomes embedded in the understandings and actions of teachers with reasonable stability. The outcomes of this process may take the ideas of the course developers in unexpected directions. In sum, I suggest that the professional relationship between teachers and the central development team is a critical issue. I am not convinced that this relationship has yet received the attention it deserves: the major effort has been directed towards the rationale of the project and its teaching resources. This is of course understandable. These were priorities for the team, under the pressing need to meet both the requirements of the regulatory body and the practical demands of the project. However, given the importance of the relationship to teachers within previous curriculum development efforts (Donnelly and Jenkins 2001), this is not in my view a trivial or merely 'academic' issue, but central to the future development and likely impact of the Core, and the project as a whole.

The previous paragraph might be rendered a little less abstract by reference to some concrete issues. First, it is possible that some teachers will see the activities associated with *Ideas about Science* as enlivening and enriching the curriculum, but as having only a limited impact on its main learning aims. This perspective has some resonance with the findings of Study 3, at least at the level of teachers' practice. Second, it is not merely possible but almost certain that teachers will bring diverse perspectives to the philosophical and sociological issues relating to science that the course seeks to engage, despite the efforts which have been made to reach a consensus (Osborne *et al*. 2003). These issues are rarely within teachers' own area of specialist knowledge, and their interpretation often remains controversial among philosophers and sociologists of science. Third, and at a yet more concrete level, it is clear that significant numbers of teachers and students took the view that the Core contains less practical work than existing courses, and were critical of this. (I ought to note here that this view is not necessarily shared by members of the development team.) At any rate it is probably fair to say that the place of practical work within the aims and preferred methods of the Core is less central than has been common in the established wisdom of science teaching. There is also evidence, from Study 2, that some students are less enthused than might have been hoped by discussions and group work. Yet there is, as the Study 3 researchers noted, a body of research which suggests that the impact of practical work on learning is moderate, at best. While concerns about the scale of practical work may have been addressed in the revision of the materials at the end of the pilot, the questions about rationale, impact and professional authority over teaching remain. As things stand we have what can appear as an uneasy accommodation between the development team and the teachers taking up the course. I perhaps ought to stress that these comments are meant to be illustrative. They are not intended as a criticism of either teachers or the development team, but to draw attention to the nature of their relationship in the context of this important issue.

My final theme in this Discussion is principally directed towards underlining the central place of scientific knowledge within the Core. In my view the distinctive aspect of the Core, at least in terms of learning outcomes, is its expansion of the treatment of scientific knowledge beyond that in traditional courses, through its handling of *Ideas about science*. Yet I hope that the account of *Ideas about* *science* which is offered within the project materials will be kept under review. I note particularly the theme of *Risk*, as it is taken up in the specification. Analysis of Risk is important in much contemporary controversy. It is also prominent in the account of *Ideas about science* given in the Core specification. But in what sense can Risk properly be called an *Idea about science*, and what might be the consequence of identifying it as such? Much more distinctive of science, in my view, is its materialistic account of the universe, and its capacity to generate agreement and a measure of confidence beyond most, arguably all, other forms of human knowledge creation. I do not see these key characteristics of science reflected strongly, or perhaps at all, in the list of *Ideas about Science*.

The central aim of the Core is to deepen students' understanding of scientific knowledge, so that they can deploy that deepened understanding in the circumstances where they encounter science in their lives. It seeks to enable them to analyse critically the methods by which scientific knowledge is created, its intellectual status and the implications which can properly be drawn from it. The Core also sets out to alter the balance of knowledge which is addressed in science courses, placing greater emphasis on those areas which relate to aspects of modern life. These are important, legitimate and demanding aims. I hope that the approach taken in the Core, and the teaching which emerges from it, will continue to acknowledge them as fundamental. I am pointing towards a tension here, which I hope I do not overstate, between these aims, and those concerned with students' understanding of sociological, political and ethical issues related to science, as well as with themes such as Risk. These latter issues and themes are increasingly incorporated into the science curriculum, and not just within C21. This may be occurring for pragmatic reasons, in the hope that it will improve motivation and interest, or because these issues cannot be separated out when considering scientific questions which impact on modern life. Yet, in my view, their presence within science education ought always to derive from and be subordinate to its foundational purpose: to promote a critical understanding of scientific knowledge proper, and its place in the world.

REFERENCES

Evaluation Studies are referenced as follows:

OCR: A technical evaluation of the OCR GCSE Sciences suite pilot examinations

PMR 2006: *GCSE 21st Century Science 2006 Centre Coursework Report: 4934/1, 4934/2 and 4938/1* (Principal Moderator's Report, 2006)

QCA: Qualifications and Curriculum Authority (2005). *Evaluation and analysis of the 21st Century science pilot GCSEs*. London, QCA

Archenhold, F., J. Bell, J. Donnelly, S. Johnson and G. Welford (1988). *Science at Age 15. A Review of APU Survey Findings 1980-84*. London, HMSO.

Bartholomew, H., J. Osborne and M. Ratcliffe (2004). Teaching students 'ideas-about-science': five dimensions of effective practice. *Science Education* **88**(5): 655-682.

Donnelly, J. and E. W. Jenkins (2001). *Science Education: Policy, Professionalism and Change*. London, Paul Chapman/SAGE.

HM Treasury, DTI and DfES (2004). *Science and Innovation Framework 2004-14*. London, HMSO.

Jenkins, E. W. and N. W. Nelson (2005). Important but not for me: students' attitudes towards secondary school science education in England. *Research in Science and Technological Education* **23**(1): 41-57.

Laugksch, R. C. (2000). Scientific literacy: a conceptual overview. *Science Education* **84**(1): 71-94.

Millar, R. (2006). Twenty First Century Science: Insights from the design and implementation of a scientific literacy approach in school science. *International Journal of Science Education* **28**(13): 1499-1521.

Osborne, J., S. Collins, M. Ratcliffe, R. Millar and R. Duschl (2003). What 'ideas-about-science' should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching* **40**(7): 692-720.

Pawson, R. (2003). Nothing as practical as a good theory. *Evaluation* **9**(4): 471-90.

Qualifications and Curriculum Authority (2005). *Evaluation and analysis of the 21st Century science pilot GCSEs*. London, QCA.

Appendix A Commissioned evaluations relating to the C21 Pilot

The evaluation of *Twenty First Century Science* will focus on three aspects of the development.

1 The understanding which students acquire of the two 'pillars' of the Core Science course: *Science Explanations* and *Ideas about Science*.

2 The views of science and of school science, and attitudes towards science and school science, of students taking *Twenty First Century Science* course(s).

3 The extent to which teachers are successful in handling the more novel requirements of Core Science, in particular the teaching of *Ideas about Science* and the management of classroom discussions of science-related issues which may involve a range of social, economic, political and ethical ideas, as well as technical ones.

(Source: documentation sent to evaluation teams.)

Appendix B Evaluation Co-ordinator's Brief

1 To act as a 'buffer' between the *Twenty First Century Science* development group and those carrying out the evaluation studies, to increase the confidence of external audiences that the evaluation studies are independent of the developers.

2 To bring fresh insights to the evaluation exercise as a whole.

3 To ensure so far as possible that the different strands of the evaluation remain focused on issues of general importance and interest.

4 To help ensure that links, and points of contact, between the separate evaluation projects are identified and exploited as far as is possible.

(Source: documentation sent to evaluation teams.)

Appendix C Aims of the project

I From the Core specification:

The aims of this GCSE specification are to encourage candidates to:

• acquire a systematic body of scientific knowledge, and the skills needed to apply this in new and changing situations;

• acquire an understanding of scientific ideas, of how they develop, of the factors which may affect their power, and of their limitations;

• consider and evaluate critically their own data and conclusions, and those obtained from other sources, using ICT where appropriate;

• evaluate, in terms of their scientific knowledge and understanding and their understanding of the processes of scientific enquiry and of the nature of scientific knowledge, the benefits and drawbacks of scientific and technological developments, including those related to the environment, personal health and quality of life, and considering ethical issues;

• select, organise and present information clearly and logically, using appropriate scientific terms and conventions, and ICT where appropriate;

• interpret and evaluate scientific data and conclusions from a variety of sources;

• demonstrate that they can select, organise and present information relevantly and appropriately to a given brief using appropriate scientific terms and conventions, and using ICT where appropriate;

• use electronic (internet, CD ROMs, databases, simulations etc.) and/or more traditional sources of information (books, magazines, leaflets etc.) to collect data and ideas on a topic of scientific interest.

OCR GCSE Science E: Single Award Pilot (2003) p.10

II The Twenty First Century Science website site identifies the aims of the Core under its account of scientific literacy:

We would expect a scientifically literate person to be able to:

• appreciate and understand the impact of science and technology on everyday life;

• take informed personal decisions about things that involve science, such as health, diet, use of energy resources;

• read and understand the essential points of media reports about matters that involve science;

• reflect critically on the information included in, and (often more important) omitted from, such reports; and

• take part confidently in discussions with others about issues involving science.

http://www.21stcenturyscience.org/rationale/scientific-literacy,903,NA.html
(accessed 19 October 2006)

APPENDIX D
FROM THE PRINCIPAL MODERATOR'S REPORT (2006)

(Excerpt from OCR, Science E (Twenty First Century Science) General Certificate of Secondary Education ***GCSE 7797/7798/7799*** *Reports on the Units June 2006.)*

Data Analysis

The assessment involves the application of the same Strand I and E criteria as for investigations, and similar comments apply as indicated above. Interpretation and evaluation of data are clearly difficult skills for candidates and the assessment criteria are more rigorous than the current Sc1.2. This consequently produced a greater spread of marks reflecting the different abilities of candidates in this area. It is most important that information is provided by the Centre about how the task is presented to candidates. It is also important that candidates record the data they have collected or are going to use and not just plot a graph etc. without any reference to the original data.

Many Centres took marks for this assessment from full investigations. In many ways these candidates appeared to be better placed to make realistic evaluations of their procedures and data collected. However, in the case of weaker candidates, the data collected was often poor in quality and quantity so that they found interpretation difficult. Therefore in these cases data collection activities involving whole class participation were generally the most successful. These included pollution surveys, fitness studies and habitat surveys. In these activities, the whole class can be involved in the planning stage. Each candidate takes some measurements, so that they are familiar with the practical difficulties involved. The total body of data collected can be very large, so that there is plenty for students to say about accuracy, validity and reliability.

Overall, both in these exercises and in full investigations there was often only limited evidence that candidates were applying ideas about the quality and reliability of data. Outliers were generally recognised as being present but not always clearly identified on the graph or in the text of the report.

In strand I, the separate requirements to identify and describe patterns in data and then to link them to a scientific explanation meant that marks were often lower than would have been awarded using the Sc1.2 criteria. In Sc1.2, only the very weakest candidates score less than 4, and almost all get 5 or 6. Here marks of 1, 2 or 3 were not uncommon. A similar, but less marked, effect occurred in strand E. The result of this was that marks were more widely spread than is usual with Sc1.2, giving clearer differentiation, and consequently more secure grading. Grade boundaries were chosen to take account of this difference in marking, so were rather lower than originally anticipated.

Case Studies

General comments

It is clear that Case Studies have been a very successful aspect of coursework assessment. They have drawn a most positive and enthusiastic response from candidates of all abilities. In some Centres all candidates were given the same title whereas in others a broader range of opportunities were given. In general the latter were more successful. However, whatever arrangements were adopted it was clear that students showed a sense of 'ownership' of the study, and even very weak students managed to produce coherent reports. "Copying" in the sense of relevant direct quotations from texts, or cut-and-paste from web sources is allowable. Reports from the weakest candidates often consisted of little more than two or three 'cut-and-paste' sections, with minimal editorial comment from the candidate. Thus candidates in this group had selected relevant material from a source, made some attempt to link the facts together and present a report achieving 5 or 6 marks.

Even in Centres where all candidates worked on a single topic provided for them, there was almost no evidence of collusion or copying between candidates. Centres had taken heed of the advice in the specification to ensure that a substantial proportion of the final writing up took place under supervision.

It would be most helpful for moderation if more annotation or commentary was provided for each candidate in the sample selected so that the moderator could support the Centre marks. In many cases only the final mark awarded was recorded.

Choice of subjects for Case Studies

Case Studies should always be framed in terms of a question to provide a focus in an area of controversy in which either the quality of the scientific evidence or the actions that should be taken in a particular situation are in doubt. For example, 'Is it safe to use your mobile phone?' rather than 'How do mobile phones work?' This will encourage candidates to look for different opinions and views, and to consider the evidence base for claims or the reliability of sources. Studies which were presented as questions to answer were always more effective than those which simply *described* a topic. The more successful candidates also described the relevant science needed to understand their chosen topics, and took care to evaluate sources and compare different views.

Popular topics included:

- Natural disasters e.g. "Is it worth while to plan ways of avoiding an asteroid strike?"; "Should people be allowed to live near volcanoes / in earthquake zones / on river flood plains?"
- Aspects of diet e.g. "Is obesity inherited?" Is the Atkins diet good for you?"
- Food additives – are they good or bad?
- Should GM crops be allowed?
- Should parents be allowed to have designer babies?
- Is cloning ethical?

- Should stem cell research be allowed?
- Are mobile phones bad for your health?
- Should we spend more developing alternative energy resources?
- Is the MMR jab safe?
- Should smoking be banned in public places?
- Is there life on other planets?
- Does motor traffic cause asthma?
- Is hormone replacement therapy good or bad?

Quality A: Selecting information

This Quality was often over-marked. Where possible, students should use multiple and different types of sources of information e.g. web-sites, encyclopaedias, library books, course textbook and their own notes. There should be a list of the sources used and these should be detailed, referring to for example, book and page number and full URL, not just to the homepage of the particular website.

Material from the sources should be selectively used, not just a collection of ‘cut-and-paste’ extracts. This will usually include some direct quotation, but should also involve some re-structuring of information. Where sections of text are directly quoted, this should be made clear e.g. quote-marks or different font from the main body of the text if word processed. Some candidates highlighted in colour quotations from sources – a novel and effective way of showing where information came from.

Most candidates wrote a bibliography of sources at the end of their reports. However, rather fewer candidates used references within the text to show the source of particular information or opinions.

Some candidates wrote a bibliography at the end of the report and then put page references that matched the page number in the Case Study. Some included a useful glossary of scientific terms within the report. Several candidates handed in full print-outs of their sources but added little or no editorial comment of their own and therefore a match to the 4 mark description was not secure. Some candidates gathered information from self-constructed questionnaires which also added to the pool of material for their Case Study, but occasionally this distracted them from the underlying science and scientific evidence.

At the 4 mark level, any obvious disagreement between sources should be recognised and also their reliability should have been considered. One candidate wrote a short sentence at the bottom of the appropriate pages describing the reliability or bias nature of the sources quoted on that page – a simple but very effective way.

Failure to discuss reliability and failure to indicate within the text where particular pieces of information or opinions had come from prevented many candidates from being awarded 4 marks in this quality.

Quality B: Science Understanding

This tended to be the area of weakest performance. This section allows students to show what they understand of the background scientific knowledge and understanding which will help them evaluate the information in their sources. Students should explain the basic scientific facts, principles and concepts of the topic. Their student textbook or own notes will be a good source. This can be backed up by extra information from more advanced or specialist books, encyclopaedias or web-sites.

Candidates also usually failed to consider the scientific content in the sources they quoted. Reporting was often at the 'headline level', simply repeating claims without looking beyond the headline for the experiments or data the claims were based on.

Candidates generally responded with enthusiasm but the studies were often short of information about the basic scientific knowledge and understanding which underpin the topics chosen. Candidates should be urged to describe basic KS3 or KS4 ideas about the topic, as well as any new information they discover, to provide a scientific basis for the opinions they develop in quality C. For example, in the many variations of the popular 'Genetically Modified Food' case study, a brief introduction to what genes are, followed by a clear definition of what is meant by genetic modification would have been helpful. In 'power generation' case studies, there should be explanations of the nuclear fission process and the nature of the nuclear waste and the associated disposal problems. Also many candidates included various schematic and illustrative labelled diagrams but little of the information presented was used and developed in any way.

Quality C: Conclusions

There should be evidence that the sources used have been compared to check for consistency and to identify areas of conflict or disagreement. There should also be evidence that the underlying science has been used to try to resolve any differences.

There should be evidence that points 'for and against' have been compared and that candidates have given their own viewpoint or position in relation to the original question, and have justified this by reference to the sources.

Several candidates scored fewer marks than they were probably capable of, particularly in quality C, because they simply chose to report information about their topic, without any real consideration of the scientific evidence they were based on. Opinions from a variety of sources were often quoted but without reference to the source or to the evidence that the claims were based on. Although most candidates made an effort to give two different views in their studies, these were rarely compared, and conclusions often seemed to lack any clear basis in the evidence shown. This rarely leads to marks above 4 and it was very rare indeed for even the best candidates to attempt any judgement of the quality, validity or reliability of any of the scientific evidence offered by their sources.

Quality D: Presentation
The majority of reports included headings and/or sub-headings to provide the necessary structure but the visual impact was often variable. Suitable diagrams and graphics should be incorporated as appropriate to clarify difficult ideas and encourage effective communication. Too often, pictures were used simply to look pretty – illustrative rather than informative. Rather too little use was made of diagrams, charts, tables or graphs as compact ways of conveying large amounts of information, or to visualise difficult concepts. The best candidates always made good use of explanatory diagrams.

Some candidates included a table of contents which gave structure to the report and also helped to guide readers quickly to particular sections.

Those reports which were presented simply as *PowerPoint* printouts often lacked sufficient detail for high marks in the other qualities. However, those which had notes to accompany each slide were much more successful in matching the higher mark descriptions across the other qualities.

EVALUATION OF THE TWENTY FIRST CENTURY SCIENCE PILOT: A PROJECT RESPONSE

Jenifer Burden, Peter Campbell, Andrew Hunt and Robin Millar
Twenty First Century Science project team
(University of York and Nuffield Curriculum Centre)

What have we learned from these evaluation studies? Given the very modest resources available for each study, we share Donnelly's caution about reading too much into their findings. In particular, we share his view that it would be risky to make strong inferences about the impact and outcomes of the *Twenty First Century Science* approach and courses from September 2006 onwards, as these have been modified considerably in the light of the pilot. The data collected during the pilot, including that from the three commissioned studies, were used to inform changes to the teaching materials and the teacher support package during the pilot itself and, more thoroughly in the revision of the specifications and teaching materials for publication in 2006. So the courses now being taught are not those that were evaluated.

With the benefit of hindsight, perhaps our greatest regret is that resource limitations obliged us to focus so narrowly on one element of the *Twenty First Century Science* pilot – the Core Science course. As a result we have little objective evidence of the outcomes of Core Science plus Additional Science (General) or of Core Science plus Additional Science (Applied). Informal evidence collected by the project suggests that the Additional (Applied) option has been seen in many schools as a distinctively new and attractive option, offering a solution to a long-standing problem of making suitable provision for one particular (and quite large) group of students who can be interested in science, provided it is presented in a way that makes the practical usefulness of the knowledge and skills more apparent. It would be useful now to have more systematically collected evidence of the implementation of this course in schools, and of its outcomes. (See the table on page 3 for what each of these courses is now called.)

Amongst these outcomes is the extent of its impact of students' post-16 science choices – an indirect rather than a direct aim of the project, in Donnelly's terms. Post-16 subject choice is also of considerable interest in the case of those students taking the Core plus Additional (General) option – a combination likely to be taken by many students who would previously have studied Double Award GCSE Science. The fact that in many pilot schools only some of the year group followed *Twenty First Century Science* courses makes to very difficult to interpret the information we have so far collected on post-16 subject choices. This will become considerably easier from September 2006, as more schools will be using *Twenty First Century Science* courses with their entire year cohort. The project intends to collect data on this, though it will be several years before any firm conclusions can be safely drawn.

Where Studies 1, 2 and 3 compare *Twenty First Century Science* and non-*Twenty First Century Science* students, the great majority of the differences observed are not statistically significant. Whilst it would be nice to report statistically significant gains in understanding of important ideas, or attitudes towards science, the absence of measured effects that attain statistical significance is not surprising. We concur with Donnelly's comments on the matching of student samples and on domain sampling in Study 1 particularly. These are, to a large extent, problems that face any quantitative study using an experimental design – and are here exacerbated by the small resources available, the limitations of access, and variations in the way the pilot was implemented. So were we unwise to ask these researchers to carry out studies of such an explicitly comparative kind? We think not, since questions about the learning and attitudinal outcomes are likely to be among the first that many people will ask. The results of Study 1 provide reassuring evidence that, despite its stronger greater emphasis on other aspects of science learning, *Twenty First Century Science* students appear to acquire a level of understanding of scientific ideas that is similar to that of students following other GCSE Science courses. It is also pleasing that the findings of Study 2 suggest that the course has encouraged greater interest in science stories and issues that students hear about through informal channels. In the light of Study 3's findings, we think that learning and attitudinal gains are likely to be greater once teachers have had enough time to gain familiarity and confidence with the course content and the teaching approaches it requires.

On the other hand, we also think that these evaluation studies of the pilot also highlight the need for rigorous research studies of a different kind, if we are to understand the effects of an innovation like *Twenty First Century Science* and learn from it. A set of detailed case studies of schools as they take up the *Twenty First Century Science* courses could provide insights that would significantly augment the kinds of evidence that can be gained from more quantitative studies. The project team has informal anecdotal evidence that the pilot had a very striking impact on teaching and learning in many schools. These reinforce our sense that it is a mistake to attach the label 'works' or 'doesn't work' to a new course or approach; instead we see that some teachers are able to 'make it work', particularly those for whom the emphasis and approach of a course resonates with their own view of the purpose of science education. It would, therefore, be useful to know more of the characteristics of teachers and schools where *Twenty First Century Science* is seen to be working well, and others where this is less clearly the case. We hope that such studies might be carried out in the next few years, though we also believe that they should ideally be done by researchers less closely associated with the development of the project than we are.

Our overall view, therefore, is similar to Donnelly's, that many of the findings of the evaluation studies on the pilot are inconclusive about its effects, and that they 'invite, indeed demand, replication and development during the years after the completion of the pilot'.