Collecting and using student feedback on quality and standards of learning and teaching in HE

A report to HEFCE by the Centre for Higher Education Research and Information (Open University), NOP Research Group and SQW Ltd

Collecting and using student feedback on quality and standards of learning and teaching in higher education

A report to the Higher Education Funding Council for England

John Brennan	CHERI
Robin Brighton	SQW Limited
Nick Moon	NOP Research Group
John Richardson	Institute for Educational Technology, the Open University
Jane Rindl	SQW Limited
Ruth Williams	CHERI

SQW Limited
economic development consultants

Enterprise House
Vision Park
Histon
Cambridge CB4 9ZR
Tel: +44 (0) 1223 209400
Fax: +44 (0) 1223 209401
www.sqw.co.uk

The Open University
Centre for Higher Education Research and Information
344-354 Gray's Inn Road
London WC1X 8BP

Tel: +44 (0) 20 7447 2552
Fax: +44 (0) 20 7837 0290
www.open.ac.uk/cheri

NOP Research Group

245 Blackfriars Road,
London SE1 9UL

Tel: +44 (0) 20 7890 9830
Fax: +44 (0)
www.nopworld.com

TABLE OF CONTENTS

Chapter **Page**

A Invitation to tender
B Questionnaire for HEIs on internal student feedback processes
C Institutions visited
D Aide Memoire for discussions with students
E Proposed questionnaire for the National Survey

Executive Summary

Introduction

1. This report has been prepared for the Higher Education Funding Council for England (HEFCE) by a project team comprising SQW Limited, the Centre for Higher Education Research and Information (CHERI) at the Open University and NOP Research Group. The study had two main components:

   - to identify good practice by higher education institutions (HEIs) in collecting quantitative and qualitative feedback from students and to make recommendations on the design and implementation of mechanisms for use by individual institutions. The focus of this part of the study is quality enhancement

   - to make recommendations on the design and implementation of a national survey of recent graduates, the results of which would be published. This part of the study is focused on providing comparative information to assist applicants to higher education (HE)[1].

2. Fieldwork for the study was undertaken between September and December 2002. Written information on institutional processes was requested from all HEIs in England and visits were made to 20 HEIs. During these visits, discussions were held with staff and current students on both feedback procedures within the institutions and the potential value of a national graduate survey. More focused discussions on the National Survey were held with a further 50 students and a small pilot survey was undertaken over the Christmas period.

Institutional processes for collecting and using student feedback

3. Virtually all higher education institutions (HEIs) possess quite elaborate mechanisms for the collection of student feedback information. While there are considerable variations in detail, all institutions use a range of mechanisms, both quantitative and qualitative. There is considerable variation in the detail of questionnaire design but considerable commonality to the topics covered. A number of institutions have introduced a degree of standardisation to their questionnaires while allowing faculties and departments some discretion to add or

[1] CHERI was primarily responsible for the first component and SQW Limited, with advice from NOP Research Group, for the second. However, the two components are related and staff from CHERI and SQW Limited worked together on both.

indeed modify content. The most common focus of questionnaires is the individual module although a majority of institutions also gather feedback data at other levels.

4. Many institutions have some kind of central unit with responsibility for student feedback although in practice much is often devolved to faculties or departments. Even where there is a central unit, analysis of feedback data is often quite limited, with little use of comparative analysis or relating feedback data to other institutional datasets.

5. In many institutions, more use could be made of feedback data. This would require additional resource but consideration might be given to collecting less data and analysing it more thoroughly. It may not be necessary to administer a module questionnaire every time the module is offered.

6. Greater analysis and more imaginative presentation of feedback data might encourage more use to be made of it, which in itself would increase the commitment of staff and students to the importance of feedback processes. We detected that this commitment was slight in some places.

7. One way of making greater use of the data is to enable students to take it into account in choosing modules.

8. Whether it is being used for internal or external purposes, it is important to place data in context, to present it in the light of other forms of data on quality and standards and with reference to the distinctive features of the institution. If this contextualisation of information is at the expense of some direct comparability, this is preferable to the publication of data that is misleading or meaningless.

9. Although we believe that it would be possible and in principle desirable for institutions to publish summary information, appropriately contextualised, of student feedback information at programme[2] or equivalent level, we do not believe that such a requirement should be made mandatory. There are genuine concerns within institutions about the effects of publication upon the quality enhancement role of feedback and about the resources that would be required to produce good quality publications that would genuinely be useful to prospective students. However, some institutions might well wish to publish if their feedback systems can provide information in a suitable form, if the needs of quality enhancement can be protected, and if the resource can be made available. The purpose of publishing such information would be to increase the prospective student's knowledge of the likely student experience at *that* institution, not to rank order all institutions in the land. Therefore, enunciation of context and use of common themes are more important than direct comparability of numerical data.

10. This study has confirmed that HEIs are devoting considerable resource to obtaining feedback from their students. Some of the good practice discovered by the present study will be

[2] Programme, in this report, refers to the collection of units, modules or courses that lead to an award.

included in the *Good Practice Guide* to be published by HEFCE later in 2003. However, we believe that there remains a need to do more to share experiences and good practice both within and between institutions, especially with regard to the analysis, presentation and uses of student feedback data. In some institutions, consideration may need to be given to whether more data is being collected than is needed and whether greater effort needs to be devoted to use of existing data than to collecting more of it. At the same time, the overall cost and burden of student feedback should be monitored to ensure that the value is commensurate with the costs.

11. Detailed recommendations have been made throughout the report and these are presented at the end of each of the following sections:

- purposes
- mechanisms
- collection
- analysis and interpretation
- actions and decision-making
- presentation and publication
- dissemination to students.

The National Survey

12. Applicants to HE would find a well designed National Survey, which provided information at below the institutional level, useful. There are many aspects of the HE experience which are relevant to applicants when selecting which HEI to study at. Our view is that the National Survey should focus on teaching and learning narrowly defined.

13. There are real methodological issues surrounding student feedback on teaching quality. We recognise the importance of these, but believe that there is sufficient evidence to suggest that the Course Experience Questionnaire (CEQ), currently used in Australia, has addressed these successfully. We recommend that the National Survey should be based closely on the CEQ.

14. We recommend that the National Survey should not collect information on employment related issues. To do so would mean the survey would be administered at least two years after graduation. This would compromise its usefulness as an indicator of comparative teaching

quality and also reduce response rates. If employment related issues are not included in the survey then it could be administered before graduates leave the HEI where they are studying.

15. We recommend that the National Survey should be census-based and a postal survey. The main reason for a census approach is the need to report results by institution and also by subject area. We think it likely that responses in some 'cells' would be unacceptably low in number if the National Survey was sample based.

16. We think there would be merit in all aspects of the survey (mail out, analysis and reporting) being contracted to an independent organisation, but this may not be immediately acceptable to the HEIs because of concerns relating to the Data Protection Act 1998. We therefore recommend that, initially, questionnaires are mailed out by the institutions but data is processed and analysed by a single contractor. The only additional demand on HEIs would be mail outs of the questionnaire and reminders.

17. We recommend that information is reported at the subject level using the Joint Academic Coding System (JACS). If the number of responses is sufficiently large then information should be reported at both the 19 subjects and the next level down. The lower level of reporting will, almost certainly, not be feasible for some subjects at some HEIs.

18. There is a question as to whether averages for the institution as a whole should also be provided. Neither students, nor other stakeholders, expressed interest in information at this level and we see little point in reporting it. There is also an argument that publication at this level would feed directly into league tables which are considered to be misleading.

19. There should be hyperlinks from the National Survey site to HEI web sites, either to the HEI's home page or a special page where the HEI has chosen to provide a commentary on the CEQ scores. Alternatively the link could be to the relevant department. Users should be able to browse the site, by subject or HEI, and also define searches, for example by geographical region and possibly also entry requirements.

20. The National Survey will need to evolve and this strategic process needs to be overseen by a steering group. We would expect it to comprise, inter alia, representatives from HEFCE, QAA, Universities UK, SCOP, NUS and HESA, and also the other HE funding bodies should they decide to participate in the survey.

21. We recommend that various aspects of the National Survey should be tested through a pilot exercise. This would comprise:

- initial testing and development of the questionnaire, which could be achieved quickly with a small sample

- large scale testing to further validate the questionnaire and explore a number of issues concerned with the conduct and management of the National Survey. This would require several hundred responses.

22. One option would be to run a pilot survey covering a subset of institutions and/or subject areas. This would be significantly cheaper than a census-based pilot, yet would still enable a range of approaches and options to be tested. However, it would delay the introduction of a census-based survey and therefore mean a delay before results covering all institutions could be published. In order to ensure that results are published as soon as possible, a full census survey could be undertaken after the initial development of the questionnaire. This would still be in the nature of a pilot in that the various tests described in this section would be undertaken and the questionnaire, and process, could be modified for subsequent surveys if appropriate. However, if the tests indicated that responses were robust and meaningful then results from the pilot could be published. The disadvantages are that the census-based pilot would be more expensive than a sample-based exercise and there would also be less scope to test options for survey management.

23. On the basis of 350,000 graduates from English HEIs each year we estimate the annual costs of a census-based National Survey would be in the region of £634,000. This includes an allowance of £155,000 for data processing, but does not allow for full set-up and management costs, or the involvement of staff from the HEIs in mailing out questionnaire and reminders.

24. These costs could be reduced in two ways. First, the National Survey could be a sample rather than a census survey. Second, the National Survey could be undertaken less frequently than annually. There is some merit in a biennial survey, but the National Survey would not be able to report accurately on HEIs where the quality of provision changed rapidly.

Foreword

This is a report of a study, commissioned by the Higher Education Funding Council for England (HEFCE), into the collection and use of student feedback information on the quality of teaching and learning. The study was undertaken by a consortium of organisations, led by SQW Limited and also comprising the Centre for Higher Education Research and Information (CHERI) of the Open University and NOP Research Group. The study had two main components:

- to identify good practice by higher education institutions (HEIs) in collecting quantitative and qualitative feedback from students and to make recommendations on the design and implementation of mechanisms by individual institutions

- to make recommendations on the design and implementation of a national survey of recent graduates, the results of which would be published.

CHERI has been primarily responsible for the first of these components and SQW Limited, with advice from NOP Research Group, for the second. However, there are important overlaps between the two parts of the study and staff from CHERI and SQW Limited worked on and contributed to both. Other individuals, in addition to the named authors of this report, contributed to the study. They were Neil Costello, Sarah Francis, Patrick Pringle, and Jane Rindl (SQW Limited) and Alison Ashby and John Richardson, of the Institute for Educational Technology, the Open University. John Richardson undertook the literature review which is presented as Chapter 3 of this report.

During the project we consulted with a large number of organisations and we are grateful for their assistance. Special thanks are due to the staff, students and officers of the Student Unions at the 20 HEIs visited during the study, and also to the many other students and graduates who participated in discussions and pilots of the National Survey.

1 Introduction

Context

1.1 The immediate context for the analysis and recommendations on student feedback contained in this report is the recommendations of the Task Group chaired by Professor Sir Ron Cooke whose report, *Information on quality and standards in higher education* (HEFCE 02/15), was published in November 2001 (and hereafter referred to as the Cooke Report). The Task Group was set up 'to identify the categories of data, information and judgements about quality of teaching and learning that should be available within higher education institutions (HEIs), and those which should be published'. The background to the work of the Task Group was the planned changes to the methods of quality assurance in higher education in England, as described in a consultation document, *Quality assurance in higher education* (HEFCE 01/45). This sought to replace the previous arrangements of institutional audits and subject reviews by the Quality Assurance Agency for Higher Education (QAA) with arrangements that would be characterised by a 'lightness of touch' and a greater recognition of the responsibilities of individual higher education institutions for setting, maintaining and reviewing quality and standards. The new approach built on three main principles, set out in the Cooke Report as follows:

- meeting public information needs, so that stakeholders – and above all students – can obtain information which is up-to-date, consistent and reliable about the quality and standards of teaching and learning at different HEIs

- recognising the primary responsibility of each HEI to operate robust internal mechanisms for setting, maintaining and reviewing quality and standards; for generating information about its quality and standards; and for publishing the key parts of that information

- lightness of touch, so that the burden on HEIs is reduced to the minimum consistent with proper accountability and meeting information needs, and so that the greatest value is secured from the resources used.

1.2 Essentially, the new arrangements are intended to replace processes of external subject review with a greater reliance on and utilisation of the processes and outcomes of institutions' own quality assurance procedures. In most institutions, these latter have developed substantially

over the last ten years, responding in part to the requirements of the external quality assurance procedures operated variously by the higher education funding councils, the Higher Education Quality Council (HEQC) and the QAA.

1.3 Among the pieces of information arising out of the operation of institutional quality assurance arrangements, the Cooke Report recommended that the following should be available in all HEIs:

- information on institutional context
- information on student admission, progression and completion
- information on the HEI's internal procedures for assuring academic quality and standards.

1.4 Among the latter would be 'information on student satisfaction with their higher education experience, covering the views of students on:

- arrangements for academic and tutorial guidance, support and supervision
- library services and IT support
- suitability of accommodation, equipment and facilities for teaching and learning
- perceptions of the quality of teaching and the range of teaching and learning methods
- assessment arrangements
- quality of pastoral care.'

1.5 Moreover, the Task Group went on to recommend that some aspects of information on the quality and standards of teaching and learning should be published. The two elements of such information relevant to the present study are:

- feedback from recent graduates, disaggregated by institution, collected through a national survey

- feedback from current students collected through HEIs' own surveys, undertaken on a more consistent basis than now.

1.6 Feedback from students has always played an important role in the maintenance of quality and standards in higher education. As quality assurance arrangements have themselves become more formalised, so too have the arrangements for the collection, analysis and use of student feedback. However, there are a number of other contextual factors that have influenced these arrangements.

1.7 The expansion and differentiation of British higher education has had major implications for the inner workings of HEIs and for the people who work and study in them. A steady decline in student/staff ratios has meant that the traditionally close relationships between teachers and taught have all but disappeared in most institutions, with the possible exceptions of courses in laboratory or studio based subjects. Thus, informal means of communication between students and their teachers have become less effective in securing reliable feedback. In many institutions, modular forms of course organisation have, whatever their other merits, added to the anonymity of the student experience and a further decline in the opportunities for informal interaction and communication. Other forms of pressure on academic staff including the Research Assessment Exercise (RAE) and, ironically, the pressures of external quality assurance of teaching, have also reduced the time available for informal face-to-face meetings between staff and students. All of these trends have led to the gradual replacement of the informal with the formal, of which the widespread introduction of student feedback surveys has been a conspicuous part. These and other feedback mechanisms have also had to take account, in many institutions, of a more diverse student body among which traditional and homogeneous expectations and attitudes cannot be assumed.

1.8 Other changes in higher education have mirrored changes in the wider society. What has been called the 'new public sector management' has gained ground in most parts of the public sector. Some have argued that this has entailed the replacement of 'trust' with 'accountability' within public sector organisations. There has certainly been a strong trend towards increasing consumer choice through the publication of information about service standards in different institutions. This has heralded a growth in the use of performance indicators, the construction of 'league tables' and a growth and strengthening of management cadres within public sector organisations of all kinds. So too within universities and colleges, new management positions, administrative units, procedures and codes of practice have been introduced. These have brought a measure of standardisation and centralisation into institutions that had traditionally been marked by decentralisation and the professional freedoms of their staff. Competition between institutions – whether in the recruitment of

students or for RAE grades – has been a further aspect of the growth of a new entrepreneurial and accountable world in higher education.

1.9 It is, however, possible to overemphasise these changes or at least to give insufficient attention to the continued importance of some distinctive features of higher education culture. Of these we might mention the collegiality of academic life, the traditions of academic freedoms and departmental autonomies, the strength of subject loyalties and cultures (sometimes against those of the institution) and the ambiguous position of the student in all this – junior member? consumer? stakeholder? These features still exist in HEIs and are an important factor in determining how institutions respond to external initiatives and indeed in how *internal* policies and procedures are implemented. Thus, in the case of student feedback as in much else, one can expect to find a diversity of policy and especially practice both within and between institutions.

1.10 The twin processes of institutional audit and subject review have clearly done much to build up quite extensive formal procedures of quality assurance in most HEIs. Yet, as much of the literature on the subject emphasises, quality assurance has to achieve a balance between accountability and improvement, and it has been suggested that reactions to some aspects of quality assurance have been marked by compliance rather than commitment, at least in some departments and institutions. In particular, that which is imposed from outside the unit or institution in question may be viewed with some suspicion by those within. Sensitive to these features of academic culture, many institutions have allowed considerable variation in the implementation of quality assurance procedures, partly to reflect subject differences and partly to maximise the commitment of staff. Devolved procedures to strengthen staff commitment and to achieve the improvement functions of quality assurance may limit the consistency of approach between and within institutions and the comparability of the information that is generated by such procedures.

1.11 In summary then, the context for this study includes the following:

- a growing emphasis in public policy on consumer choice and competition between institutions – creating new needs for information on quality and standards

- a considerable growth in formal institutional arrangements for quality assurance in recent years, in which student feedback surveys play an important part

- variation between institutions and between departments in the details of these arrangements, to reflect the growing diversity of higher education and to achieve ownership and commitment by staff

- increasing pressures on staff as a result of worsening staff student ratios, RAE etc

- the existence of a lot of institutional information about quality and standards but a lack of consistency and comparability in much of it.

Study objectives

1.12 The project had two main components:

- to review current good practice by HEIs in collecting quantitative and qualitative feedback from students on the quality and standards of their higher education programmes, and using that feedback to secure improvement. And to make recommendations on how individual HEIs can best design and implement their own internal mechanisms with respect to student feedback. This part of the project we refer to as *institutional processes*

- to make recommendations on the design and implementation of a national survey to collect feedback on the quality of teaching and learning from students who have recently graduated, and on publishing the results. This is referred to as the *National Survey*.

1.13 These two components are distinct and they are intended to achieve different things. A national survey should provide comparable and consistent data across all HEIs in England. It would also be concerned with the views of recent graduates. However, in order to be keep within manageable limits, it would cover only a limited number of key questions. Institutional processes, by contrast, need to reflect the particular circumstances and needs of each HEI. But they will generate a much richer and more comprehensive range of qualitative and quantitative information for the institutions to use in identifying how to raise the quality of their programmes. The intention is that the two components of the project will be complementary but the primary aim of the National Survey is to provide comparative information for applicants to HE, whereas the primary aim of institutional processes is to contribute to quality enhancement within the HEI.

Structure of the report

1.14 Following the introduction, Chapter 2 describes the work programme undertaken during the study. Chapter 3 is a review of the literature relating to student feedback. Chapters 4 and 5 present findings and recommendations in relation to institutional processes and the national

survey respectively. Conclusions are presented in Chapter 6. Additional information is presented in appendices.

2 Approach

2.1 The key fieldwork tasks were as follows:

- an invitation was sent to all HEIs to provide written information about their policies and practices on collecting and using student feedback. Accompanying this invitation was a checklist of questions. The replies were analysed to establish the range of, and main approaches to, student feedback and to identify a series of institutional cases to be explored in more detail by the research team. Student Unions were invited to submit information on student feedback processes

- a review of the literature relating to student feedback

- interviews with stakeholders to identify the key topic areas on which the National Survey should provide information

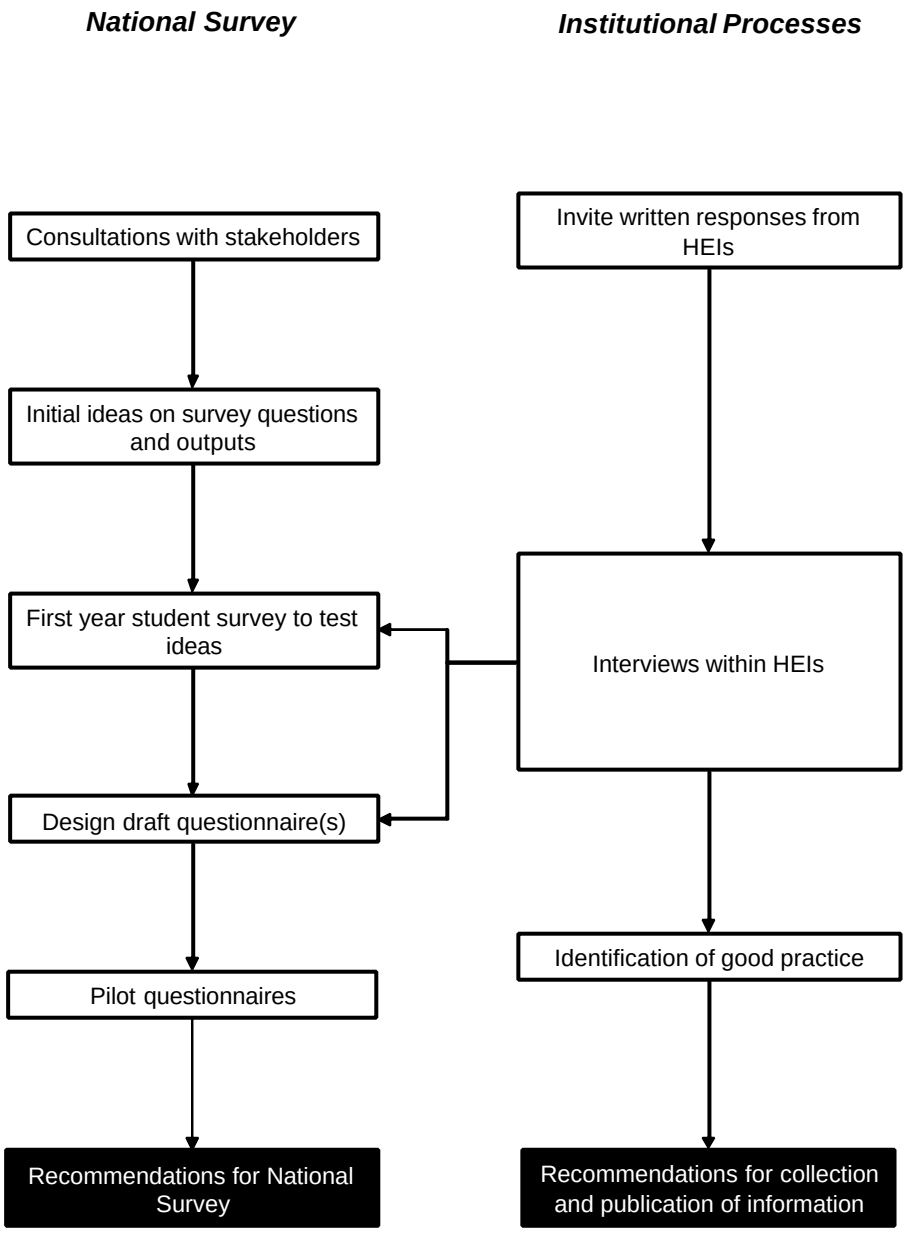
- the previous task enabled us to identify broad topic areas for the National Survey. These topics were tested and refined through discussions with 50 first year students

- a National Survey questionnaire was then piloted

- twenty institutions were selected for more detailed exploration through site visits by members of the research team. The institutions were selected in terms of the analysis of written responses and partly to reflect the diversity of institutions in terms of size, internal structures, history and tradition, mission etc. Interviews were then held with Pro Vice-Chancellors, quality managers, managers of student surveys, careers services, registrars, Student Unions and students, deans and sub-deans, heads of department and other academic staff.

2.2 Many of these tasks were undertaken in parallel and the overall approach to the study is summarised in Figure 2.1.

Figure 2.1: Summary of approach



3 A review of the literature[3]

Introduction

3.1 The purpose of this chapter is to review the published research literature concerning the use of formal instruments to obtain student feedback in higher education. The primary emphasis will be on sources that have been subjected to the formal processes of independent peer review, but there is also a 'grey' literature consisting of conference proceedings, in-house publications and technical reports that contain relevant information even if they are lacking in academic rigour.

3.2 The first part of the review will be concerned with the kinds of instruments that have been used to obtain student feedback. The relevant sources comprise two relatively discrete subsets: the predominantly North American literature concerned with students' evaluations of their teachers; and the predominantly Australian and British literature concerned with students' perceptions of the quality of their programmes. In both cases, questions can be raised about the adequacy of student feedback as a measurement tool. In daily life, tools that measure physical attributes such as length or weight can be trusted because they are both reliable (they yield consistent results) and valid (they measure what they purport to measure), but the reliability and validity of tools intended to measure perceptions and other psychological attributes must be established through empirical research. A section at the end of this chapter describes various methods of defining and assessing the psychometric properties of assessment instruments. Academic staff (not least those with specialist expertise in this field) would resist the use of any instrument that had not been shown to have adequate reliability and validity, and any agency or institution that tried to impose such an instrument would be open to justifiable criticism.

3.3 The second part of the review will be concerned with a number of practical issues involved in the collection of student feedback. Why collect feedback? Why use formal instruments? What should be the subject of the feedback? What kind of feedback should be collected?

[3] This chapter was written by Professor John T. E. Richardson, Institute of Educational Technology, The Open University. He is grateful to the following people for their comments on an earlier draft of this paper: John Brennan, The Open University; Robin Brighton, SQW Limited; Graham Gibbs, The Open University; Herbert Marsh, University of Western Sydney; Keith Trigwell, University of Oxford; Ruth Williams, The Open University and to Hamish Coates, University of Melbourne, for providing data from the work carried out by his colleagues (2001).

When should feedback be collected? Would a single questionnaire be suitable for all students? Why are response rates important? How seriously is student feedback taken? Accordingly, the focus here will be on the practical utility of the instruments that are used to obtain student feedback.

Students' Evaluations of Teaching

3.4 In North America, the practice of obtaining student feedback on individual teachers and course units is widespread. Marsh and Dunkin (1992) identified four purposes for collecting Students' Evaluations of Teaching (SETs):

- diagnostic feedback to teachers about the effectiveness of their teaching
- a measure of teaching effectiveness to be used in administrative decision making
- information for students to use in the selection of course units and teachers
- an outcome or process description for use in research on teaching.

3.5 Marsh and Dunkin noted that the first purpose was essentially universal, whereas the three others were not:

'At many universities systematic student input is required before faculty are even considered for promotion, while at others the inclusion of SETs is optional or not encouraged at all. Similarly, in some universities the results of SETs are sold to students in university bookstores as an aid to the selection of courses or instructors, whereas the results are considered to be strictly confidential at other universities. (p. 143)'

3.6 The feedback in question usually takes the form of students' ratings of their level of satisfaction or their self-reports of other attitudes towards their teachers or their course units. The feedback is obtained by means of standard questionnaires, the responses are automatically scanned, and a descriptive summary of the responses is returned to the relevant teacher and, if appropriate, the teacher's head of department. The process is relatively swift, simple and convenient for both students and teachers, and in most North American institutions it appears to have been accepted as a matter of routine. It has, however, been described as a 'ritual' (Abrami *et al.*, 1996), and precisely for that reason it may not always be regarded as a serious matter by those involved. In many institutions, the instruments used to obtain student feedback have been constructed and developed in-house and may never have been subjected to any kind of external scrutiny. Marsh (1987) described five instruments that had received some kind of formal evaluation, and others have featured in subsequent research.

3.7 The instrument that has been most widely used in published work is Marsh's (1982) Students' Evaluations of Educational Quality (SEEQ). In completing this instrument, students are asked to judge how well each of 35 statements (for instance, 'You found the course intellectually stimulating and challenging') describes their teacher or course unit, using a five-point scale from 'very poor' to 'very good'. The statements are intended to reflect nine different aspects of effective teaching: learning/value, enthusiasm, organisation, group interaction, individual rapport, breadth of coverage, examinations/grading, assignments and workload/difficulty. The evidence using this and other questionnaires has been summarised in a series of reviews (Marsh, 1982, 1987; Marsh & Bailey, 1993; Marsh & Dunkin, 1992; see also Arubayi, 1987).

3.8 The test-retest reliability of students' evaluations is high, even when there is an extended period between the two evaluations. The interrater reliability (see paragraph 3.83 below) of the average ratings given by groups of students is also high, provided that the average is based on 10 or more students. There is a high correlation between the ratings produced by students taking *different* course units taught by the *same* teacher, but little or no association between the ratings given by students taking the *same* course unit taught by *different* teachers. This suggests that students' evaluations are primarily a function of the person teaching the course unit rather than of the particular unit being taught.

3.9 Evaluations of the same teachers given by successive cohorts of students are highly stable over time. Indeed, Marsh and Hocevar (1991b) found no systematic changes in students' ratings of 195 teachers over a 13-year period. Although this does demonstrate the stability of the students' ratings, it also implies that the performance of the teachers was not improving with experience. Nevertheless, Roche and Marsh (2002) found that teachers' perceptions of their own teaching became more consistent with their students' perceptions of their teaching as a consequence of receiving feedback in the form of students' evaluations. In other words, students' evaluations may change teachers' self-perceptions even if they do not change their teaching behaviour.

3.10 The factor structure of the SEEQ has been confirmed in several studies. In particular, Marsh and Hocevar (1991a) showed that it was invariant across teachers of different status and across course units in different disciplines and at different levels. There is a consensus that students' ratings of teaching effectiveness vary on a large number of dimensions, but there is debate as to whether these can be subsumed under a single, more global dimension. Marsh (1991; Marsh & Dunkin, 1992; Marsh & Roche, 1997) argued that, although students' scores on the dimensions of the SEEQ were correlated with each other, they could not be adequately captured by a single higher-order factor. On the other hand, Abrami and d'Apollonia (1991; Abrami, d'Apollonia, & Rosenfield, 1996; d'Apollonia & Abrami, 1997) proposed that

students' evaluations of teaching were dominated by a single overarching construct that they called 'general instructional skill'.

3.11 The fact that students' evaluations of teachers are correlated with the teachers' self-evaluations also constitutes evidence for their validity. In fact, teachers' self-evaluations exhibit essentially the same factor structure as their students' evaluations, teachers' self-evaluations are correlated with their students' evaluations on each individual dimension of the SEEQ, and teachers' self-evaluations are not systematically different from their students' evaluations (see Marsh, 1987). Students' evaluations of their teachers are not highly correlated with evaluations provided by *other* teachers on the basis of classroom observation. Nevertheless, both the reliability and the validity of the latter evaluations have been questioned. There is better evidence that SETs are correlated with ratings of specific aspects of teaching by trained observers (e.g., Murray, 1983).

3.12 In principle, the validity of students' evaluations might be demonstrated by finding correlations between SETs and academic performance. However, the demands and the assessment criteria of different course units may vary, and so students' grades or examination marks cannot be taken as a simple measure of teaching effectiveness. One solution is to compare students' evaluations and attainment in a single course unit where different groups of students are taught by different teachers but receive the same form of assessment (a *multisection validity study*). In these circumstances, there is a clear relationship between SETs and academic attainment, even when the grades are assigned by an independent evaluator, though some aspects of teaching are more important in predicting attainment than others (Cohen, 1981; Marsh, 1987).

3.13 The relationship between SETs and academic attainment is stronger when students know their final grades, though there is still a moderate correlation if they provide their ratings before their final grades are known (Cohen, 1981). Greenwald and Gilmore (1997a, 1997b) noted that in the latter case the students can acquire expectations about their final grades from the results of their intermediate assessments. They found a positive relationship between students' expected grades and their overall ratings of their teaching but a negative relationship between students' expected grades and their estimated workload. They argued that students reduced their work investment to achieve their original aspirations when faced with lenient assessment on their midterm tests.

3.14 The latter research raises the possibility that SETs might be biased by the effects of extraneous background factors, a possibility that is often used to foster scepticism about the value of SETs in the evaluation of teaching in higher education (Husbands & Fosh, 1993). Marsh (1987) found that four variables were potentially important in predicting SETs: the

students' prior interest in the subject matter; their expected grades; their perceived workload; and their reasons for taking the course unit in question. Nevertheless, the effects of these variables upon students' ratings were relatively weak and did not necessarily constitute a bias. For instance, course units that were perceived to have a higher workload received more positive ratings, and the effect of prior interest was mainly on what students said they had learned from the course unit rather than their evaluation of the teaching *per se* (see Marsh, 1983).

3.15 Marsh (1987) acknowledged in particular that more positive SETs could in principle arise from the students' satisfaction at receiving higher grades (the *grading satisfaction hypothesis*) or else from other uncontrolled characteristics of the student population. The fact that the relationship between SETs and academic attainment is stronger when the students know their final grades is consistent with the grading satisfaction hypothesis. However, Marsh pointed out that, if students are taught in different groups on the same course unit, they may know how their attainment compares with that of the other students in their group, but they have no basis for knowing how their attainment compares with that of the students in other groups. Yet the correlation between SETs and academic attainment arises even when it is calculated from the average SETs and the average attainment across different groups, and even when the different groups of students do not vary significantly in terms of the grades that they expect to achieve. Marsh argued that this was inconsistent with the grading satisfaction hypothesis and supported the validity of SETs.

3.16 Although the SEEQ has been most widely used in North America, it has also been employed in investigations carried out in Australia, New Zealand, Papua New Guinea and Spain (Clarkson, 1984; Marsh, 1981, 1986; Marsh & Roche, 1992; Marsh, Touron, & Wheeler, 1985; Watkins, Marsh, & Young, 1987). The instrument clearly has to be adapted (or translated) for different educational settings, and in some of these studies a different response scale was used. Even so, in each case both the reliability and the validity of the SEEQ were confirmed. In a trial carried out by the Curtin University of Technology Teaching Learning Group (1997), the SEEQ was found to be far more acceptable to teachers than the existing in-house instrument. Coffey and Gibbs (2001) arranged for a shortened version of the SEEQ (containing 24 items from six scales) to be administered to students at nine universities in the United Kingdom. The results confirmed the intended factor structure of this inventory and also showed a high level of internal consistency. Because cross-cultural research tended to confirm the factor structure of the SEEQ, Marsh and Roche (1994) argued that it was especially appropriate for the increasingly multicultural student population attending Australian universities.

3.17 In a further study, Coffey and Gibbs (in press) asked 399 new teachers from eight countries to complete a questionnaire concerned with their approaches to teaching. They found that those teachers who adopted a student-focused or learning-centred approach to teaching received significantly higher ratings from their students on five of the six scales in the shortened SEEQ than did those teachers who adopted a teacher-focused or subject-centred approach to teaching. In the case of teachers who had completed the first semester of a training programme, Coffey and Gibbs (2000) found that their students gave them significantly higher ratings on four of the six scales in the shortened SEEQ at the end of the semester than they had done after four weeks. Nevertheless, this study suffered from a severe attrition of participants, and it is possible that the latter effect was simply an artefact resulting from sampling bias. Equally, the students may have given more positive ratings simply because they were more familiar with their teachers.

3.18 SETs are most commonly obtained when the teaching process is face-to-face and controlled by a single lecturer or instructor. It has indeed been suggested that the routine use of questionnaires to obtain students' evaluations of their teachers promotes an uncritical acceptance of traditional conceptions of teaching based on the bare transmission of knowledge and the neglect of more sophisticated conceptions concerned with the promotion of critical thinking and self-expression (Kolitch & Dean, 1999). It should be possible to collect SETs in other teaching situations such as the supervision of research students, but there has been little or no research on the matter.

3.19 One very different situation is that of distance education, where students are both physically and socially separated from their teachers, from their institutions, and often from other students too (Kahl & Cropley, 1986). To reduce what Moore (1980) called the 'transactional distance' with their students, most distance-learning institutions use various kinds of personal support, such as tutorials or self-help groups arranged on a local basis, induction courses or residential schools, and teleconferencing or computer conferencing. This support seems to be highly valued by the students in question (Fung & Carr, 2000; Hennessy, Flude, & Tait, 1999). However, it means that 'teachers' have different roles in distance education: as authors of course materials and as tutors. Gibbs and Coffey (2001) suggested that collecting SETs in distance education could help to clarify the expectations of tutors and students about the nature of their relationship.

3.20 The intellectual rights and copyright in the SEEQ belong to Professor Herbert W. Marsh of the University of Western Sydney, Macarthur. It is presented on a double-sided form that allows for the inclusion of supplementary items and open-ended questions. If the SEEQ is administered in a class setting, respondents may be asked to record the course unit and the teacher being rated, but they themselves can remain anonymous. Marsh and Roche (1994)

elaborated the SEEQ as the core of a self-development package for university teachers that incorporates a self-rating questionnaire for teachers, a guide to interpreting the students' overall evaluations and booklets on improving teaching effectiveness in areas where evaluations identify scope for improvement. They offered advice on how this package might be adopted in programmes at other institutions.

3.21 Marsh (1987) concluded that 'student ratings are clearly multidimensional, quite reliable, reasonably valid, relatively uncontaminated by many variables often seen as sources of potential bias, and are seen to be useful by students, faculty, and administrators' (p. 369). The literature that has been published in the subsequent 15 years has confirmed each of these points and has also demonstrated that student ratings can provide important evidence for research on teaching. The routine collection of students' evaluations does not in itself lead to any improvement in the quality of teaching (Kember, Leung, & Kwan, 2002). Nevertheless, feedback of this nature may help in the professional development of individual teachers, particularly if it is supported by an appropriate process of consultation and counselling (Roche & Marsh, 2002). SETs do increase systematically following specific interventions aimed at improving teaching (Hativa, 1996).

Student satisfaction surveys

3.22 However, perhaps the most serious limitation of the instruments that have been described thus far is that they have focused upon students' evaluations of particular course units in the context of highly modular programmes of study, and hence they provide little information about their experience of their programmes or their institutions as a whole. In addition to collecting SETs for individual course units, many institutions in North America make use of commercially published questionnaires to collect comparative data on their students' overall satisfaction as consumers. One widely used questionnaire is the Noel-Levitz Student Satisfaction Inventory, which is based explicitly on consumer theory and measures the students' satisfaction with their experience of higher education. It contains either 76 items (for institutions offering two-year programmes) or 79 items (for institutions offering four-year programmes); in each case, respondents are asked to rate both the importance of their expectation about a particular aspect of higher education and their level of satisfaction. Overall scores are calculated that identify aspects of the students' experience where the institutions are failing to meet their expectations.

3.23 A similar approach has been adopted in in-house satisfaction surveys developed in the United Kingdom, but most have of these have not been adequately documented or evaluated. Harvey (1997) described a general methodology for developing student satisfaction surveys based upon their use at the University of Central England. First, significant aspects of students'

experience are identified from the use of focus groups. Second, these are incorporated into a questionnaire survey in which larger samples of students are asked to rate their satisfaction with each aspect and its importance to their learning experience. Finally, the responses from the survey are used to identify aspects of the student experience that are associated with high levels of importance but low levels of satisfaction. According to Harvey (2001), this methodology has been adopted at a number of institutions in the United Kingdom and in some other countries. Descriptive data from such surveys have been reported in institutional reports (e.g., Harvey, 1995), but no formal evidence with regard to their reliability or validity has been published.

Students' perceptions of academic quality

3.24 From the perspective of an HEI seeking to maintain and improve the quality of its teaching, it could be argued that the appropriate focus of assessment would be a programme of study rather than an individual course unit or the whole institution, and this has been the dominant focus in Australia and the United Kingdom.

3.25 In an investigation into determinants of approaches to studying in higher education, Ramsden and Entwistle (1981) developed the Course Perceptions Questionnaire (CPQ) to measure the experiences of British students in particular degree programmes and departments. In its final version, the CPQ contained 40 items in eight scales that reflected different aspects of effective teaching. It was employed by Ramsden and Entwistle in a survey of 2,208 students across 66 academic departments of engineering, physics, economics, psychology, history and English. A factor analysis of their scores on the eight scales suggested the existence of two underlying dimensions: one reflected the positive evaluation of teaching and programmes, and the other reflected the use of formal methods of teaching and the programmes' vocational relevance.

3.26 The CPQ was devised as a research instrument to identify and to compare the perceptions of students on different programmes, and Ramsden and Entwistle were able to use it to reveal the impact of contextual factors on students' approaches to learning. However, the primary factor that underlies its constituent scales is open to a natural interpretation as a measure of perceived teaching quality, and Gibbs, Habeshaw, and Habeshaw (1988, pp. 29-33) argued that the CPQ could be used for teaching evaluation and course review. Even so, the correlations obtained by Ramsden and Entwistle between students' perceptions and their approaches to studying were relatively weak. Similar results were found by other researchers (Parsons, 1988), and this led to doubts being raised about the adequacy of the CPQ as a research tool (Meyer & Muller, 1990).

3.27 Ramsden (1991a) developed a revised instrument, the Course Experience Questionnaire (CEQ), specifically as a performance indicator for monitoring the quality of teaching on particular academic programmes. In the light of preliminary evidence, a national trial of the CEQ was commissioned by a research group set up by the Australian Commonwealth Department of Employment, Education and Training to examine performance indicators in higher education (Linke, 1991). In this national trial, usable responses to the CEQ were obtained from a total of 3,372 final-year undergraduate students at 13 Australian universities and colleges of advanced education (see also Ramsden, 1991b).

3.28 The instrument used in this trial consisted of 30 items in five scales which had been identified in previous research as reflecting different dimensions of effective instruction: Good Teaching (8 items); Clear Goals and Standards (5 items); Appropriate Workload (5 items); Appropriate Assessment (6 items); and Emphasis on Independence (6 items). The defining items of the five scales (according to the results of the national trial) are shown in the table below. In addition, three of the items in the Appropriate Assessment scale could be used as a subscale to monitor the perceived importance of rote memory as opposed to understanding in academic study.

3.29 The respondents were instructed to indicate their level of agreement or disagreement (along a scale from 'definitely agree', scoring 5, to 'definitely disagree', scoring 1) with each statement as a description of their programme of study. Half of the items referred to positive aspects, whereas the other half referred to negative aspects and were to be scored in reverse. This means that the instrument as a whole controlled for any systematic responses biases either to agree with all of the items or to disagree with all of the items. (Unfortunately, the items to be scored in reverse were not distributed equally across the five CEQ scales.)

Scale	**Defining item**
Good Teaching	Teaching staff here normally give helpful feedback on how you are doing.
Clear Goals and Standards	You usually have a clear idea of where you're going and what's expected of you in this course.
Appropriate Workload	The sheer volume of work to be got through in this course means you can't comprehend it all thoroughly.
Appropriate Assessment	Staff here seem more interested in testing what we have memorised than what we have understood.
Emphasis on Independence	Students here are given a lot of choice in the work they have to do.

3.30 As a result of this national trial, it was determined that the Graduate Careers Council of Australia (GCCA) should administer the CEQ on an annual basis to all new graduates through the Graduate Destination Survey, which is conducted a few months after the completion of their degree programmes. The survey of the 1992 graduates was carried out in 1993 and obtained usable responses to the CEQ from more than 50,000 graduates from 30 institutions of higher education (Ainley & Long, 1994). Subsequent surveys have covered all Australian universities and have typically obtained usable responses to the CEQ from more than 80,000 graduates, reflecting overall response rates of around 60% (Ainley & Long, 1995; Johnson, 1997, 1998, 1999; Johnson, Ainley, & Long, 1996; Long & Hillman, 2000). However, in the GCCA surveys, the original version of the CEQ has been modified in certain respects:

- in response to concerns about the employability of graduates, a Generic Skills scale was added to 'investigate the extent to which higher education contributes to the enhancement of skills relevant to employment' (Ainley & Long, 1994, p. xii). This contains six new items that are concerned with problem solving, analytic skills, teamwork, communication and work planning. Of course, similar concerns about the process skills of graduates have been expressed in the United Kingdom (Committee of Vice-Chancellors and Principals, 1998). The items in the Generic Skills scale are somewhat different from those in the rest of the CEQ, insofar as they ask respondents to evaluate the skills that they have gained from their programmes rather than the quality of the programmes themselves. Other researchers have devised more extensive instruments for measuring graduates' perceptions of their personal development during their programmes of study (e.g., Cheng, 2001; Purcell & Pitcher, 1998)

- to compensate for this and reduce the length of the questionnaire still further, the Emphasis on Independence scale was dropped, and a further seven items were removed on the grounds that they had shown only a weak relationship with the scales to which they had been assigned in Ramsden's (1991a; 1991b, p. 6) analysis of the data from the Australian national trial. This produced a revised, short form of the CEQ consisting of 23 items in five scales

- two other items were employed but not assigned to any of the scales. One measured the respondents' overall level of satisfaction with their programmes, and this has proved to be helpful in validating the CEQ as an index of perceived academic quality (see below). An additional item in the first two surveys was concerned with the extent to which respondents perceived their programmes to be overly theoretical or abstract. This was replaced in the next three surveys by reinstating an item from the Appropriate Assessment scale that measured the extent to which feedback on their work was usually provided only in the form of marks or grades. In subsequent

surveys, this in turn was replaced by a wholly new item concerned with whether the assessment methods required an in-depth understanding of the syllabus. In practice, however, the responses to these additional items have not shown a strong relationship with those given to other items from the Appropriate Assessment scale, and so they have not been used in computing the respondents' scale scores.

3.31 Wilson, Lizzio, and Ramsden (1997) proposed that for research purposes the original version of the CEQ should be augmented with the Generic Skills scale to yield a 36-item instrument. They compared the findings obtained using the short, 23-item version and this 36-item version when administered to successive cohorts of graduates from one Australian university.

3.32 Evidence concerning the psychometric properties of the 30-item version of the CEQ has been obtained in the Australian national trial (Ramsden, 1991a, 1991b) and in research carried out in individual universities in Australia (Trigwell & Prosser, 1991) and Britain (Richardson, 1994). Evidence concerning the psychometric properties of the 23-item version of the CEQ has been obtained in the GCCA surveys and in the study by Wilson *et al.* (1997); the latter also provided evidence concerning the psychometric properties of the 36-item version of the CEQ.

3.33 The internal consistency of the scales as measured by Cronbach's (1951) coefficient alpha is generally satisfactory. There is unfortunately no evidence on the CEQ's test-retest reliability. The composition of the scales according to the results of factor analyses conducted on the responses to individual items is broadly satisfactory. In the 23-item version, all of the items tend to load on distinct factors reflecting their assigned scales. The application of Rasch's (1960) measurement analysis confirms the multidimensional structure of the CEQ (Ainley, 1999; cf. Waugh, 1998). In the 30-item and the 36-item versions, most items load on factors reflecting their assigned scales, but there is a consistent tendency for a few items on the Good Teaching scale and the Emphasis on Independence scale to load on other factors.

3.34 The construct validity of the CEQ according to the results of factor analyses on the respondents' scores on the constituent scales is also broadly satisfactory. The modal solution is a single factor on which all of the scales show significant loadings. The Appropriate Workload scale shows the lowest loadings on this factor, and there is some debate over whether it could be taken to define a separate dimension (Ainley, 1999; Richardson, 1997). The criterion validity of the CEQ as an index of perceived quality can be tested by examining the correlations between the respondents' scores on the constituent scales and their responses to the additional item concerned with their overall satisfaction. Typically, all of the scales show statistically significant correlations with ratings of satisfaction, but the Appropriate Workload scale shows the weakest associations.

3.35 The discriminant validity of the CEQ is shown by the fact that the respondents' scores on the constituent scales vary across different academic disciplines and across different institutions of higher education offering programmes in the same discipline. In particular, students produce higher scores in departments that pursue student-centred or experiential curricula through such models as problem-based learning (see also Eley, 1992; Sadlo, 1997). Conversely, Ainley and Long (1995) used results from the 1994 GCCA survey to identify departments of psychology in which there was 'the possible need for review of teaching and assessment practices' (p. 50). Long and Hillman (2000, pp. 25-29) found in particular that ratings on the Good Teaching scale as well as students' overall level of satisfaction varied inversely with the size of their institution.

3.36 As mentioned earlier, Ramsden and Entwistle (1981) were originally concerned to demonstrate a connection between students' perceptions of their programmes and the approaches to learning that they adopted on those programmes. The weak relationships that they and other researchers found cast doubt upon the concurrent validity of the CPQ. In contrast, investigations carried out at the Open University have shown an intimate relationship between the scores obtained on the CEQ by students taking different course units and their self-reported approaches to studying, such that the students who evaluate their course units more positively on the CEQ are more likely to adopt a deep approach to learning (Lawless & Richardson, 2002; Richardson & Price, 2003). Typically, the two sets of measures share between 45% and 70% of their variance. Similar results were obtained by Trigwell and Ashwin (2002), who used an adapted version of the CEQ to assess perceptions of the tutorial system among students at an Oxford college.

3.37 Wilson *et al.* (1997) demonstrated that students' scores on the 36-item version of the CEQ were significantly correlated with their cumulative grade point averages. The correlation coefficients were highest for the Good Teaching scale and the Clear Goals and Standards scale, and they were lowest for the Generic Skills and Appropriate Workload scales. Of course, these data do not imply a causal link between good teaching and better grades. As mentioned above, Marsh (1987) pointed out that more positive student ratings could result from students' satisfaction at receiving higher grades or from uncontrolled characteristics of the student population. In both these cases, however, it is not clear why the magnitude of the relationship between CEQ scores and academic attainment should vary across different scales of the CEQ.

3.38 Finally, Lizzio, Wilson, and Simons (2002) constructed a theoretical model of the relationships between CEQ scores, approaches to studying and academic outcomes. They interpreted scores on the Generic Skills scale and students' overall ratings of satisfaction as outcome measures, as well as grade point average. In general, they found that students'

scores on the other five scales of the CEQ were positively correlated with all three outcome measures. Students' perceptions of their academic environment according to the CEQ had both a direct influence upon academic outcomes and an indirect influence that was mediated by changes in the students' approaches to studying. In contrast, students' academic achievement before their admission to university had only a weak influence on their grade point average and no effect on their overall satisfaction.

3.39 Although the CEQ has been predominantly used in Australia, it has also been employed in other countries to compare graduates from different programmes and to compare current students. For instance, Sadlo (1997) used the CEQ to compare students taking undergraduate programmes in occupational therapy at HEIs in six different countries. In the United Kingdom, the 30-item version of the CEQ has been used both for academic review (Richardson, 1994) and for course development (Gregory, Harland, & Thorley, 1995; Gregory, Thorley, & Harland, 1994). Wilson et al. (1997) advised that the CEQ was not intended to provide feedback with regard to individual subjects or teachers. Nevertheless, Prosser, Trigwell, Hazel, and Gallagher (1994) adapted the CEQ to refer to particular topics (such as mechanics in a physics programme or photosynthesis in a biology programme), and a modified version of the CEQ concerned with students' perceptions of individual course units has been used to compare their experience of large and small classes (Gibbs & Lucas, 1996; Lucas, Gibbs, Hughes, Jones, & Wisker, 1997). The Curtin University of Technology Teaching Learning Group (1997) reworded the 23-item version of the CEQ to refer to the lecturer teaching a specific course unit, and they proposed that it might complement the SEEQ in the evaluation of individual lecturers.

3.40 The intellectual rights and the copyright in the CEQ belong to Professor Paul Ramsden of the University of Sydney, Australia, the Graduate Careers Council of Australia (GCCA) and the Australian Commonwealth Department of Education, Training and Youth Affairs. Like the SEEQ, it can be conveniently presented on a double-sided form and the responses automatically scanned. In the GCCA surveys, a descriptive summary of the average ratings given to each programme at each institution is published, provided that the response rate at the institution in question has exceeded 50%. (Normally this is achieved by all except a few private institutions.) Once again, the process seems to have been accepted as a matter of routine in most Australian institutions, and some are using versions of the CEQ to monitor their current students. At the University of Sydney, for example, the average ratings obtained on an adapted Student Course Experience Questionnaire determine a portion of the financial resources that are allocated to each faculty.

3.41 One problem with the CEQ is that the wording of the constituent items may not be suitable for all students. For instance, Johnson *et al*. (1996, p. 3) remarked that the appropriateness of

some items was questionable in the case of respondents who had completed a qualification through a programme of research, since in this case the notion of meeting the requirements of a particular 'course' might be quite tenuous. In response, a separate instrument, the Postgraduate Research Experience Questionnaire was developed (Johnson, 1999, p. 11). Initial findings obtained with this instrument indicated that it had reasonable internal consistency and a consistent structure based on six dimensions: Supervision, Skill Development, Intellectual Climate, Infrastructure, Thesis Examination, and Goals and Expectations. This instrument is now employed across the Australian university system, and the results are returned to institutions but are not published.

3.42 Nevertheless, further research demonstrated that the questionnaire did not discriminate among different universities or among different disciplines at the same university (Marsh, Rowe, & Martin, 1999). As a result, there is considerable scepticism about whether it provides an adequate basis for benchmarking universities or disciplines within universities. One difficulty is the lack of a coherent research base on the experiences of postgraduate research students, and this has encouraged the use of totally ad hoc instruments to measure their perceptions of quality. Another difficulty is that evaluations of research training typically confound the overall quality of the research environment with the practice of individual supervisors. It is only very recently that researchers and institutions have recognised the need to distinguish institutional monitoring from enhancing supervisory practice (Chiang, 2002; Pearson, Kayrooz, & Collins, 2002).

3.43 The GCCA surveys also embrace students who have studied by distance education, for whom items referring to 'lecturers' or 'teaching staff' might be inappropriate. As mentioned earlier, academic staff in distance-learning institutions have two rather different roles: as the authors of course materials and as course tutors. Richardson and Woodley (2001) adapted the CEQ for use in distance education by amending any references to 'lecturers' or to 'teaching staff' so that the relevant items referred either to teaching materials or to tutors, as appropriate. The amended version was then used in a postal survey of students with and without a hearing loss who were taking course units by distance learning with the Open University. A factor analysis of their responses confirmed the intended structure of the CEQ, except that the Good Teaching scale split into two scales concerned with good materials and good tutoring. Similar results were obtained by Lawless and Richardson (2002) and by Richardson and Price (2003), suggesting that this amended version of the CEQ is highly robust in this distinctive context.

3.44 Although the CEQ was intended to differentiate between students taking different programmes of study, the GCCA surveys have also identified apparent differences related to the demographic characteristics of the respondents, including gender, age, first language and ethnicity. However, the authors of the annual reports from the GCCA surveys have been at

pains to point out that these effects could simply reflect the enrolment of different kinds of student on programmes in different disciplines with different teaching practices and different assessment requirements. In other words, observed variations in CEQ scores might arise from respondents taking different programmes rather than from inherent characteristics of the respondents themselves. Indeed, in research with Open University students taking particular course units (Richardson & Price, 2003), students' demographic characteristics such as gender, age and prior education did not show any significant relationship with their perceptions of the academic quality of their courses.

3.45 One potential criticism of the CEQ is that it does not include any items relating to the pastoral, physical or social support of students in higher education. In principle, it is entirely possible to include additional items concerned with institutional facilities, such as computing and library resources. In fact, some institutions involved in the Australian graduate surveys have included extra items regarding administrative matters, student services and recreational facilities, but these additional items were not considered in the published analysis of results from the CEQ (Johnson *et al.*, 1996, p. 3). An initial analysis suggested that students' satisfaction with their facilities was a much weaker prediction of their overall satisfaction than the original scales in the CEQ (Wilson *et al.*, 1997). As Johnson *et al.* (1996, p. 5) noted, the CEQ does not claim to be comprehensive but seeks information about dimensions of teaching and learning that appear to be central to the majority of academic subjects taught in HEIs.

3.46 Nevertheless, discussions in focus groups with stakeholders and analyses of the responses to open-ended questions included in the CEQ motivated further research. McInnis, Griffin, James, and Coates (2001) devised six new scales, each containing five items, to measure the following domains: Student Support, Learning Resources, Course Organisation, Learning Community, Graduate Qualities and Intellectual Motivation. The properties of the Course Organisation scale proved to be unsatisfactory, but McInnis et al. suggested that the other five scales could be used by institutions in the annual surveys of their graduates. This would yield an extended version of the CEQ containing 50 items. McInnis et al. found that students' scores on the new scales were correlated with their scores on the five original scales of the 23-item CEQ, and they concluded that the inclusion of the new scales had not affected their responses to the original scales (p. x).

3.47 However, McInnis et al. did not examine the constituent structure of their extended instrument in any detail. They have provided a table of correlation coefficients among the scores of 2,316 students on the five original scales of the 23-item CEQ and their six new scales. A factor analysis of the students' scores on all 11 scales also yields a single underlying dimension, but this is mainly dominated by the new scales at the expense of the original scales. This indicates that the extended 50-item version of the CEQ is perceived by

students as mainly concerned with informal aspects of higher education (such as resources and support systems). Like the Generic Skills scale, the new scales were introduced for largely pragmatic reasons and are not grounded in research on the student experience. Accordingly, although the extended CEQ taps a broader range of students' opinions, it may well be less appropriate for measuring their perceptions of the more formal aspects of the curriculum that are usually understood to define teaching quality.

3.48 As in the case of students' evaluations of teaching, there is little evidence that the collection of student feedback using the CEQ in itself leads to any improvement in the perceived quality of programmes of study. However, the proportion of graduates who agreed that they were satisfied with their programmes of study in the GCCA surveys has gradually increased from 60% in 1995 to 68% in 2001, while the proportion who disagreed decreased from 14% to 10% over the same period ('Graduate Satisfaction', 2001). By analogy with the limited amount of evidence on the value of SETs, students' scores on the CEQ might assist in the process of course development, especially if used in a systematic process involving consultation and counselling (Gregory et al., 1994, 1995), and they might also be expected to improve following specific interventions aimed at improving the quality of teaching and learning across entire programmes of study.

Practical issues in obtaining student feedback

Why obtain student feedback?

3.49 In principle, student feedback can be obtained for at least three different reasons: to monitor the quality of teaching and learning; to improve the quality of teaching and learning; and to advise potential students about the quality of teaching and learning. Clearly, both students' evaluations of teaching and their perceptions of academic quality have been investigated in different studies with each of these aims in mind. The research evidence suggests: that student feedback provides an important source of evidence for assessing quality; that it can be used to inform attempts to improve quality (but simply collecting such feedback is unlikely to lead to such improvements); and that student feedback can be communicated in a way that is informative to future students.

Why use formal instruments?

3.50 Student feedback can be obtained in many ways other than through the administration of formal questionnaires. These include casual comments made inside or outside the classroom, meetings of staff-student committees, and student representation on institutional bodies. Good practice would encourage the use of all these means to maintain and enhance the quality of

teaching and learning in higher education. However, surveys using formal instruments have two advantages: they provide an opportunity to obtain feedback from the entire population of students; and they document the experiences of the student population in a more or less systematic way.

3.51 In principle, one could obtain student feedback using open-ended questionnaires. These might be especially appropriate on programmes in education, the humanities and the social sciences, where students are often encouraged to be sceptical about the value of quantitative methods for understanding human experience. Nevertheless, the burden of analysing open-ended responses and other qualitative data is immense, even with only a relatively modest sample. The process of data analysis becomes quite intractable with larger samples unless there are a limited number of response alternatives to each question that can be encoded in a straightforward manner. The use of quantitative inventories to obtain student feedback has therefore been dictated by purely organisational constraints, particularly given the increasing size of classes in higher education. The content of such instruments could, of course, be based on results from qualitative research, as in CEQ, or from focus groups, as in Harvey's (1997) student satisfaction methodology.

3.52 In addition, informal feedback is mainly available when teachers and learners are involved in face-to-face situations. In distance education, as mentioned earlier, students are both physically and socially separated from their teachers and their institutions, and this severely constrains the opportunities for obtaining student feedback. In this situation, the use of formal inventories has been dictated by geographical factors as much as by organisational ones (Morgan, 1984). It can be argued that it is not appropriate to compare the reports of students at institutions (such as the Open University) which are wholly committed to distance education, with the reports of students at institutions which are wholly committed to face-to-face education. However, it presumably *is* appropriate to compare the reports of distance-learning and campus-based students taking the same programmes at the large number of institutions that offer both modes of course delivery, and this provides a further constraint on the choice of methods for obtaining student feedback.

What should be the subject of the feedback?

3.53 Student feedback can be obtained on teachers, course units, programmes of study, departments and institutions. At one extreme, one could envisage a teacher seeking feedback on a particular lecture; at the other extreme, one might envisage obtaining feedback on a national system of higher education, especially with regard to controversial developments such as the introduction of top-up fees. Nevertheless, it is clearly sensible to seek feedback at a level that is appropriate to one's basic goals. If the aim is to assess or improve the quality of

particular teachers, they should be the subject of feedback. If the aim is to assess or improve the quality of particular programmes, the latter should be the subject of feedback. There is no evidence that obtaining feedback at one level is useful or effective in monitoring or improving quality at another level.

What kind of feedback should be collected?

3.54 Most of the research evidence has been concerned with students' perceptions of the quality of the teaching that they receive or their more global perceptions of the academic quality of their programmes. Much less evidence has been concerned with students' level of satisfaction with the teaching that they receive or with their programmes in general. Consumer theory maintains that the difference between consumers' expectations and perceptions determines their level of satisfaction with the quality of provision of a service. This assumption is embodied in American instruments such as the Noel-Levitz Student Satisfaction Inventory and also in Harvey's (1997) student satisfaction methodology. (Indeed, one could, in principle, modify the CEQ to measure students' expectations when embarking on a programme as well as their subsequent perceptions of its academic quality.) This approach was extended by Narasimhan (2001) to incorporate the expectations and perceptions of teachers in higher education as well as those of their students.

3.55 One fundamental difficulty with this approach is that it privileges satisfaction as a notion that is coherent, homogeneous and unproblematic. In fact, the limited amount of research on this topic suggests that student satisfaction is a complex yet poorly articulated notion that is influenced by a wide variety of contextual factors which are not intrinsically related to the quality of teaching (Wiers-Jenssen, Stensaker, & Grøgaard, 2002). In the case of the CEQ, in contrast, satisfaction ratings are simply used as one way to validate students' perceptions of academic quality, which are themselves regarded as being multidimensional in nature.) The discomfort that is associated with the process of intellectual development during higher education has been well documented in interview-based research by Perry (1970) and Baxter Magolda (1992). It is, in any case, hard to justify the satisfaction of students as a fundamental goal of higher education in its own right, and to that extent higher education should not be likened to a commodity or service. This is not to argue that satisfaction ratings are wholly irrelevant to institutions (positive ratings may prove very useful for marketing purposes), simply that they are uninformative about issues of quality.

3.56 A different issue is whether student feedback should be concerned solely with curricular matters or whether it should also be concerned with the entire range of facilities available at institutions of higher education (including computing, library, recreational and sporting facilities). Although the latter considerations are undoubtedly important in evaluating the

student experience, it can be argued that they are not intrinsic to the quality of teaching and learning. There is research evidence that students' perceptions of institutional facilities are less important as predictors of their overall satisfaction than their perceptions of the academic features of their programmes. Moreover, including additional scales about the broader institutional environment in feedback questionnaires might undermine those instruments as indicators of teaching quality. It would be preferable to evaluate institutional facilities as an entirely separate exercise, and in this case an approach that was orientated towards consumer satisfaction might well be entirely appropriate.

When should feedback be collected?

3.57 In principle, it would seem sensible to collect feedback on students' experience of a particular educational activity at the completion of that activity, since it is presumably their experience of the entire activity that is of interest. In other words, it would be most appropriate to seek student feedback at the end of a particular course unit or programme of study. Nevertheless, some other suggestions have been put forward. Narasimhan (2001) noted that obtaining feedback at the end of a course unit could not benefit the respondents themselves and that earlier feedback would be of more immediate value. Indeed, Greenwald and Gilmore (1997a, 1997b) found that students' perceptions in the middle of a course unit influenced their subsequent studying and final grades.

3.58 Others have suggested that the benefits or otherwise of having completed a programme of study are not immediately apparent to the new graduates, and hence feedback should be sought some time after graduation. Indeed, from a purely practical point of view, it would be both convenient and economical to obtain feedback from recent graduates as part of the First Destination Survey (FDS). Concern has been expressed that this might reduce the response rate to the FDS and thus impair the quality of the information that is available about graduate employment. However, the converse is also possible: that incorporating the FDS might reduce the response rate to a survey of graduates' perceptions of the quality of their programmes. This would be a serious possibility if the questionnaire to be used for the FDS were perceived as either cumbersome or intrusive.

3.59 Of course, the longer the interval for obtaining feedback from students about their educational experiences, the more likely it is that their responses will be vulnerable to forgetfulness. This may be simply because of the passage of time, but there is a particular problem for those graduates who immediately enrol on further programmes of study. In general, it is unreasonable to seek feedback on one educational activity when the students are involved in a subsequent activity, as the students will find it progressively difficult to separate their experiences of the two activities.

Would a single questionnaire be suitable for all students?

3.60 Experience with the SEEQ and the CEQ in America and Australia suggests that it is feasible to construct questionnaires that have a very wide range of applicability. The results have been used to make meaningful comparisons across a variety of institutions and a variety of disciplines. In addition, many institutions that use the SEEQ to obtain feedback from students about teachers or course units and institutions that use the CEQ to obtain feedback from recent graduates about their programmes seem to accept these surveys as entirely sufficient sources of information and do not attempt to supplement them with other instruments. This suggests that the use of a single national questionnaire to survey recent graduates would largely supplant other instruments that are currently used for this purpose by individual institutions. Instead, they might be induced to focus their efforts elsewhere (for instance, through more extensive surveys of current students).

3.61 It is clearly necessary that such a questionnaire should be motivated by research evidence about teaching, learning and assessment in higher education and that it should be properly assessed as a research tool. The only existing instruments that satisfy these requirements are the SEEQ (for evaluating individual teachers and course units) and the CEQ (for evaluating programmes). It has been argued that instruments like the SEEQ take for granted a didactic model of teaching, and this may be true of any questionnaire that focuses on the role of the teacher at the expense of the learner. Conversely, course designers who adopt student-centred curricula may find that these instruments are unhelpful as evaluative tools (Kember et al., 2002). In a similar manner, Lyon and Hendry (2002) claimed that the CEQ was not appropriate for evaluating programmes with problem-based curricula. However, their results may have been due not to inadequacies of the CEQ but to difficulties in introducing problem-based learning (Hendry, Cumming, Lyon, & Gordon, 2001), and the CEQ has been successfully used with other problem-based programmes.

3.62 In the GCCA surveys, the CEQ seems to be appropriate for assessing the experience of students on both undergraduate and postgraduate programmes. (Students taking joint degree programmes are asked to provide responses for each of their disciplines separately.) However, it does not seem to be useful for assessing the experiences of students working for postgraduate research degrees, and no suitable alternative has yet been devised. It may prove necessary to evaluate the quality of postgraduate research training using a quite different methodology from the CEQ. In distance education, it has proved necessary to amend the wording of many of the items in the CEQ, and the constituent structure of the resulting questionnaire reflects the different roles of staff as the authors of course materials and as tutors. The wording and the structure of any instrument that was adopted for use in a national survey of graduates would have to accommodate the different practices in campus-based and

distance education. More generally, it would have to be able to accommodate any variations in practice in higher education that might arise in the future.

Why are response rates important?

3.63 Some might argue that the purpose of feedback surveys was simply to provide students with an opportunity to comment on their educational experience. On this argument, students who do not respond do not cause any difficulty because they have chosen not to contribute to this exercise. Nevertheless, most researchers assume that the purpose of feedback surveys is to investigate the experience of all the students in question, and in this case those who do not respond constitute a serious difficulty insofar as any conclusions have to be based on data contributed by a sample.

3.64 Inferences based upon samples may be inaccurate for two reasons: sampling error and sampling bias. Sampling error arises because, even if a sample is chosen entirely at random, properties of the sample will differ by chance from those of the population from which the sample has been drawn. In surveys, questionnaire responses generated by a sample will differ from those that would be generated by the entire population. The magnitude of the sampling error is reduced if the size of the sample is increased, and so efforts should be made to maximise the response rate.

3.65 Sampling bias arises when a sample is not chosen at random from the relevant population. As a result, the properties of the sample may be misleading estimates of the corresponding properties of the population as a whole. In surveys, sampling bias arises if relevant characteristics of the people who respond are systematically different from those of the people who do not respond, in which case the results may be at variance with those that would have been found if responses had been obtained from the entire population.

3.66 Research has shown that students who respond to surveys are different from students who do not respond in terms of demographic characteristics, study behaviour and academic attainment (Astin, 1970; Nielsen, Moos, & Lee, 1978; Watkins & Hattie, 1985). It is therefore reasonable to expect that students who respond to feedback surveys will be systematically different from those who do not respond to such surveys in their educational experience. This kind of bias is unavoidable, but its impact can be reduced by minimising the number of non-respondents.

3.67 In social research, a response rate of 50% is considered satisfactory for a postal survey (Babbie, 1973, p. 165; Kidder, 1981, pp. 150-151). As mentioned earlier, the Australian GCCA surveys require that this response rate be achieved by individual institutions if their

average ratings are to be published. Indeed, the vast majority of participating institutions do achieve this response rate, and at a national level the GCCA surveys regularly achieve response rates of around 60%. In other words, this is the kind of response rate that can be achieved in a well-designed postal survey, although it clearly leaves ample opportunity for sampling bias to affect the results. The position of the Australian Vice-Chancellors' Committee (2001) is that an overall institutional response rate for the CEQ of at least 70% is both desirable and achievable.

3.68 Student feedback at the end of course units is often collected in a class situation, and this could be used to obtain feedback at the end of entire programmes (in both cases, presumably, before the assessment results are known). This is likely to yield much higher response rates and hence to reduce the impact of sampling error and sampling bias. There is an ethical issue as to whether students should be required to contribute feedback in this manner. In a class situation, students might feel under pressure to participate in the process, but the guidelines of many professional bodies stipulate that participants should be able to withdraw from a research study at any time. It will be important for institutions to clarify whether the collection of feedback is a formal part of the teaching-learning process or whether it is simply tantamount to institutional research.

3.69 With the increasing use of information technology in higher education, institutions may rely less on classroom teaching and more upon electronic forms of communication. This is already the case in distance learning, where electronic means of course delivery are rapidly replacing more traditional correspondence methods. Information technology can also provide a very effective method of administering social surveys, including the direct electronic recording of responses (see Watt, Simpson, McKillop, & Nunn, 2002). It would be sensible to administer feedback surveys by the same mode as that used for delivering the curriculum (classroom administration for face-to-face teaching, postal surveys for correspondence courses and electronic surveys for on-line courses). Little is known about the response rates obtained in electronic surveys, or whether different modes of administration yield similar patterns of results. Nevertheless, it is both good practice and arguably a legal requirement under the 2001 Special Education Needs and Disability Act (SENDA) to make feedback questionnaires available in various formats for use by students with disabilities.

3.70 To achieve high response rates, it is clearly necessary to ensure the cooperation and motivation of the relevant population of students. Those who have satisfactorily completed a course unit or an entire programme may be disposed to complete feedback questionnaires, but this may not be the case for students who have failed and particularly for those who have withdrawn from their studies for academic reasons. At the Open University, students who drop out of course units are automatically sent a questionnaire to investigate the reasons for

their withdrawal. This provides useful information, but the response rates are typically of the order of 25%. One could therefore not be confident that the data were representative of students who withdraw from course units.

How seriously is student feedback taken?

3.71 It is often assumed that the publication of student feedback will help students to make decisions about the choice of programmes and course units, that it will help teachers to enhance their own professional skills and that it will help institutions and funding bodies to manage their resources more effectively. None of these assumptions has been confirmed by empirical research, though it should be noted that most of the evidence relates to the use that is (or is not) made of SETs.

3.72 There have been consistent findings that students believe SETs to be accurate and important, although they constitute only one of the sources of information that students use when choosing between different course units (Babad, 2001). However, students may be sceptical as to whether attention is paid to the results either by the teachers being assessed or by senior staff responsible for appointments, appraisal or promotions, because they perceive that teachers and institutions attach more importance to research than to teaching. Indeed, unless students can see that the expression of their opinions leads to concrete changes in teaching practices, they may make little use of their own ratings (Spencer & Schmelkin, 2002). However, the development needs that students ascribe to their teachers may be driven by a didactic model of teaching and may differ from the teachers' own perceived needs (Ballantyne, Borthwick, & Packer, 2000).

3.73 From the teachers' perspective, the situation is a similar one. In the past, some resistance to the use of student ratings has been expressed, based on the ideas that students are not competent to make such judgements or that student ratings are influenced by teachers' popularity rather than their effectiveness. Both sociability and competence contribute to the idea of an 'ideal teacher' (Pozo-Muñoz, Rebolloso-Pacheco, & Fernández-Ramirez, 2000), but most teachers do consider SETs to be useful sources of information (Schmelkin, Spencer, & Gellman, 1997). Left to their own devices, however, they may be unlikely to change their teaching in the light of the results, to make the results available for other students, to discuss them with more senior members of staff, or to refer them to institutional committees or administrators (Nasser & Fresko, 2002).

3.74 Even in institutions where the collection of student feedback is compulsory, teachers may make little attempt to make use of the information that it contains. Once again, this may be because institutions are perceived to attach more importance to research than to teaching,

despite having formal policies that implicate teaching quality in decisions about staff appointments, appraisal and promotions (Kember et al., 2002). There seems to be no published research evidence on the use that senior managers of institutions make or do not make of student feedback in such cases, but there are four main reasons for the apparent lack of attention to this kind of information.

3.75 The first reason is the lack of guidance to teachers, managers and administrators on how such information should be interpreted. In the absence of such guidance, there is little or no scope for any sensible discussion about the findings. Potential users of student feedback need to be helped to understand and contextualise the results (Neumann, 2000). The second reason is the lack of external incentives to make use of such information. In the absence of explicit rewards for good feedback or explicit penalties for poor feedback (or at least for not acting upon such feedback), it is rational for both teachers and students to infer that their institutions do not take the quality of teaching seriously and value other kinds of activities such as research (Kember et al., 2002).

3.76 A third point is that the results need to be published to assure students that action is being taken, although care should also be taken that to ensure they are not misinterpreted or misrepresented. The Australian Vice-Chancellors' Committee (2001) issued a code of practice on the release of CEQ data. This cautions against making simplistic comparisons between institutions (because of variations in the student populations at different institutions); aggregating the results from different disciplines to an institutional level (because of variations in the mix of disciplines at different institutions); and attaching undue importance to trivial differences in CEQ scores. In the United Kingdom, it would arguably be appropriate to report student feedback data for each institution in the 19 broad subject groupings (JACS) used by HESA.

3.77 The final reason for the lack of attention to student feedback is the under-researched issue of the ownership of feedback data. Teachers may be less disposed to act on the findings of feedback, and students may be more disposed to be sceptical about the value of providing feedback to the extent that it appears to be divorced from the immediate context of teaching and learning. This is more likely to be the case if student feedback is collected, analysed and published by their institution's central administration and even more so if it is collected, analysed and published by an impersonal agency that is wholly external to their institution. The collection of feedback concerning programmes or institutions for quality assurance purposes certainly does not reduce the need to obtain feedback concerning teachers or course units for developmental purposes.

Conclusions

- student feedback provides important evidence for assessing quality; it can be used to support attempts to improve quality; and it can be useful to prospective students

- the use of quantitative instruments is dictated by organisational constraints (and in distance education by geographical constraints, too)

- feedback should be sought at the level at which one is endeavouring to monitor quality

- the focus should be on students' perceptions of key aspects of teaching or on key aspects of the quality of their programmes

- feedback should be collected as soon as possible after the relevant educational activity

- it is feasible to construct questionnaires with a wide range of applicability. Two groups are problematic: postgraduate research students and distance-learning students. Curricular innovations might make it necessary to reword or more radically amend existing instruments

- response rates of 60% of more are both desirable and achievable for students who have satisfactorily completed their course units or programmes. Response rates may well be lower for students who have failed or who have withdrawn from their course units or programmes

- many students and teachers believe that student feedback is useful and informative, but many teachers and institutions do not take student feedback sufficiently seriously. The main issues are: the interpretation of feedback; institutional reward structures; the publication of feedback; and a sense of ownership of feedback on the part of both teachers and students.

Psychometric properties of inventories and questionnaires

Reliability

3.78 The most fundamental requirement of a research instrument is that it should be *reliable* in the sense that it would yield consistent results if used repeatedly under the same conditions with the same participants and is therefore relatively unaffected by errors of measurement. This can be measured by a number of different coefficients of reliability, all of which vary in principle between zero (reflecting total *un*reliability) or one (reflecting perfect reliability). (In practice, instruments of poor reliability may actually yield estimates that are less than zero.)

3.79 One such measure is *test-retest reliability*: this involves calculating the correlation coefficients between the scores obtained by the same individuals on successive administrations of the same instrument. However, this suffers from two kinds of problem. With relatively short intervals between the two administrations, the participants will become familiar with the instrument and may even recall the responses that they gave at the first administration; as a result, its test-retest reliability may be spuriously high. This problem can be ameliorated by constructing equivalent or parallel forms of the same instrument for administration on the different occasions, but this is not a solution that has been adopted in the case of student feedback questionnaires.

3.80 In contrast, with relatively long intervals between two administrations of the same instrument, the participants are more likely to be exposed to contextual influences that lead to changes in the personal qualities being measured; as a result, the instrument's test-retest reliability may be spuriously low. In this situation, the correlation coefficient between the scores obtained at the two administrations is more a measure of its stability than its reliability, and variability in the scores obtained on different occasions need not cast doubt on the adequacy of the instrument. Moreover, longitudinal studies of this sort are hard to carry out because of the high probability of attrition: the participants may decline to participate in the follow-up session, or they may no longer be available for inclusion (for instance, in the case of students who have withdrawn from their studies in the interim). As a result, the participants who contribute data from the follow-up session may be unrepresentative of the original sample.

3.81 An alternative approach is to estimate an instrument's reliability by examining the consistency between the scores obtained on its constituent parts at a single administration. (It is also clearly less arduous to administer an instrument on a single occasion than on two separate occasions.) One such measure is *split-half* reliability: the items are divided into two distinct subsets, and a correlation coefficient is calculated between the scores obtained on the

two halves. (Formally, this is similar to comparing parallel forms of an entire instrument.) For instance, in evaluating tests of ability, one might compare the total score obtained on the odd-numbered items and the total scores obtained on the even-numbered items. However, this may not be appropriate when evaluating student feedback questionnaires and other instruments measuring attitudes.

3.82 The most common measure of reliability is Cronbach's (1951) coefficient alpha. This estimates the internal consistency of an instrument by comparing the variance of the total scores with the variances of the scores on the constituent items. It is formally equivalent to the average value of split-half reliability across all the possible ways of dividing the items into two distinct subsets. This is generally felt to be a useful indicator of the reliability of a test instrument, although low values of internal consistency may arise *either* because an instrument is unreliable *or* because it is not measuring a single personal quality or trait. Other procedures such as factor analysis or Rasch's (1960) measurement analysis would be needed to explore which of these was the case (see below). Moreover, it is not widely appreciated that this measure is itself subject to error and variability from one sample of participants to another (see Fan & Thompson, 2001).

3.83 Finally, when an instrument is used to obtain assessments of a particular individual by a number of different judges, it is appropriate to ask whether the judges are consistent in their evaluations. This is referred to as *interrater reliability*. The interrater reliability of the *average* rating of the same individual given by a group of judges increases with the number of judges in the group.

Validity

3.84 The other fundamental requirement of a research instrument is that it should be *valid* in the sense that it measures the personal qualities or traits that it purports to measure. This can be judged in a number of different ways: some depend upon the properties of the instrument itself; some depend upon the relationships between the scores on different items; and others depend upon the relationships between scores on the instrument itself and scores on other measures.

3.85 One approach to assessing the validity of an instrument is to examine the wording or structure of the constituent items. This might be carried out at a relatively superficial level, simply by asking whether the contents of the instrument appear to be appropriate; this is known as *face validity*. On the other hand, it might be carried out by a more thorough process of analysis and comparison of the items, and this is known as *content validity*. Both techniques are

limited in so far as they rely upon subjective and qualitative judgements rather than objective procedures.

3.86 Another approach is to examine the relationships between the scores obtained by a sample of participants on the constituent parts of an instrument. This is known as *construct validity* and is usually addressed by means of factor analysis. This can provide evidence that the instrument measures one or more distinctive traits or constructs. Rasch's (1960) measurement analysis can also be used where the constituent items are assumed to measure a single construct. A different use of factor analysis is to examine the relationships between the scores on the constituent parts of one instrument and the scores obtained by the same participants on other instruments. This can provide evidence of an instrument's *convergent validity*: that it is measuring the same traits that are being measured by the other instruments. In both applications, factor analysis employs formal statistical procedures but also relies upon an element of subjective interpretation.

3.87 A further approach is to examine the correlations between the scores on an instrument and the scores obtained on some independent criterion. This is known as criterion (or criterion-related) validity and yields coefficients varying between zero (reflecting a total lack of validity) and one (reflecting perfect validity). However, the value of any coefficient of validity will be limited by the reliability both of the instrument itself and of the relevant criterion. (In practice, once again, instruments of poor validity may actually yield estimates that are less than zero.) The criterion may be measured at the same time as the instrument is administered (concurrent validity), or it may be measured at some later point, so that the instrument is essentially being used to predict the criterion in question (predictive validity). However, the fact that scores on an instrument can be used to predict some outcome measure does not necessarily mean that the trait or traits being measured by the instrument are causally responsible for the observed outcome.

3.88 A related form of validity is *discriminative validity*: the extent to which an instrument yields different scores on groups of participants who would be expected to differ in the underlying trait or traits. In the case of students, the groups might differ on demographic characteristics (such as age, gender or educational background) or on contextual characteristics (such as their academic discipline, department or institution).

References

Abrami, P. C., & d'Apollonia, S. (1991). Multidimensional students' evaluations of teaching effectiveness—Generalizability of '*N* = 1' research: Comment on Marsh (1991). *Journal of Educational Psychology, 83,* 411-415.

Abrami, P. C., d'Apollonia, S., & Rosenfield, S. (1996). The dimensionality of student ratings of instruction: What we know and what we do not. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research*, vol. 11 (pp. 213-264). New York: Agathon Press.

Ainley, J. (1999, August). *Using the Course Experience Questionnaire to draw inferences about higher education*. Paper presented at the conference of the European Association for Research on Learning and Instruction, Göteborg, Sweden.

Ainley, J., & Long, M. (1994). *The course experience survey 1992 graduates*. Canberra: Australian Government Publishing Service.

Ainley, J., & Long, M. (1995). *The 1994 Course Experience Questionnaire: A report prepared for the Graduate Careers Council of Australia*. Parkville, Victoria: Graduate Careers Council of Australia.

Arubayi, E. A. (1987). Improvement of instruction and teacher effectiveness: Are student ratings reliable and valid? *Higher Education, 16,* 267-278.

Astin, A. W. (1970). The methodology of research on college impact, Part Two. *Sociology of Education, 43,* 437-450.

Australian Vice-Chancellors' Committee & Graduate Careers Council of Australia (2001). *Code of practice on the public disclosure of data from the Graduate Careers Council of Australia's Graduate Destination Survey, Course Experience Questionnaire and Postgraduate Research Experience Questionnaire*. Canberra: Australian Vice-Chancellors' Committee. Retrieved November 28, 2002, from http://www.avcc.edu.au

Babad, E. (2001). Students' course selection: Differential considerations for first and last course. *Research in Higher Education, 42,* 469-492.

Babbie, E. R. (1973). *Survey research methods.* Belmont, CA: Wadsworth.

Ballantyne, R., Borthwick, J., & Packer, J. (2000). Beyond student evaluation of teaching: Identifying and addressing academic staff development needs. *Assessment and Evaluation in Higher Education, 25,* 221-236.

Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development.* San Francisco: Jossey-Bass.

Cheng, D. X. (2001). Assessing student collegiate experience: Where do we begin? *Assessment and Evaluation in Higher Education, 26,* 525-538.

Chiang, K.-H. (2002, December 11). *Relationship between research and teaching in doctoral education in UK universities*. Unpublished paper presented at the Annual Conference of the Society for Research into Higher Education, University of Glasgow.

Clarkson, P. C. (1984). Papua New Guinea students' perceptions of mathematics lecturers. *Journal of Educational Psychology, 76,* 1386-1395.

Coffey, M. and Gibbs, G. (2000). 'Can academics benefit from training? Some preliminary evidence', *Teaching in Higher Education* 5, 385–389.

Coffey, M., & Gibbs, G. (2001). The evaluation of the Student Evaluation of Educational Quality Questionnaire (SEEQ) in UK higher education. *Assessment and Evaluation in Higher Education, 26,* 89-93.

Coffey, M. and Gibbs, G. (in press). 'New teachers' approaches to teaching and the utility of the Approaches to Teaching Inventory', *Higher Education Research and Development*.

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research, 51,* 281-309.

Committee of Vice-Chancellors and Principals (1998). *Skills development in higher education: A short report*. London: Committee of Vice-Chancellors and Principals.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297-334.

D'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist, 52,* 1198-1208.

Eley, M. G. (1992). Differential adoption of study approaches within individual students. *Higher Education, 23,* 231-254.

Fan, X., & Thompson, B. (2001). Confidence intervals about score reliability coefficients, please: An *EPM* guidelines editorial. *Educational and Psychological Measurement, 61,* 517-531.

Fung, Y., & Carr, R. (2000). Face-to-face tutorials in a distance learning system: Meeting student needs. *Open Learning, 15,* 35-46.

Gibbs, G., & Coffey, M. (2001). *Developing an associate lecturer student feedback questionnaire: Evidence and issues from the literature* (Student Support Research Group Report No. 1). Milton Keynes: The Open University.

Gibbs, G., Habeshaw, S., & Habeshaw, T. (1988). *53 interesting ways to appraise your teaching*. Bristol: Technical & Educational Services.

Gibbs, G., & Lucas, L. (1996). Using research to improve student learning in large classes. In G. Gibbs (Ed.), *Improving student learning: Using research to improve student learning* (pp. 33-49). Oxford: Oxford Centre for Staff Development.

Graduate satisfaction (2001, December). *GradStats*, No. 6. Retrieved November 27, 2002, from http://www.gradlink.edu.au/gradlink/gcca/GStats2001.pdf

Greenwald, A. G., & Gilmore, G. M. (1997a). Grading leniency is a removable contaminant of student ratings. *American Psychologist, 52,* 1209-1217.

Greenwald, A. G., & Gilmore, G. M. (1997b). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology, 89,* 743-751.

Gregory, R., Harland, G., & Thorley, L. (1995). Using a student experience questionnaire for improving teaching and learning. In G. Gibbs (Ed.), *Improving student learning through assessment and evaluation* (pp. 210-216). Oxford: Oxford Centre for Staff Development.

Gregory, R., Thorley, L., & Harland, G. (1994). Using a standard student experience questionnaire with engineering students: Initial results. In G. Gibbs (Ed.), *Improving student learning: Theory and practice* (pp. 529-538). Oxford: Oxford Centre for Staff Development.

Harvey, L. (2001). *Student feedback: A report to the Higher Education Funding Council for England*. Birmingham: University of Central England, Centre for Research into Quality.

Harvey, L. (with Geall, V., Mazelan, P., Moon, S., & Plummer, L.). (1995). *Student satisfaction: The 1995 report on the student experience at UCE*. Birmingham: University of Central England, Centre for Research into Quality. (ERIC Document Reproduction Service No. ED388166)

Harvey, L. (with Plimmer, L., Moon, S., & Geall, V.). (1997). *Student satisfaction manual*. Buckingham: SRHE & Open University Press.

Hativa, N. (1996). University instructors' ratings profiles: Stability over time, and disciplinary differences. *Research in Higher Education, 37,* 341-365.

Hendry, G. D., Cumming, R. G., Lyon, P. M., & Gordon, J. (2001). Student-centred course evaluation in a four-year, problem based medical programme: Issues in collection and management of feedback. *Assessment and Evaluation in Higher Education, 26,* 327-339.

Hennessy, S., Flude, M., & Tait, A. (1999). *An investigation of students' and tutors' views on tutorial provision: Overall findings of the RTS Project (Phases I and II)*. Milton Keynes: The Open University, School of Education.

Husbands, C. T., & Fosh, P. (1993). Students' evaluation of teaching in higher education: Experiences from four European countries and some implications of the practice. *Assessment and Evaluation in Higher Education, 18,* 95-114.

Johnson, T. (1997). *The 1996 Course Experience Questionnaire: A report prepared for the Graduate Careers Council of Australia*. Parkville, Victoria: Graduate Careers Council of Australia.

Johnson, T. (1998). *The 1997 Course Experience Questionnaire: A report prepared for the Graduate Careers Council of Australia*. Parkville, Victoria: Graduate Careers Council of Australia.

Johnson, T. (1999). *Course Experience Questionnaire 1998: A report prepared for the Graduate Careers Council of Australia*. Parkville, Victoria: Graduate Careers Council of Australia.

Johnson, T., Ainley, J., & Long, M. (1996). *The 1995 Course Experience Questionnaire: A report prepared for the Graduate Careers Council of Australia*. Parkville, Victoria: Graduate Careers Council of Australia.

Kahl, T. N., & Cropley, A. J. (1999). Face-to-face versus distance learning: Psychological consequences and practical implications. *Distance Education, 7,* 38-48.

Kember, D., Leung, D. Y. P., & Kwan, K. P. (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment and Evaluation in Higher Education, 27,* 411-425.

Kidder, L. H. (1981). *Research methods in social relations.* New York: Holt, Rinehart, & Winston.

Kolitch, E., & Dean, A. V. (1999). Student ratings of instruction in the USA: Hidden assumptions and missing conceptions about 'good' teaching. *Studies in Higher Education, 24,* 27-42.

Lawless, C. J., & Richardson, J. T. E. (2002). Approaches to studying and perceptions of academic quality in distance education. *Higher Education, 44,* 257-282.

Linke, R. D. (1991). *Performance indicators in higher education: Report of a trial evaluation study commissioned by the Commonwealth Department of Employment, Education, and Training*. Canberra: Australian Government Publishing Service.

Lizzio, A., Wilson, K., & Simons, R. (2002). University students' perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education, 27,* 27-52.

Long, M., & Hillman, K. (2000). *Course Experience Questionnaire 1999: A report prepared for the Graduate Careers Council of Australia*. Parkville, Victoria: Graduate Careers Council of Australia.

Lucas, L., Gibbs, G., Hughes, S., Jones, O., & Wisker, G. (1997). A study of the effects of course design features on student learning in large classes at three institutions: A comparative study. In C. Rust & G. Gibbs (Eds.), *Improving student learning: Improving student learning through course design* (pp. 10-24). Oxford: Oxford Centre for Staff and Learning Development.

Lyon, P. M., & Hendry, G. D. (2002). The use of the Course Experience Questionnaire as a monitoring evaluation tool in a problem-based medical programme. *Assessment and Evaluation in Higher Education, 27*, 339-352.

Marsh, H. W. (1981). Students' evaluations of tertiary instruction: Testing the applicability of American surveys in an Australian setting. *Australian Journal of Education, 25*, 177-192.

Marsh, H. W. (1982). SEEQ: A reliable, valid and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology, 52*, 77-95.

Marsh, H. W. (1983). Multidimensional ratings of teaching effectiveness by students from different academic settings and their relation to student/course/instructor characteristics. *Journal of Educational Psychology, 75*, 150-166.

Marsh, H. W. (1986). Applicability paradigm: Students' evaluations of teaching effectiveness in different countries. *Journal of Educational Psychology, 78*, 465-473.

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*, 253-388.

Marsh, H. W. (1991). Multidimensional students' evaluations of teaching effectiveness: A test of alternative higher-order structures. *Journal of Educational Psychology, 83*, 285-296.

Marsh, H. W., & Bailey, M. (1993). Multidimensional students' evaluations of teaching effectiveness. *Journal of Higher Education, 64*, 1-18.

Marsh, H. W., & Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional perspective. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research*, vol. 8 (pp. 143-233). New York: Agathon Press.

Marsh, H. W., & Hocevar, D. (1991a). Multidimensional perspective on students' evaluations of teaching effectiveness: The generality of factor structures across academic discipline, instructor level, and course level. *Teaching and Teacher Education, 7*, 9-18.

Marsh, H. W., & Hocevar, D. (1991b). Students' evaluations of teaching effectiveness: The stability of mean ratings of the same teachers over a 13-year period. *Teaching and Teacher Education, 7*, 303-314.

Marsh, H. W., & Roche, L. A. (1992). The use of student evaluations of university teaching in different settings: The applicability paradigm. *Australian Journal of Education, 36*, 278-300.

Marsh, H. W., & Roche, L. A. (1994). *The use of students' evaluations of university teaching to improve teaching effectiveness*. Canberra: Department of Employment, Education, and Training. Retrieved November 28, 2002, from http://fistserv.uws.edu.au/seeq/Report/seeq1to5.htm

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist, 52*, 1187-1197.

Marsh, H. W., Rowe, K. J., & Martin, A. (1999, December 13). *Postgraduate students' evaluations of research supervision: Benchmarking universities*. Unpublished paper presented to the Postgraduate

Research Experience Questionnaire (PREQ) Seminar, University of Melbourne. Retrieved November 28, 2002, from http://edweb.uws.edu.au/self/preq.htm

Marsh, H. W., Touron, J., & Wheeler, B. (1985). Students' evaluations of university instructors: The applicability of American instruments in a Spanish setting. *Teaching and Teacher Education, 1,* 123-138.

McInnis, C., Griffin, P., James, R., & Coates, H. (2001). *Development of the Course Experience Questionnaire (CEQ)*. Canberra: Department of Education, Training and Youth Affairs.

Meyer, J. H. F., & Muller, M. W. (1990). Evaluating the quality of student learning: I. An unfolding analysis of the association between perceptions of learning context and approaches to studying at an individual level. *Studies in Higher Education, 15,* 131-154.

Moore, M. G. (1980). Independent study. In R. D. Boyd, J. W. Apps, & Associates, *Redefining the discipline of adult education* (pp. 16-31). San Francisco: Jossey-Bass.

Morgan, A. (1984). A report on qualitative methodologies in research in distance education. *Distance Education, 5,* 252-267.

Murray, H. G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. *Journal of Educational Psychology, 75,* 138-149.

Narasimhan, K. (2001). Improving the climate of teaching sessions: The use of evaluations by students and instructors. *Quality in Higher Education, 7,* 179-190.

Nasser, F., & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment and Evaluation in Higher Education, 27,* 187-198.

Neumann, R. (2000). Communicating student evaluation of teaching results: Rating Interpretation Guides (RIGs). *Assessment and Evaluation in Higher Education, 25,* 121-134.

Nielsen, H. D., Moos, R. H., & Lee, E. A. (1978). Response bias in follow-up studies of college students. *Research in Higher Education, 9,* 97-113.

Parsons, P. G. (1988). The Lancaster Approaches to Studying Inventory and Course Perceptions Questionnaire: A replicated study at the Cape Technikon. *South African Journal of Higher Education, 2,* 103-111.

Pearson, M., Kayrooz, C., & Collins, R. (2002, December 11). *Postgraduate student feedback on research supervisory practice*. Unpublished paper presented at the Annual Conference of the Society for Research into Higher Education, University of Glasgow.

Perry, W. G., Jr. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart, & Winston.

Pozo-Muñoz, C., Rebolloso-Pacheco, E., & Fernández-Ramirez, B. (2000). The 'ideal teacher': Implications for student evaluation of teacher effectiveness. *Assessment and Evaluation in Higher Education, 25,* 253-263.

Prosser, M., Trigwell, K., Hazel, E., & Gallagher, P. (1994). Students' experiences of teaching and learning at the topic level. *Research and Development in Higher Education, 16,* 305-310.

Purcell, K. & Pitcher, J. (1998) *Great expectations: A new diversity of graduate skills and aspirations* (Manchester: Higher Education Careers Services Unit).

Ramsden, P. (1991a). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education, 16,* 129-150.

Ramsden, P. (1991b). Report on the Course Experience Questionnaire Trial. In R. D. Linke, *Performance indicators in higher education: Report of a trial evaluation study commissioned by the Commonwealth Department of Employment, Education, and Training. Volume 2: Supplementary Paper* (pp. 1-85). Canberra: Australian Government Publishing Service.

Ramsden, P., & Entwistle, N. J. (1981). Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology, 51,* 368-383.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Denmarks paedagogiske Institut.

Richardson, J. T. E. (1994). A British evaluation of the Course Experience Questionnaire. *Studies in Higher Education, 19,* 59-68.

Richardson, J. T. E. (1997). Students' evaluations of teaching: The Course Experience Questionnaire. *Psychology Teaching Review, 6,* 31-45.

Richardson, J. T. E., & Price, L. (2003). Approaches to studying and perceptions of academic quality in electronically delivered courses. *British Journal of Educational Technology, 34,* 45-56.

Richardson, J. T. E., & Woodley, A. (2001). Perceptions of academic quality among students with a hearing loss in distance education. *Journal of Educational Psychology, 93,* 563-570.

Roche, L. A., & Marsh, H. W. (2002). Teaching self-concept in higher education: Reflecting on multiple dimensions of teaching effectiveness. In N. Hativa & P. Goodyear (Eds.), *Teacher thinking, beliefs and knowledge in higher education* (pp. 179-218). Dordrecht: Kluwer.

Sadlo, G. (1997). Problem-based learning enhances the educational experiences of occupational therapy students. *Education for Health, 10,* 101-114.

Schmelkin, L. P., Spencer, K. J., & Gillman, E. S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education, 38,* 575-592.

Spencer, K. J., & Schmelkin, L. P. (2002). Student perspectives on teaching and its evaluation. *Assessment and Evaluation in Higher Education, 27,* 397-409.

Trigwell, K., & Ashwin, P. (2002, September). *Evoked conceptions of learning and learning environments*. Paper presented at the 10th International Improving Student Learning Symposium, Brussels.

Trigwell, K., & Prosser, M. (1991). Improving the quality of student learning: The influence of learning context and student approaches to learning on learning outcomes. *Higher Education, 22,* 251-266.

Watkins, D., & Hattie, J. (1985). A longitudinal study of the approaches to learning of Australian tertiary students. *Human Learning, 4,* 127-141.

Watkins, D., Marsh, H. W., & Young, D. (1987). Evaluating tertiary teaching. A New Zealand perspective. *Teaching and Teacher Education, 2,* 41-53.

Watt, S., Simpson, C., McKillop, C., & Nunn, V. (2002). Electronic course surveys: Does automating feedback and reporting give better results? *Assessment and Evaluation in Higher Education, 27,* 325-337.

Waugh, R. F. (1998). The Course Experience Questionnaire: A Rasch measurement model analysis. *Higher Education Research and Development, 17,* 45-64.

Wiers-Jenssen, J., Stensaker, B., & Grøgaard, J. B. (2002). Student satisfaction: Towards an empirical deconstruction of the concept. *Quality in Higher Education, 8,* 183-195.

Wilson, K. L., Lizzio, A., & Ramsden, P. (1997). The development, validation and application of the Course Experience Questionnaire. *Studies in Higher Education, 22*, 33-53.

4 Institutional processes for collecting and using student feedback

Introduction

4.1 The terms of reference for this strand of the project were to:

Review current good practice by HEIs in collecting quantitative and qualitative feedback from students on the quality and standards of their higher education programmes, and using that feedback to secure improvement.

Make recommendations on how individual HEIs can best design and implement their own internal mechanisms for:

- *collecting quantitative and qualitative feedback data from current students*
- *following up that feedback to secure improvement and address students' concerns.*

4.2 There were three aspects to the work which fed into both strands of the project:

- a consultation with all HEFCE-funded HEIs on their internal student feedback processes and their views on a set of presumptions about good practice. The consultation resulted in a 60% response rate (81 responses). The questionnaire used in the consultation is reproduced as Appendix B
- a literature review (see Chapter 3)
- visits by the project team to 20 HEIs. The institutions were selected to represent the diversity of the higher education sector and took into account the size and type of institution ('old' and 'new' universities and specialist and non-specialist institutions) as well as a geographical spread. All of the 20 institutions visited had been among the institutions that had responded to the consultation. Meetings were held with senior managers, deans/heads of department, members of central administrative units responsible for student feedback (where they existed), and students and Student Unions. The institutions visited are listed in Appendix C.

4.3 There are a number of key inter-related issues to be considered in reviewing institutional processes for collecting and using student feedback:

- purposes
- mechanisms
- collection
- analysis and interpretation
- actions and decision-making
- presentation and publication
- dissemination to students.

4.4 The rest of this part of the report is structured in terms of the above issues. Each section reviews current practice in institutions and recommends ways in which practice might be improved in the future. The recommendations will be elaborated in the *Good Practice Guide*[4] to be published separately.

Purposes

What is meant by student feedback?

4.5 Student feedback can be defined broadly as obtaining information about:

- student satisfaction with specific programmes/units or services
- student views about whether their objectives have been met
- student accounts of their learning and study methods.

4.6 These are rather different. The third would imply the collection of largely descriptive behavioural information. The first and second require students to make evaluations. Neither

[4] The Higher Education Funding Council for England (HEFCE) has contracted the Centre for Higher Education Research and Information to produce a Good Practice Guide based on the project but separate to the report of the project. The Guide will draw upon the experience of the project in more detail and provide examples of the interesting practice identified.

separately nor collectively can these different facets of student feedback be equated with 'quality'.

Purposes as perceived by institutions

4.7 Among the main purposes of obtaining feedback cited by the institutions responding to the consultation are 'enhancing the students' experience of learning and teaching' and 'contributing to monitoring and review of quality and standards'. Other purposes cited include 'to ensure the effectiveness of course design and delivery', 'enabling a dialogue with students', 'helping students reflect on their experiences', 'as part of the teaching and learning process', 'identifying good practice', 'measuring student satisfaction' and 'contributing to staff development'. In many cases student feedback is used for a multiplicity of purposes.

4.8 Student feedback is regarded as playing an important role in institutional quality assurance; the overwhelming majority of HEIs stated that student feedback is central to their annual monitoring and periodic review processes. However, student feedback is just one source of information to support quality assurance processes, and it needs to be understood and interpreted in the context of other information (such as student profiles, progression rates, and external examiner reports). Indeed, a number of institutions cautioned against using student feedback as an indicator of quality and standards.

4.9 While most of the purposes cited imply a relationship with quality enhancement in some sense of that term, there may be need for greater precision of purpose on some occasions. These in part reflect the different levels at which information is intended to be used and in part the needs of different users. Both issues are considered in later sections.

A new purpose?

4.10 An important new purpose is the need to inform prospective students about the quality and standards of higher education programmes of study. The rationale emanates from the new national arrangements for quality assurance that require HEIs to make public certain information about their higher education provision, including the views of students about the quality and standards of their programmes. The QAA will also expect institutions to possess certain kinds of information (on which the QAA intends to place greater reliance) to support institutional quality assurance.

4.11 This somewhat new role for student feedback has many implications for HEIs, including:

- balancing market and quality considerations

- maintaining data integrity in new contexts of data use
- designing new instruments for collecting student feedback or the modification of existing ones
- developing new ways of presenting and summarising data.

4.12 These issues are dealt with in more detail in later sections of this report.

4.13 A related purpose is the 'internal' publication of student feedback to help inform current students about their choice of modules. This does not currently appear to be a very common practice. It raises a number of practical issues, not least the timing of collection and publication of feedback.

Levels

4.14 Student feedback is collected at a number of different levels – as became clear from our consultation and visits to HEIs. The use of student feedback at different institutional levels may reflect differences in purpose. Levels include:

- an individual teacher or class
- a module[5] or unit
- a semester or year of study
- a programme of study[6]
- a subject
- a department
- a faculty

[5] For the rest of this report, the term 'module' will be used to refer to the components or units of curricula structures that are unitised or modularised.

[6] Programme of study is used to cover the units, modules or courses that lead to an award. It is recognised that many students will be combining modules into programmes that lead to joint degrees.

- an institution.

4.15 The following examples (Figure 4.1) show the different *uses* that student feedback might have at different levels (e.g. module, programme, support service) and who might be interested in using it.

4.16 The use to which feedback will be put and the level at which it is collected and/or analysed will have implications for the timing of collection in relation to decision-making cycles. For example, feedback collected at the module level will be used by the module teacher(s) to check how things are going and to enable relatively immediate modifications, if necessary. This might imply collection at a mid-point in the delivery of the module. However, module feedback will also feed into the 'programme team', department or subject discussions about longer-term developments and improvements, as well as contributing to the monitoring and review process. For these purposes, feedback needs to be collected near to the end of the module and implies some comparison between modules in terms of the data collected. Especially where modules are semester length, the most common practice seems to be to administer end of module questionnaires. This means that the results cannot feed into improvements that will benefit the current students, and we found that this could affect students' commitment to the feedback process.

Figure 4.1: The uses and users of student feedback

Users of student feedback	Uses of student feedback	Level
A teacher	To improve teaching	Module
A programme team	To check that learning objectives have been met, to check coherence of a programme as a whole, and to improve the student learning experience in general	Programme, plus individual modules
A department or faculty	To help satisfy responsibilities for quality and standards and to help plan future provision	Programme
Institutional leadership and senior academic committees	To help satisfy responsibilities for quality and standards and to help plan future strategy	Programme, subject and institutional support services (i.e. institution-wide)
Current students	To inform module choices	Module
Prospective students	To inform choice within and between HEIs	Programme, subject

4.17 There is undoubtedly a potential for tension between the requirements of different users and different levels. As we argue in the section on 'collection', it is certainly the case that aggregating module feedback does not convert it into feedback on the student experience of the programme as a whole. It does, however, provide some information on the module 'set'

that constitutes the programme, and this is likely to be of value to those with responsibilities at this level. This might be regarded as preferable to collecting feedback separately at *both* module and programme levels. Moreover, module level feedback can be of use at the programme level in conjunction with other relevant data (i.e. student profiles, progression data, external examiners' reports, programme level student feedback). An alternative approach would be to conduct institution-wide surveys that could be disaggregated to programme level. We recognise that reference to 'programme' level is not applicable to all institutions and all forms of curriculum organisation. In some forms of modular organisation, no level beyond that of the individual module may be identifiable, although groupings can always be derived from the actual choices made by students.

Recommendations

4.18 Clarity of purpose is key:

- there is a need for clarity about the differences between student feedback on satisfaction, on learning processes (study methods), and on student objectives and their achievement. The dangers are that questionnaires muddle these different kinds of feedback, although it is not impossible to combine them into a single instrument

- all involved in the collection and use of student feedback data need to be clear about the purposes and intended uses of the data. This is especially important for the students themselves if their commitment to the process is to be maximised. Thus, the purposes and use to which the information will be put should be stated at the start of questionnaires

- consideration should be given to the range of alternative ways in which purposes can be achieved. For example, questionnaire fatigue among both students and staff is clearly a danger. There may be some potential for reducing the total burden by sampling or by collecting feedback in alternate years or only when other quality indicators have suggested cause for concern

- use of a range of feedback mechanisms will be more effective than reliance on questionnaires. For example, the existence of a well-publicised complaints procedure or a discussion during class may provide more effective ways of checking that a module is going well than asking students to fill out questionnaires

- the needs of users at different levels in the institution should be recognised at the time that data is collected. They will have implications for what is collected and when, and for what forms of aggregation might subsequently be required.

Mechanisms

4.19 The main mechanisms for collecting student feedback include:

- questionnaires at various levels
- student representation on local and institutional committees
- staff/student liaison committees
- other (e.g. discussion groups)
- during a lecture/seminar
- personal tutorial system
- informal.

Current practice

4.20 Most institutions operate a mix of mechanisms at various levels. Institutions appear to find qualitative and quantitative methods equally useful, although some institutions feel that qualitative feedback is by far the most useful form of feedback. The most commonly used methods of feedback are staff/student liaison committees (or their equivalent), student representation on committees and questionnaires. A minority mentioned discussion groups, tutorials and informal methods.

4.21 Student representation systems are universal, although their effectiveness appears to vary (see below). Many institutions run staff/student liaison committees comprising staff and students of a particular academic unit (a programme, department, faculty, school or subject). Questionnaire feedback (or student opinion surveys) is almost universal except in some small specialist institutions that rely on other and less formal forms of feedback. Questionnaires are not always popular with staff or students, but they may be inevitable in view of the decline in staff/student ratios, the growth of modular forms of curricula organisation, and other internal

and external pressures on staff that have combined to reduce the opportunities for informal face-to-face contacts with students. The increased need for 'hard data' for accountability purposes has also shifted feedback from informal to formal modes, and provides another reason for the resistance to questionnaires among some staff and students.

4.22 However, there remains something of a consensus that reliance cannot be placed on any single method of feedback. A variety of mechanisms will continue to be used and these will take account of factors such as the purposes, levels and available resources.

Recommendations

- the mechanisms used should take account of the form of curriculum organisation, including the length of modules, and the numbers of students enrolled on modules, on programmes, in departments and so on

- reliance should not be placed on any one mechanism for collecting and using feedback

- reliance on informal feedback, while important, is not recommended as a sole mechanism for obtaining student feedback

- mechanisms used will also need to relate to purpose, which should be clearly stated and communicated to staff and students. Although traditionally related to quality and enhancement, additional purposes need to be recognised and accommodated.

Collection

Current practice

4.23 The overwhelming majority of HEIs indicated in their responses to the consultation that institution-wide policies on the collection and use of student feedback have been established. (A number were being reviewed or revised and a few commented that this action was in light of the Cooke recommendations – HEFCE report 02/15.)

4.24 As we have already seen, the main levels at which student feedback is obtained and the purposes for which it is used are many. The most common level at which feedback is collected is the module, followed by the programme level. Many institutions collect feedback at both levels. However, the module is felt to be the most effective level for gathering and

using feedback because it is closest to the student experience and therefore most appropriate to ensuring fairly immediate improvements to the teaching and learning process. This is especially so in modular structures, where students can more easily locate their immediate frame of reference and experience with the module than with the programme.

4.25 The other main levels where student feedback is sought and used are at department/faculty/school level for the purposes of annual monitoring and review. Around a third of institutions responding to the consultation collect feedback at institutional level (and sometimes reported at the programme level) to gauge overall student satisfaction with institutional services. A minority indicated that feedback was collected at the level of the individual teacher, and an additional three specifically cited that such feedback is obtained and used for the purposes of staff appraisal, promotion and salary review.

4.26 In addition to information on the teaching and learning process, the majority of HEIs gather feedback on library and IT resources and facilities (sometimes as part of module/programme questionnaires and always as part of institution-wide surveys, although the purposes will, of course, differ). Around half mentioned that information was gathered on other services such as careers, counselling, admissions, induction, catering and accommodation. A minority (seven) specifically cited that information was gathered on the role of the Student Union.

4.27 Well over half the HEIs that collected student feedback at more than one level stated that the levels are related (e.g. responses are aggregated for the purposes of deciding what action, if any, to take). The main reasons stated by institutions are for preparing annual monitoring and review reports for the committee cycles. A minority said that the extent to which this is done or could be done was limited, and 15 respondents provided a negative response. Some institutions sounded cautionary notes about aggregating student feedback because they felt that information needs at module and programme levels are different. Therefore they argued that it is inappropriate to (numerically) aggregate module feedback, and the level at which information is gathered should relate to the level at which it is to be used and where authority is located.

4.28 However, it remains the case that if it is felt impossible or undesirable to aggregate module data then additional data will need to be collected to meet needs at higher levels in the institution.

4.29 Institutions were asked how students are able to influence the issues on which, and the method by which, feedback is sought. The main ways cited by HEIs are through staff/student liaison committees and the student representation system. (Varying views were obtained from our interviews with students as to the effectiveness of such arrangements.) Just over 20

institutions stressed the importance of the Student Union's role through its regular meetings with senior management and its training of student representatives. Other ways cited included through quality assurance processes and working groups, which had been set up to look at student feedback issues.

Specific issues relating to questionnaires

Level

4.30 The level(s) at which questionnaires are administered will be related to purpose. For the purposes of gathering information on student views about the quality and standards of teaching and learning, we have already indicated that most HEIs administer questionnaires at the module and/or the programme level (and often both), and to a lesser extent at institutional level. The majority of institutions feel strongly that the module is the most important level at which to capture student feedback because this is where the teaching and learning takes place and where immediate improvements can be made. The wider perspective/experience is also seen as beneficial (i.e. programme, institution levels), but there are concerns relating to 'generalisation' of the experience (especially in a modularised structure) and whether and how quickly action can be taken on the results (see below). A number of institutions have discontinued programme/institutional level questionnaires because of the low response rates and because the information generated was too broad to be useful. This division of view on the importance of different levels may reflect variations in the effectiveness of different levels of the quality assurance procedures within institutions. If it appears difficult to take action at the programme level, this may be indicative of problems in the quality assurance procedures at this level. It may also suggest that data have not been analysed or presented to be useful at this level (see the next section).

4.31 The consultation with HEIs included a set of presumptions about good practice, which were based on the work of the Task Group, and institutions were asked for their views. One of these presumptions relates to the issue of level:

For the purposes of reporting back the results of student opinion surveys within the HEI, it should be possible to disaggregate the results to the level of individual programme, because a primary purpose of getting the information is in order that the quality of individual programmes can be improved.

4.32 As mentioned above, most feedback through questionnaires (student opinion surveys) is gathered at the module level. Programme level is less common where modularisation has been established because students' frames of reference and experience are not so easily identified at that level, and the information collected cannot readily be related to structures

that are meaningful to staff. Thus, at the module level, it is an issue of *aggregation* rather than disaggregation.

4.33 Where institution-wide surveys are undertaken, these can often be disaggregated to programme level. However, we have noted that purpose and focus at this level will differ. Furthermore institutions may not wish to disaggregate to the programme level because of difficulties in identifying students on 'meaningful' programmes in a modularised curriculum, because the information at that level is not of interest to senior managers or staff, or because numbers are too small to make judgements.

4.34 If the need (whether it be an external or an internal one) is for programme level information, it is technically feasible to aggregate if common questions have been used. (Data may need to be weighted and the problem of defining programmes and programme boundaries overcome.) It needs to be clear, however, that such aggregation does not produce data on the student experience and coherence of the overall programme, but does indicate something of student views about the set of modules that constitute it. In certain circumstances, aggregation may be possible and useful but, in doing so, there are factors that will need to be taken into account.

4.35 It should also be recognised that module and institution-wide surveys have largely different purposes. Module is bottom-up teacher/student driven evaluation of the immediate teacher-learner interface. Institution-wide is much more of a management information instrument designed to provide an overview, with the possibility of disaggregating to programme level to provide external information on programmes and, for programme directors (or equivalent), a more holistic perspective on the student experience than could be obtained by assembling feedback from a set of modules.

4.36 The consultation with HEIs revealed that **over half agreed in principle with the above presumption** and a minority of institutions disagreed altogether. Many, including those who agreed in principle, expressed reservations or concerns. Some of these are indicated in Figure 4.2.

Figure 4.2: Institutional concerns about the good practice presumption regarding programme level feedback

Obtaining feedback at the module level is more important than the programme level and collection at this level will continue regardless of external pressures.
Programme level information is different from and less detailed than that collected for the module and therefore not useful for improvement purposes.
Where institutional level surveys are carried out, it would be inappropriate to disaggregate to programme level because of subject differences and sample sizes.
Alternatively, where module level feedback is the norm, to aggregate to programme level can be

administratively complex and resource intensive, and of questionable value.
Low response rates will create meaningless and questionable data.
In modular structures it is difficult to 'capture' students on a programme, hence low response rates.
Views will differ on what constitutes a programme.

4.37 However, programme level feedback may become more important as the use of programme specifications and student progress files increases.

4.38 The issue of level is clearly an important one and cannot be disassociated from issues of purpose. Much current practice favours concentration at the module level and this is the level that clearly has relevance to the individual teacher. But even if feedback is restricted to module level, a number of questions arise as to the information needs of programme teams or committees if they are to discharge responsibilities for quality and standards, including ensuring that actions are taken by individual staff members. Here the issue may be one of data reduction more than aggregation. If they are not to be overwhelmed with data at the module level, committees need to receive this data in a reduced form but which nevertheless enables them to reach judgements on the quality of teaching. In this way, comparison between modules will be possible and will aid in the interpretation of results and help to identify good practice. The question of aggregation of data to programme or other institutional level will in part be dependent on curriculum and organisational structures; although if enhanced uses of student feedback are to be achieved, ways will need to be found to make data meaningful at these other levels.

Questionnaire fatigue

4.39 A related issue to that of the level at which questionnaires are administered is that of questionnaire fatigue, a common complaint that is made by staff (for themselves) and for their students (although our interviews with students revealed a different point of view – see paragraph 4.46). As we have already mentioned, module level feedback is most common and popular for monitoring and improving the teaching and learning process. However, as we discovered, in a modularised system based on semesters, a student might be required to complete up to 12 module questionnaires. And if questionnaires are used in the way their purpose intends, there will also be an added burden on staff to process, analyse, interpret and take action based on the results. There was some evidence from our interviews with teaching staff and students that both completion and use of data from these questionnaires can easily become ritualistic for both parties. One solution might be to administer such questionnaires less frequently, for example every other time the module is offered (unless it is new or substantial changes have been introduced). Another solution would be to adopt a more holistic and integrated approach to feedback with a much smaller total number of

questionnaires administered but covering a broader area of student experience than the individual module.

Questions asked

4.40 Once again, purpose and level will determine the types of questions asked. The questionnaires we received (module and programme levels) as part of the consultation exercise varied in the types of questions and in their length, although questions tend to cover a similar range of aspects. These include organisation and content, teaching, learning and assessment practice, appropriateness of methods used, clarity of presentations and handouts, preparation of staff, workload, feedback, support and guidance provided, the learning environment, skills development, level of difficulty, and overall ratings about the module/programme. Many institutions try to keep questionnaires to two sides of A4 but many were double this length and sometimes more.

4.41 Of the examples provided, few questionnaires at the module and programme levels ask for details about the student. Thus, it is not possible to analyse the data according to such considerations as age or entry qualifications, other modules taken etc. Given the increased diversity of higher education, it is important to ascertain whether courses are equally successful for the different types of students who are taking them. Rather than collect this sort of data with each questionnaire, it might be possible to link student feedback data to other institutional datasets (although institutions will need to satisfy themselves about any issues concerning student anonymity and confidentiality). This is being explored in a number of institutions.

4.42 A further presumption of good practice related to the types of questions that students should be asked:

Surveys should include quantifiable 'tick-box' elements, capable of being analysed electronically. But they should not be limited to 'tick-boxes', but should give students opportunities to comment, expand and explain in their own words. This is valuable to enable the staff responsible for each programme to interpret and understand the results.

4.43 **The vast majority of institutions agreed with the presumption**. However, a minority of institutions (seven) highlighted the relative costs of collecting, analysing and reporting on fixed choice and open-ended elements. Some smaller/specialist HEIs made related but contradictory statements: two felt that open-ended questions and face-to-face contact were more important, and that fixed choice elements produce invalid results because students are being asked to comment on complex situations that cannot be reduced to tick boxes. Another institution, however, commented that it would only use fixed choice elements because of administrative and resource issues. In practice, the inclusion of a few open-ended questions

need not add substantially to the burden of analysis if they are briefly surveyed to ascertain whether new issues are being raised.

Response rates

4.44 The kinds of response rates normally obtained at different levels varied greatly from institution to institution and by level. The module level was by far the most effective (ranging from 60-100%), although there are issues relating to the 'quality' of responses, which are discussed below. There was some evidence in the responses provided by HEIs that institution/programme level questionnaires have a much lower response rate (between 20-30%) than those administered at other levels. This is not altogether surprising because questionnaires at module level are more focussed and therefore relevant to the immediate student experience, and tend to be paper-based and administered 'in-class', which ensures higher response rates. However, questionnaires administered 'in-class' will be completed only by those students who attend the lecture or seminar and who may not be typical. A number of HEIs reported that they were experimenting with questionnaires on-line; this was normally associated with the introduction of 'virtual or managed learning environments', although eight reported that their pilots had resulted in lower response rates than when they had been paper-based.

Administration of questionnaires

4.45 Thus, the administration of questionnaires will have an effect on response rates. There are two main issues involved: how seriously the collection and use of feedback is taken by both staff and students, and the logistics of administration. The first issue is perhaps more important. Response rates will be affected by how questionnaires are presented to students. If staff appear to believe they are of no importance or they are a bureaucratic imposition on their teaching time, and students believe that staff never take notice of the results, it will not be surprising if students do not take them seriously. This will show up in terms of the 'quality' of responses received (ritualistic/mechanistic answers) and/or in the response rates. As noted above, staff at the institutions we visited were concerned about questionnaire fatigue and its effect on students. However, in discussions with students, it was not the number of questionnaires they were asked to complete that was the problem, but the feeling that the exercise was a waste of time and that no action would be taken in response to their views.

4.46 In terms of logistics, the vast majority of questionnaires are administered to students 'on-campus' at the end of class with time allotted for their completion. Many of the students we spoke to during our visits felt they were given insufficient time for completion, the purpose and use were not always fully explained, and they were not made to believe that their

feedback was important and welcome. There also seems to be varying practice about who distributes and collects the questionnaires. Some institutions involve a student representative in the process, others use staff not connected with the teaching. Using someone other than the teacher helps demonstrate a commitment to independence and promotes confidence in the process. However, an alternative view presented to us is that by using an independent person, it makes the process appear to be a bureaucratic exercise and disengages teaching staff.

4.47 In some circumstances questionnaires are sent by post, especially to 'off-campus' students; here the response rates tend to be much lower. Successful efforts to raise response rates through reminders are made by some institutions, although this is resource intensive. Around a third of respondents said they had used or are using the intranet or email to engage students with questionnaires. However, as reported above, response rates tend to be lower because students are able to choose whether or not they complete them.I If they feel their responses will not be taken seriously, there is no reason or motivation to complete them unless they feel very strongly about an issue. One institution has achieved high response rates to web-based questionnaires by linking the activity with another (e.g. module choices for the following year); here students are brought together at an appointed time in one location. However, this is dependent on access to a suite of terminals, which is not an option for every institution. Others are integrating questionnaire feedback with the introduction of managed learning environments. Again, this has its drawbacks as some students have concerns about the confidentiality of web-based and emailed questionnaires; institutions using these media will need to demonstrate to students that their responses are treated confidentially.

4.48 Another presumption of good practice relates to the administration of questionnaires as follows:

In order that the results may carry credibility with students and others, surveys need to be administered, and the results analysed, in a way which is, and is seen to be, free from the risk of manipulation and distortion.

4.49 Again, **the vast majority of institutions agreed with this presumption** and indeed a few mentioned their use of external agencies. A number of comments were raised, however, regarding the practicalities and realities involved. Where questionnaire use is widespread and frequent, the practicalities (including burden and cost) would preclude the independent issuing and collection of questionnaires. As already mentioned, if teachers are removed from the process, gathering feedback might be seen as a management tool and imply lack of trust in staff. One institution suggested that the only way to guarantee against manipulation and distortion is through student involvement in the process.I Indeed a number of institutions reported using students (especially student representatives) to explain the purpose of the survey and to distribute and collect questionnaires.

4.50 A further issue relates to the extent to which systems for processing questionnaire feedback (i.e. collection and analysis) are centralised or devolved within an institution. Where processes are devolved, there is a risk that manipulation and distortion can take place and a few institutions reported such instances. Centralised processes of collecting and analysing feedback can help ensure against manipulation and distortion by providing a confidential and independent service to academic units and services. Centralised processes can also be cheaper by creating economies of scale and freeing up teaching staff time, whereas decentralisation may hide costs but not save them. If information is to be used at different levels in the institution, a central unit may be better able to service their different needs. While centralisation does not necessarily imply 'standardisation', they are related and this is discussed below.

Standardised v non-standardised questionnaires

4.51 The use of standardised questionnaires within and across HEIs is another presumption of good practice:

> *Student opinion surveys should be conducted consistently within each HEI across its different schools, faculties and departments, in order to generate a consistent set of results. We expect that a core set of standard questions will need to be identified, which all HEIs should include in surveys and which would be reported publicly on a standard basis. This may well imply central administration within the HEI, at least of some elements. We would, however, expect that surveys also allow for individual tailoring to the circumstances of different programmes, departments and units. This gives the staff concerned (academic, support and administrative) the opportunity to obtain the information they believe will be of most value to them in assessing current performance and how it can be improved. So we need to balance consistency in feedback on core issues across the institution without damaging the flow of more specific and tailored information to address local issues for individual groups.*

4.52 The extent to which questionnaires are standardised (i.e. use of a common set of questions) varies between institutions and often within institutions. Some institutions recommend the use of a standard questionnaire, but allow flexibility in terms of additional questions or flexibility in whether individual academic/service units use a standard questionnaire or their own. Others (15) have a central unit that administers, processes and analyses a common questionnaire(s). It may be worth differentiating between 'standard' module level and 'standard' institution or programme level. As we have noted previously, the purposes of feedback tend to be different at these different levels. Standardisation is likely to be more important at the latter level, where comparability of data is important; whereas, at the module level, a sense of ownership and engagement by both teachers and students may require an element of differentiation by subject or course.

4.53 Reasons for standardisation include:

- ensuring high standards and competence levels of questionnaire design and data analysis

- allowing comparisons (both internal and external)

- allowing linkages to other institutional datasets.

4.54 It must be emphasised that questionnaire design, data analysis and interpretation require particular skills and competencies; not all staff can be expected to have these. Staff will make more use of results of feedback the more professionally it is done, and this suggests the need for greater standardisation and maybe centralisation. However, where there is standardisation and centralisation, it is crucial that staff responsible for data collection and analysis are clear about the needs of different users, are aware about how these are changing, and are able to anticipate future needs as well as meet existing ones. One-off special surveys as opposed to 'regular' ones may not require the same level of standardisation; however, they should meet the same levels of professional standards of competence.

4.55 Reasons against standardisation include:

- different purposes of feedback

- different types of provision, delivery and mode of study

- different types of student

- different learning experiences.

4.56 Standardisation can also undermine local ownership and commitment and cut against the grain of traditional governance structures in devolved institutions. In many HEIs, institutional feedback mechanisms have evolved over a number of years and these mechanisms appear to work for the individual institution concerned.

4.57 While **well over half of the respondents to the consultation agreed in principle with the presumption** in favour of a degree of standardisation, a sizeable number disagreed or had serious reservations. The messages reflect some of the issues raised above and are summarised in Figure 4.3.

Figure 4.3: Institutional concerns about the good practice presumption regarding

standardisation
Standardised questionnaires are too generic and will mask diversity.
A core set of questions across the sector cannot take account of the diversity of institutions, provision and students in UK higher education.
Consistency will not allow for flexibility in relation to improvement and will be ineffective in obtaining detailed feedback. Sector-wide comparable data would have little or no relevance to enhancing quality within institutions.
Standardisation implies centralisation, which is an added cost.
Likewise, the use of tailored questions to a standardised questionnaire increases costs and creates technical difficulties.
A standardised questionnaire breaches good survey design, which requires a clear focus and purpose to the exercise and a questionnaire that is not too long.
Response rates will have to be high and consistent to have sector-wide value.
Staff closest to the students should be responsible for identifying the most appropriate method for gathering feedback.
Standardisation and publication imply another purpose other than enhancement.
A solution might be to have a core set of issues rather than questions.

Validity and reliability

4.58 Good practice in questionnaire design requires that the questions asked are clear, succinct and unambiguous. It is essential to eliminate the possibility that a question might mean something different from one student to another. Staff should be able to demonstrate that their questionnaires meet requirements of reliability and validity. A questionnaire that is reliable 'would yield consistent results if used repeatedly under the same conditions with the same participants and is therefore relatively unaffected by errors of measurement'. A valid questionnaire is one that measures what it purports to measure. (See Chapter 3 for a fuller discussion of reliability and validity.) It is unclear how far institutional questionnaires have been tested for their reliability and validity.

Advantages and disadvantages of questionnaires

4.59 One of the main advantages of questionnaires is that they can ensure 'inclusivity' in that all students are given the chance to provide feedback. Questionnaires are also relatively inexpensive to administer, process and analyse, especially when using optical mark readers (OMR) or scanners, although the extent to which questionnaires are standard will determine the extent to which economies of scale can be achieved. Questionnaires provide 'real' evidence in that they document evidence in a relatively systematic way. They also allow comparisons and analysis of trends. Many users of questionnaires in our visits to institutions stated that questionnaires rarely throw up any surprises; their main value is that they provide confirmation about what is already known.

4.60 The negative aspects are that questionnaires tend to be 'ex-post' in that students are asked for their views at the end of a module or programme; students rarely get to know the results, let alone any actions taken as a consequence of their feedback, and in any case would be unaffected by those actions. Questionnaire fatigue can be a problem for both students and staff. Students also complain of boredom and compliance, especially if they believe nothing is being done with the information they provide. Low response rates will affect the extent to which actions and decisions can legitimately be made.

Specific issues relating to student representation systems

4.61 All institutions operate systems of student representation on institutional and local committees. Many institutions also have staff/student liaison committees, which tend to comprise student representatives and staff of a particular academic unit. These meet at regular times during the year. While student representation is universal and institutions value the role, the effectiveness of the system appears to vary between and within institutions.

4.62 The NUS has done much work with local Student Unions in this area. Local Student Unions and senior managers at many of the institutions we visited have good working relations. There is no doubt that much work is being done by the NUS, local unions and senior managers in institutions, but from our discussions with staff and students, more needs to be done. One of the main challenges for institutions is to ensure clarity about the role of the student representative and the responsibilities that go with it, and to communicate and promote the benefits to both staff and students. Another challenge is to motivate students to participate in the system and to be shown that the role has value (skills development etc). A further challenge is that while senior management accept the importance of the student role, not all staff at other levels see the benefits and some actually ignore it, thus undermining the whole system. For example, we were told of instances where student representatives were not given committee agendas in advance of meetings, departments that did not forward student representatives' details to the Student Union for inclusion in the training sessions, and departments that did not have student representatives. If staff at all levels accept the student role as important, students will be more motivated to participate in the representation system.

4.63 Another challenge to make the system work is to ensure that other students know who their student representatives are and what can be expected of them. Student representatives need to be visible to the rest of the student body they represent. Clearly, the onus is on the student representative, but staff can help smooth the way. For example, time should be made available in lectures or seminars when representatives can consult on issues and report back on discussions and actions to the students they represent. Representatives might also have a role in administering feedback questionnaires (i.e. explaining their importance and

distributing and gathering them). Representatives should be treated as equal members at committees and the like, and made to believe that their views are welcomed and valued. Representatives should also be kept up to date with developments and actions arising from committee meetings. As with questionnaires, staff must be seen to take the system seriously.

Advantages and disadvantages of student representation systems

4.64 The advantage of student representation is the 'feed forward' nature of feedback compared to questionnaires. Moreover, some issues can be resolved relatively quickly. The disadvantages are that students lack motivation to participate because they see the role as a chore and they do not feel their views are valued by all staff; and student representatives may be unrepresentative of the wider student body and have few opportunities for effective communication with other students.

Specific issues relating to other forms of feedback (discussion groups)

4.65 Few HEIs reported using other forms of feedback. Often this is because of their resource intensity (e.g. the need for good and independent facilitators, ensuring that a group is representative of the student body, and the time taken to obtain, analyse and interpret the results). However, there are advantages (and disadvantages) to these types of feedback mechanisms.

4.66 A discussion group is organised discussion with a selected group to gather information on their views and experiences of a selected topic(s). Discussion groups are particularly useful when the purpose is to explore whether there is consensus on a particular topic. In terms of student feedback, discussion groups have been used to help determine the types of questions or themes for inclusion in questionnaires.

4.67 Groups may not be representative and can be intimidating for less articulate and self-effacing individuals. Moreover, by their very nature, confidentiality and anonymity cannot be ensured in discussion groups. The role of facilitator is critical to their success. He/she needs to be able to communicate effectively, facilitate discussion, make people feel at ease, challenge group members, tease out differences in views and meanings. The challenges of the role, together with the resource intensity, are probably some of the main reasons why this form of feedback is not common. However, discussion groups can provide a rich source of information and their use in certain circumstances should be considered. They may have a 'one-off' value, for example to discuss some proposed changes or to investigate the nature of a problem already identified by other means.

Specific issues relating to the role of the Student Union

4.68 Student Unions can be effective in helping to represent student views through their:

- involvement in policy formulation/development (e.g. working parties regarding student feedback)
- involvement in the design and review of questionnaires
- role in the student representation system – on institutional committees and in training student representatives for local committees
- promotion of the importance of the student voice in general.

4.69 Many institutions delegate responsibilities and involve Student Unions in their processes concerned with student feedback – and this must be a good idea. However, institutions and Student Unions will need to be clear about the purposes of doing so and any delegation of responsibility should be monitored and reviewed. Indeed a number of institutions monitor the role by including questions about the Student Union in their questionnaires to students.

Recommendations

- reliance should not be placed on a single mechanism for gathering feedback
- different mechanisms may be needed for different purposes, levels and contexts (see section on mechanisms above)
- the purpose of collecting student feedback and how it will be used, including how results/actions will be disseminated to students, should be clearly stated in guidelines to staff and students, and especially at the point when feedback is being requested
- students and staff should be made aware of the benefits of gathering feedback, and the processes involved should be fully explained and understood by all parties. In addition, this will be most effectively done if students and staff find that they are using the results of feedback data e.g. by students in choosing their options, by staff in revising their modules/programmes, by management in planning new programmes, and by the marketing department in promoting the institution

- discussion groups are an alternative to questionnaires and student representation systems (although not necessarily a replacement). They can provide a rich source of information and their uses in certain specific circumstances should be considered.

Specific recommendations on questionnaires

- sufficient time should be allowed for students to complete questionnaires when given out 'in-class', their purpose and use should be fully explained, and students made to believe that their feedback is important and welcome

- if the module is the level at which feedback is collected, consideration should be given to frequency and/or sampling (of modules) to counter questionnaire fatigue

- consideration should be given to capturing student profile data as well as views and opinions to check how far responses vary between types of students

- questionnaires should be standardised (with a set of common questions) within institutions as far as possible to provide a basis for comparison, both within and between institutions. However, it is likely that there will always be some need for special questions to reflect different purposes and contexts, especially at module level

- if questionnaires are not completely standardised, a common core and limits to acceptable variation should be set

- response rates should always be published, and where they are below, say, 60% the results should be treated with some caution, especially if not presented alongside other sources of information. This is not to say that information might not still be valuable, but much will depend on the degree of local knowledge and the availability of other information. However, whatever the response rate, it will be important to check the typicality of respondents (e.g. age, entry qualifications and so on – as mentioned above)

- information on reliability and validity that can be claimed for the data should be provided to all users

- questionnaire feedback should not be used in isolation, but should take account of the existence of and messages from other forms of feedback

- questionnaires should include open-ended questions to provide students with an opportunity to raise issues not covered by fixed choice questions

- a system should be established to ensure that answers to open-ended questions are at least read by teaching staff, if not processed and analysed

- where questionnaires are administered 'in-class', efforts should be made to obtain responses from those students who are not present

- consideration should be given to using students in the process of distributing and collecting questionnaires 'in-class' to ensure against manipulation and distortion of results and to promote independence and confidence in the system

- where web-based and email systems are used to administer and collect questionnaire responses, every effort should be made to demonstrate to students that their responses are treated confidentially

- above all, the collection of feedback information must take account of its intended uses and the nature of the institutional quality assurance and enhancement procedures.

Specific recommendations on student representation systems

- the importance of the role of student representatives should be recognised by staff at all levels (i.e. not just senior management) and by students, and this should be communicated to students

- consideration should be given to involving the Student Union in awareness raising and training in the student representative role

- where training for student representatives is provided by the Student Union, there should be full co-operation between staff at all levels and the Student Union to ensure that students are able to take advantage of the training

- agendas and other papers should be made available to student representatives in advance of meetings and, if necessary, a briefing session held prior to the meeting to discuss issues to be raised

- institutions, students and Student Unions might wish to consider the feasibility of rotating the role between students to share the experience

- time should be made available to student representatives to enable them to gather and feed back issues to the student body.

Analysis and interpretation

Who is responsible?

4.70 Responsibility for the *analysis* of student feedback questionnaires will depend on the level at which it is sought and whether there are centralised systems for analysis. As analysis is often little more than a frequency count, it is a task that faculty or departmental administrators can readily perform when provided with suitable software. Whether this form of analysis is really exploiting the data to the full is another matter. Analysis of student sub-groups or cross-tabulation against other institutional data and, depending upon the nature of the questionnaire, use of techniques such as factor or cluster analysis would probably require the help of a specialist central unit. Similarly, trend analysis or comparative analysis – comparing modules, programmes or subjects, within or between institutions – is likely to be best performed centrally. Analysis should indicate range and standard deviations if averages are being reported. Local analysis, however, helps to ensure that the 'right' questions are being asked of the data and can bring the processes of analysis and interpretation together. Analysis of more qualitative student feedback most frequently takes the form of recording in committee minutes and, at the very least, reading open-ended answers in questionnaires. Discussion groups would permit more sophisticated forms of qualitative analysis, especially if the group discussion has been recorded. However, the analysis of feedback data is probably best not regarded as a 'research task' and will generally be kept simple.

4.71 Responsibility for the *interpretation* of the results of the analysis of feedback data will depend on the purpose of gathering it. If the purpose is to inform and improve the teaching and learning process, interpretation best resides with those who do the teaching (i.e. the teacher and/or the programme team). Certainly, it would be desirable to obtain an initial commentary on the data from those most closely associated with it, so that data *with commentary* would be received by others in the institution. If questions have been asked in questionnaires that relate to teacher performance, then the results should be treated confidentially and their interpretation will remain with the teacher and his/her line manager. Otherwise module and/or programme level feedback will be analysed and interpreted by teaching staff responsible for modules and programme teams. One of their responsibilities will be to summarise the data for use by other groups within the institution. Normally, where

institution-wide surveys are undertaken, a central unit will take on the role of analysis, but may involve other academic and service units to interpret the results, usually through the committee structure.

Current practice

4.72 In the consultation, institutions were asked whether there was a central unit responsible for the collection, analysis and presentation of student feedback, how many staff it had and where it was located organisationally in the institution. In those institutions that operated a common questionnaire across the institution (i.e. module and/or programme levels), central units had been established (around 15); otherwise, there was no central unit to deal with all student feedback (although a central service could be available to departments if they wished to make use of it). In those institutions that operated institution level surveys, these were supported at the central level.

4.73 Institutions were also asked the extent to which the *formulation* of policy about student feedback was devolved within the institution. The vast majority of HEIs have an institution-wide policy that is formulated by the academic board or equivalent. Very few (around seven) allowed policy formulation to be devolved below the central level. In four instances respondents claimed that it was a shared process. The vast majority of HEIs also indicated that the *implementation* of central policy about student feedback is devolved within the institution. A minority (10) had no devolution of policy implementation and one institution stated it had no central policy. Where responsibility for student feedback is devolved, over half the institutions issued detailed guidelines from the centre. Compliance with the guidelines is monitored through the annual monitoring and periodic review processes.

Figure 4.4: Issues to take into account with centralised and devolved systems of questionnaire administration
Is there clarity about the possibly diverse needs of users across the institution in centralised systems?
To what extent can centralised systems be tailored to the needs of individual academic units? What limits should be set to local modifications?
What systems will be put in place to analyse the results of questionnaires (OMR, web-based – or manual if devolved)?
Who should analyse the data if the system is devolved?
How will open-ended questions be handled – both centrally and locally? (There may be anonymity issues if analysis is local.)
What types of analyses will be undertaken (range as well as averages, trends, comparisons of different data sets)?
How will surveys with particularly low response rates be treated?
What other centrally- and/or locally-held information might be generated to triangulate results and inform interpretation (e.g. learning objectives, progression rates, external examiner reports, trends, comparison with other modules/programmes)?
Who will interpret the data and is there standard guidance and set criteria to guide interpretation?

Do the relevant staff (centrally and locally) have the time and expertise?
Can the whole job of analysis and interpretation be done in time to fit in with the decision-making cycles and ensure feedback to students?
Would staff (central and local) benefit from further specialist training and staff development?
Independence of a central unit can help ensure that results are free from manipulation and distortion (see section on collection), but will it lessen the sense of 'ownership' of the work at departmental/subject levels?

Figure 4.5: Issues to take into account in student representation and other feedback mechanisms
Does the system ensure that all students are represented, especially when the student body is diverse and student representatives may not be 'typical'?
Do student representatives have adequate opportunities to consult (privately) the rest of the student group?
Are student representatives fully motivated to take part in the system?
Are they given sufficient encouragement, opportunity and support to participate fully by all staff at all levels?
Are they being used appropriately and could their role be enhanced?
How much weight, compared with other sources of information, should be given to student opinion?
Are discussions and actions reported back to student representatives and do the representatives report back to the rest of the students?

Recommendations

- the question of whether to adopt centralised or devolved systems for the analysis of student feedback data will need to reflect institutional structures and circumstances

- nevertheless, it would seem desirable to maintain some central resource, both as a centre of technical expertise and advice, and to provide capacity to undertake more sophisticated (especially comparative) analyses and to meet the needs of the institution's central authorities on matters of quality and standards

- analysis of feedback data by a central unit can, providing that the unit's independence is safeguarded, help protect against manipulation or distortion of results

- those undertaking the analysis, whether centrally or locally, should ensure that they are informed about the needs of the users of the data and the purposes of collecting it

- if feedback is to inform and improve the teaching and learning process, interpretation should reside with the teaching staff responsible, although such staff might reasonably be expected to summarise and to comment on feedback data for use elsewhere in the institution

- guidance and criteria should be set for such summaries

- feedback should be interpreted in context and with other sources of information that are available.

Actions and decision-making

Current practice

4.74 Typically, in the responses to the consultation, institutions reported that the results of questionnaires administered where the teaching is delivered are seen by module teachers and/or teaching teams, passed to heads of department/school/faculty, and processed through the annual monitoring and review committee cycle. Likewise institution-wide surveys will be processed through the committee cycles.

4.75 Around half the institutions reported that the results of student feedback could be aggregated to institutional level, although non-numeric summaries are often used, especially where common questionnaires were not used. Student feedback collected at whatever level was part of the overall annual monitoring and review process. This allowed the construction of an institutional picture, as the results of student feedback along with other forms of information were passed up the committee structure. Decisions and action points occur at all levels of the structure. Actions at one level are reported to the next and the consequences of the actions reported at subsequent cycles of the committee process.

4.76 The consultation revealed that the vast majority of institutions use the annual monitoring and review process and committee cycles as the main means for following up results, deciding what action to take, checking whether action is actually taken, and monitoring the effect it has. As such, the established formal method for addressing areas of concern is based on the committee cycle, although issues that are of a serious nature or those that can be rectified immediately are often taken outside the committee cycle, but subsequently reported as action taken.

4.77 This somewhat idealised picture of collegial decision-making can, however, disguise the importance of the roles played by key individuals. As indicated previously, individual teachers have much in their power to rectify problems when they occur and are identified. Departmental heads and programme/subject leaders have a role to ensure that individual teachers are properly responsive to feedback from students. If committees are to properly discharge their responsibilities, their secretaries and chairs must ensure that information is provided to them in an accessible and digestible way, that decisions are accurately recorded

and implemented, and the effects of the resultant actions duly monitored. Although the above might appear to be 'stating the obvious', there is a belief among many students and some staff that committees represent a 'black hole' into which issues disappear rather than a route that can provide action and decisions.

4.78 At many points, we have noted the multiple purposes of student feedback. Monitoring the effectiveness of existing teachers is but one purpose. Student contributions to strategic decisions on many issues are also important and these include academic review and planning. It is not necessarily the case that the annual monitoring and review cycle of committees will ensure the use of student feedback for other purposes within the institution, whether centrally or locally within faculties. Data may need to be provided in other forms to reflect other purposes. This is where a central unit can be valuable in having the resource and the expertise to meet a variety of possibly 'one-off' needs and to go beyond routine data processing.

4.79 The main mechanisms for reporting feedback results, and actions taken in response to them, back to students are through staff/student liaison committees and student representation on other committees. Other mechanisms include notice boards, email, intranet and the Student Union. Much of this reporting back depends on the effectiveness of the student representation system and, of course, on whether students decide to consult notice boards etc. One of the main criticisms received from meetings with students was that they never found out what the results of their feedback were and what actions were taken as a result (see section on dissemination). Issues of data presentation and publication are dealt with in the next section but we note that a table of data on a notice board does not necessarily constitute an effective means of communication. There is a real danger that student cynicism may endanger the potentially very valuable functions that student feedback data can perform.

4.80 A further presumption of good practice relates to actions and decision-making as follows:

If students are to be willing to keep completing surveys, it is important that the HEI has rapid and effective mechanisms for deciding, and reporting, what follow up action has been, and will be, taken to enhance quality and standards and to address areas of concern identified by students.

4.81 As mentioned above, some issues, especially at module level, can be resolved immediately. Other issues, especially those related to policies and resources, need time for discussion by various parties and often need to be considered further up the committee chain. The latter takes time and may result in issues being shelved or not being tackled at all.

4.82 **The vast majority of institutions responding to the consultation agreed with the above presumption**. However, some reservations were highlighted as indicated in Figure 4.6

below. While undoubtedly valid, these kinds of comments if repeated often enough can convey to students a sense of 'ritualised excuses' for not responding to their concerns and a feeling that it is a waste of time to express them in the first place.

Figure 4.6: Institutional concerns about the good practice presumption regarding acting on feedback
If the feedback process becomes centralised this may lengthen the decision-making cycle.
Low response rates can distort the true picture and it might be appropriate *not* to take action.
Responses to feedback are not always automatic or axiomatic – some student opinion can be isolated, erratic (and without consensus), or unreasonable.
It may not be possible to make changes as a result of feedback even if repeated year on year (e.g. some programmes may require a mathematics or statistics component and groups of students will often provide negative feedback on these).
Certain issues may require a careful and considered response, which will compromise rapidity. Immediate responses or improvements are rarely possible, but where action does result, current students will be helping future ones.
Surveys are only one means of obtaining feedback and decisions/actions should be taken in the wider context of quality assurance.
Student views can change from year to year and be diametrically opposed. Feedback needs to be contextualised against other evidence, built up over time (while not neglecting serious issues).

Recommendations

- feedback to students is as important as feedback from students. Institutions need to ensure that students are told of the results of feedback and of any actions taken in response to it. When actions cannot be taken, the reasons need to be conveyed to students and, if possible, their further reactions obtained

- virtually all institutions have annual monitoring and review cycles involving key academic committees that can provide an effective means for receiving and acting upon the results of student feedback. Their secretaries need to ensure that feedback data are presented in a digestible way and chairs need to ensure that they are given due consideration in decision-making processes

- notwithstanding the important role of committees, certain key individuals also play important roles. Probably most important is the individual teacher whose response and actions can deal with many issues raised and convey to students the sense that their feedback is valued by the institution as a whole

- for student feedback to feed into more strategic aspects of decision-making within HEIs, it may need to be analysed and presented in different ways and here the role of a central unit is likely to be important. This is also true of qualitative feedback as recorded in committee minutes and reports. There is a danger that such information

can be 'lost' to the decision-making process if not analysed, brought together and placed in context.

Presentation and publication

Current practice

4.83 In their responses to the consultation, all except a few institutions indicated that information is published internally relating to both the results and the actions taken in response to student feedback. Most is published through reports or minutes posted on notice boards or on the intranet/email or through posters. Only a few HEIs said that information was not published. A very small minority publish the results externally.

4.84 However, publication can take a range of different forms. Few institutions appeared to publish the results of feedback in such a way as to inform the choices of modules made by existing students. Comparative and trend data were only infrequently published and there was little attempt to relate feedback data to programme specifications, i.e. to consider the extent to which the latter were being achieved.

4.85 Figure 4.7 and Figure 4.8 indicate some of the main issues concerning the presentation and internal publication of student feedback.

Figure 4.7: Issues to take into account in the internal publication of questionnaire data
What does publication mean (web, notice boards, posters)?
What should be published (results, frequencies, main issues, summaries and/or actions arising from results)?
Feedback that identifies individual teachers' performance needs to be handled sensitively.
Should data with low response rates (say, below 50%) not be published?
Should information be published in a standard form for internal comparability purposes and to assist students in their choice of modules and programmes?
In order to be useful to different users within the institution, data may need to be presented in different ways.

Figure 4.8: Issues to take into account in the internal publication of information derived from student representation (and other forms of feedback)
How and where will discussions and actions relating to student feedback be recorded and published?
How will students (former and current) be informed about their publication? Who should take responsibility – staff or student representatives?
Should student representatives be given a formal (and private) opportunity to report back results and actions to students 'in-class'?

4.86 The following presumption of good practice relates to publication of student feedback:

The consultation paper, 01/66, envisaged that the results of student opinion surveys would be published in summary form. This could be done in HEI prospectuses, or through links on the HEI's website from the prospectus to summary results of the latest survey, how it compares with the previous survey, and the improvement actions taken since that previous survey.

4.87 Institutions in their responses were very concerned about this presumption. Less than a third (25) of institutions agreed in principle with this presumption; a smaller number (17) expressed outright disagreement and the rest (39) raised serious concerns. The issues raised are summarised in Figure 4.9.

Figure 4.9: Institutional concerns about the good practice presumption regarding publication of feedback data
Publication will involve significant resources.
Publication of results is inappropriate; external examiners, reviewers and QAA audit should see results on behalf of the public.
The more results are aggregated, the less useful/meaningful they become – 'it will become an exercise to average out diversity'.
Such information might be useful to prospective students but not for enhancement.
Publication will compromise the purpose of collecting feedback.
There is a danger that module/programme feedback systems will become distorted in order to produce summaries and another set of performance indicators.
The presumption assumes that information will be collected at the level of the programme.
The appropriate place for publication is internal monitoring and review reports.
Publication will be difficult to achieve in large diverse institutions.
Some actions have long time spans which cannot be immediately resolved.
Summaries (including questions asked and methods used) will need to be in a format that cannot be distorted and are comparable across institutions, but can they take account of large diverse institutions and small specialist ones?
Variations in terms of course size, response rates, mission, and subjects would give rise to meaningless comparisons.
Will results be intelligible to audiences outside the institution concerned?
Publication will invite manipulation and distortion.

4.88 Taken together, the presumptions of good practice imply that student opinion surveys should be standardised within institutions (including a core set of questions for comparability across HEIs), be undertaken at programme level, and the results published. Throughout this report, a number of institutional concerns about these presumptions have been raised, including the following:

- the majority of HEIs believe that the module is the most important and useful level at which to capture student feedback for the purposes of improving the quality and standards of teaching and learning. Moves away from that level could undermine purposes and established procedures

- there is a fear among HEIs that context and diversity will be 'averaged out' if standardisation is imposed within and across institutions. Many HEIs have developed feedback instruments over a number of years and would be unwilling to move from practice that supports quality enhancement to practice that was perceived to serve publicity. If steps towards standardisation are taken – and this has advantages for comparative purposes (both internal and external) – they will need to be taken in ways that recognise the importance of contexts

- publication will involve costs and burden in order to summarise results of student feedback and place them in a context that can easily be understood for public consumption (e.g. by developing good practice in the presentation of summary statistics in non-technical ways to encourage sensible use by non-specialists). At the same time systems will need developing to ensure common standards and integrity

- student feedback is only one aspect of institutional quality assurance processes and therefore the dangers of taking it out of context and publishing the results/actions on their own should be borne in mind.

Options for the publication of data on student feedback

4.89 There are a number of options for the external publication of institutional data:

- publish the summary data from the results of student feedback

- publish summary data but set it in an institutional context

- publish only the institutional mechanisms and processes for collecting and using student feedback

- no requirement to publish feedback data

- optional publication of feedback data but no external requirement to do so.

4.90 According to which option is adopted, some degree of standardisation of data collection and analysis will be required. The first option would require some core questions, although these could be augmented by additional questions to reflect local context and interests. The second option could be based on core questions but could also be achievable providing that questionnaires addressed some agreed core themes. Setting data in its institutional contexts would limit its direct comparability but might actually make it more useful. Students typically want to decide between specific programmes at specific institutions (according to their entry qualifications and interests) and need to compare data at a limited number of places, not the entire higher education system. To be helpful to students, published data will need to meet some minimum standards to ensure integrity and will need to contain sufficient common elements to allow some degree of comparison. But it may be sufficient for institutions to describe honestly and accurately what they have done (regarding data collection and analysis) than do exactly the same things – providing the minimum common elements are there. The third option would not seem to provide much information of direct use to

prospective students but would provide reassurance that student views were taken into account by the institution. At the present time, it is clear that the majority view in institutions would favour the fourth option, i.e. not to publish. There is a real fear that publication would lead to distortions in the data collection and analysis process and could undermine the quality enhancement value of feedback data. In addition, as we have already noted, many institutions do not currently have good feedback data at programme level – the most appropriate level at which to publish. It will take time and resource to develop ways of presenting feedback data in forms that will be helpful to prospective students whilst avoiding the dangers mentioned above. We do not think that this should be the highest priority at the present time although we believe that option 2 above – publish some data but set in the local context – would be a desirable development in the longer term. There is, of course, no reason why institutions should not take this step immediately if they feel that their existing feedback systems could support it.

Levels at which feedback data should be published

4.91 At what level should data be presented and published? Module level information may well be beneficial to current students making their module choices for the following year, but this level of detail would probably be too much for prospective students. A solution is to aggregate, although this is not without difficulty, as has been argued in previous sections. However, if questionnaires at module level are designed at the outset with the possibilities of aggregation in mind, many of the problems can be avoided. Certainly, it appears that for institutions who wish to consider publication of feedback data, programme (or equivalent) level is likely to be the most useful level at which information could be published to inform the choices of prospective students.

Recommendations

4.92 Recognising the real concerns that exist in institutions concerning the publication of student feedback data, we believe that external publication of such data should not be made a requirement at the present time. In principle, however, we believe that information derived from feedback data, suitably contextualised, could be published and that it would be useful to existing and prospective students and also to the institutions themselves in suggesting benchmarks and pointing to good practice. In many institutions this would require modification to existing arrangements for student feedback, including safeguards to protect quality enhancement functions. Quite reasonably, institutions will give different priority to taking such steps. Like other published data on quality and standards, any published information derived from student feedback will be subject to periodic audit by QAA.

4.93 However, while we do not recommend a requirement to publish feedback data externally at the present time, we do believe that improvements could be made in the ways data is published internally.

4.94 Publication might include the following:

- student feedback on individual modules might be published within the institution in order to help inform module choice

- some institutions may wish to consider the publication of information derived from student feedback at programme or equivalent level on the institution's website in order to inform choices of prospective students. (Equivalent levels might be subjects or departments.) Feedback would need to be set in its institutional context and could be complemented by other kinds of information, for example retention rates and employment data

- if information is to be published (internally or externally), readers should be provided with information on response rates, reliability and validity and when the information had been collected. The chief consideration should be that information is meaningful to the reader and is not ambiguous or misleading.

Dissemination to students

Current practice

4.95 A separate but important issue concerning publication of feedback data concerns its dissemination to existing students, especially those who have provided the data in the first place. Publication of the results and/or actions taken as a result of student feedback on the web and notice boards does not necessarily imply effective dissemination. Students are selective in the information they access. Often, once feedback has been collected, students have 'moved on' and the assumption is made that they are no longer interested in the results or in any actions taken. However, in our discussions with students, it became clear that they *are* interested – especially in the results of feedback – and understand that actions are not always immediate. They see considerable efforts going into collecting feedback data but they often fail to see similar efforts going into its analysis and use. Below are some possibilities that might be adopted more widely.

Questionnaires

4.96 Feeding back directly to the student group involved is problematic, especially in modularised systems when end of module feedback predominates. Apart from logistics and timing, there are other issues (which have been raised in the section on actions and decision making) concerned with the good practice presumption of effective and rapid institutional follow-up mechanisms. Where surveys fail to meet requirements of reliability and validity, it may indeed be appropriate that no action is taken on their results. When, for example, they have achieved low response rates or the sample returns are biased, it may be unwise to give much weight to the results. But even here it would be wise to keep students informed in order to secure commitment to any future surveys.

4.97 Feedback is more easily given to students if the original feedback is collected before the end of the module or programme. Sometimes this does happen and we were given an example of a teacher who presented students with the 'highlights' of the responses to the questionnaire. This was appreciated by the students concerned. In reporting highlights to students, it may also help the teacher to interpret and understand with students what lies behind their views. At the very least, some indication that the questionnaires have been looked at is likely to be appreciated by students.

4.98 When feedback is collected at the module or programme end, some institutions report feedback from the previous student group to the new group at the start of the module. However, if it is to be achieved, students need to believe that their feedback matters and is taken seriously, especially by the staff who teach them. We heard from students that many staff do not give this impression.

Student representation and other feedback mechanisms

4.99 The main issues here are whether student representatives are effective in reporting back discussions and actions to the student group they represent. Our concern is that often this is not done very effectively. This is not primarily a fault of the representatives themselves – although briefing and training about their role can help their effectiveness – but of the failure by institutions to provide a time and forum for such feedback. The student representative role needs the full support of teaching staff if it is to be successful. A different kind of problem is when student representatives are themselves insufficiently informed to be able to inform their colleagues. This can occur when actions are taken outside of formal meetings and are not reported on until much later, if at all. The student representative is effectively kept 'in the dark'. Again, the onus is upon staff to respect and to make use of student representatives.

4.100 One idea worth giving serious attention to is that of involving student representatives in the administration of questionnaire surveys, both in collecting the information and reporting back the results. They would bring a different perspective to the process and several we spoke to would welcome the additional responsibilities that would be involved.

Recommendations

4.101 In most institutions, actions need to be taken to improve the feedback to students. These might include:

- to encourage students (and staff) to take the process seriously, face-to-face feeding back of the 'highlights' of results/issues raised should be built into feedback processes
- additionally, results and actions of previous feedback can be added to module handbooks or discussed with students at the start of a module
- the timing of feedback (i.e. collection and reporting) will need to be considered if feedback to students is to include information on actions taken
- opportunity should be provided for student representatives to discuss with and report issues to the students they represent in teaching time
- action sheets from meetings should be prepared so that representatives know who is responsible for following through actions and are updated on progress
- student representatives might play an enhanced role in the administration of student feedback processes, including surveys.

5 The National Survey

Introduction

5.1 The Cooke Report recommended that feedback from recent graduates on the quality of teaching and learning on their programmes should be collected and published. This information would be in addition to that derived from institutional surveys and published by the institutions themselves. The primary aim of the National Survey would be to inform individuals when deciding where to study but it might also be of use to other stakeholders, such as employers and professional bodies. The National Survey might also provide information to HEIs which would feed into quality enhancement procedures. The Cooke Report suggested that the National Survey should have two other characteristics:

- responses should be reported at the level of the whole institution, rather than be disaggregated by programme or subject

- the survey might collect feedback on the value of the programme undertaken to subsequent careers as well as perceptions of teaching and learning quality.

5.2 This chapter reports on the design of the National Survey. The fieldwork on which it was based was undertaken in parallel with the institutional survey part of the project. It began with a literature review and consultations with stakeholders, including student bodies, employers' organisations and professional bodies. Issues relating to the National Survey were also discussed with staff and students at the 20 institutions visited during the project. A list of possible topics for the survey was then developed and this was discussed with 50 students.

5.3 Following these discussions a pilot questionnaire was designed and posted to 210 graduates. Concerns over the Data Protection Act meant that institutions were unwilling to divulge contact details of graduates to us, but seven institutions kindly agreed to mail questionnaires on our behalf. They were each asked to send questionnaires to 10 graduates who had graduated in each of the years 2000, 2001, 2002[7]. The response was disappointing, in that only 20 graduates responded (10%), but some useful information was nevertheless collected. We believe the response rate reflected, in part, the time since graduation (even the most recent

[7] The spread of graduation years was chosen because of the interest in feedback on the value of the programme to subsequent employment and, therefore, a possible need to decide how long after graduation the questionnaire should be administered.

graduated around six months previously) and movement of graduates from their last known address. The pilot survey was also undertaken, unavoidably, over the Christmas vacation period and this may also have affected response rates. In one case, the questionnaires were mailed well after the closing date specified and no responses were received from graduates of that HEI. Also, there was no follow-up of non-respondents which would have improved response rates significantly.

Users of the National Survey

5.4 We have considered three groups of potential users: those wishing to enter HE; staff in HEIs; and other stakeholders such as employer organisations and professional bodies. Each of these groups is discussed below.

Those wishing to enter HE

5.5 The National Survey is targeted at this group and they were the primary focus of fieldwork. The National Survey is potentially of major interest to this group since, with the discontinuation of the QAA Subject Reviews, there is no independent survey which directly assesses teaching quality on a consistent basis between institutions.

5.6 The process through which students selected which HEI(s) to apply to was explored with students in order to define the context within which a National Survey might be used. These discussions confirmed the findings of previous studies. In particular the key choice variable for most is the programme offered by the institution. In a previous study[8] undertaken for HEFCE it was found that over one third of the respondents to a postal survey of potential applicants specified this as the single most important factor, and over 80% specified it as important. Teaching quality is one characteristic by which programmes are judged, but not the only one. The earlier study found that quality of teaching as the *single most important factor* influencing choice is low on the ranking – selected by only 6% of the postal survey sample. However, when asked to identify any factor which influenced choice, quality of teaching was specified as important by two-thirds of applicants. A larger scale study[9] found that 42% made use of teaching quality ratings and, on average, rated their usefulness as 2.7 (1 = not at all useful to 4 = very useful). Teaching quality is, therefore, an important selection

[8] SQW Limited (1999) Providing Public Information on the Quality and Standards of Higher Education Courses *Final Report to HEFCE, HEFCW, DENI, QAA, SHEFC.*

[9] Connor H, Burton R, Pearson R, Pollard E, Regan J (1999) Making the Right Choice: How Students Choose Universities and Colleges

criterion, but not a decisive one for many. Other aspects of the programme which had a significant influence on choice are:

- what is taught, i.e. course content and intended outcomes
- how it is taught, covering a range of factors such as staff-student ratios, work experience opportunities and so on
- how students are assessed.

5.7 The aspects of the programme are, however, qualitatively different from teaching quality in that they are largely factual issues which can be communicated through prospectuses and other material. Teaching quality is much more subjective. Degree results reflect the capabilities of students, as well as teaching, and current classifications limit the extent to which applicants can differentiate between HEIs.

5.8 As mentioned above, the programme is not the only consideration for applicants: location, reputation of the institution, cost of living, accommodation and other facilities, and perceptions of the social life all feature heavily. Many were also interested in employment prospects, as indicated by the employment record of previous graduates, especially but not only when vocational courses were considered.

5.9 Teaching quality is therefore perceived as only one aspect of programme suitability, and the programme is only one of the key choice variables, albeit the main one for many applicants. Nevertheless, there was a widespread view that data from the National Survey would be useful because:

- information on quality is considered to be patchy at present. Many rely on advice from teachers/family but this is often recognised to be imperfect. In addition, with widening access, a greater proportion of applicants will not be able to access information from family sources. League tables/guides are widely used, but also felt to be imperfect, and in some cases misleading
- even when applicants can access advice it can be difficult to compare HEIs on a consistent basis
- the views of previous graduates are considered especially valuable since they are felt to be independent and informed.

5.10 There was considerable diversity between those students interviewed, in the ways in which National Survey data would have been used had it been available at the time they applied to HE. To a large extent this diversity reflected factors such as access to informal advice, location constraints and the availability of specific programmes. However, the main uses are likely to be as follows:

- as a check on selection decisions made on the basis of other criteria
- helping to define a longish list of HEIs which the applicant would then investigate in more detail
- helping to decide between two or three programmes which appear equally attractive on the basis of other criteria
- checking claims made by the HEI in prospectuses
- and related to the last point, prompting questions which the applicant might raise with the HEI at open days. Open days are widely considered to be helpful, and influential, in making choices. However, they are perceived to be 'managed' in terms of the information provided and the existing students to which applicants are given access.

5.11 It is worth emphasising that the vast majority of students interviewed took a realistic view of the extent to which information derived from any kind of national survey could provide accurate assessments of teaching quality. None would make a decision solely on the basis of National Survey information, and we consider fears that the National Survey might distort choices to be misplaced.

5.12 There is obviously a requirement that National Survey outputs should be a useful reflection of teaching quality but the students interviewed also emphasised two other criteria the National Survey should meet:

- it should not be conducted by the HEIs themselves. Most of those interviewed would accept that individual HEIs might have a role in administering the survey document to their own graduates but design of the questions, interviewing (if not a postal survey), analysis and reporting should be done by an independent body. Otherwise the results will lack credibility
- the information needs to be up-to-date. There was some diversity of views here but most felt that the survey needed to be conducted at least every two years, and many

that it should be annual. Rightly or not, the perception is that quality can change quite rapidly and less frequent surveys could be misleading.

5.13 Students were, almost unanimously, of the view that information needed to be provided at a lower level of aggregation than the institution. This follows from the key interest in programmes discussed above. Institutional averages could obscure differences between high and low quality subjects/programmes and would therefore be of limited use to applicants. The key points to emerge from the discussions were:

- there was little interest in information at the module level. It would sometimes be difficult to map modules to programmes and the volume of information could be overwhelming

- applicants, ideally, require information on programmes but, as was mentioned above, many thought that the quality of a specific programme might change fairly quickly

- related to the last point, many thought that information at the 'subject' level might be a suitable compromise, since quality is perceived to be more stable at this level and subjects can be mapped to various programmes. There is a major issue as to the level of disaggregation of subject classifications and this is discussed further below.

Staff in HEIs

5.14 There are two possible uses of the National Survey by staff within HEIs. First, but of secondary importance to the current project, the National Survey data could be valuable to those undertaking research in the educational field. There is a good deal of evidence in the research literature[10] that the outputs from student surveys can be used to investigate a range of issues such as academic performance and approaches to learning.

5.15 The second possible use is for quality enhancement purposes. National Survey data could be used by an institution for benchmarking purposes or perhaps to identify specific areas which need to be addressed. Some students, but not staff, also suggested that publication could have a direct impact on quality enhancement since if their institution was rated poorly in comparison to others then the university would take action to improve in the future. Despite these potential impacts on quality enhancement the majority of staff we interviewed did not believe the National Survey could contribute in practice, and many of these were quite strongly against the survey being undertaken. This was consistent with responses to the

[10] See Chapter 3

Cooke Report – *'Only a few institutions supported the introduction of a national student feedback questionnaire'*.[11]

5.16 There appear to be three main reasons for this generally negative view of the National Survey. First, a belief that a National Survey of the broad type envisaged could not provide valid assessments of teaching quality. This is an important methodological issue and is discussed further below.

5.17 Second, a belief that the National Survey could add little if anything to internal procedures, including student feedback. Where HEIs were more positive this often reflected an interest in information on the employment related benefits of the programme, rather than teaching quality per se. The more general belief reflects a number of issues and concerns:

- any National Survey must relate to student experiences of a programme as a whole which will comprise several modules/units of study. Even if the National Survey provided valid assessments it will not identify which parts of a programme might be problematic, and can therefore provide, at best, a very limited basis for action and enhancement. In contrast, most internal feedback mechanisms are targeted at a specific level, often the module, and provide a more useful guidance for remedial action. Some of those consulted recognised the importance of considering programmes as an entity in order, for example, to assess how well individual modules were fitting together. However, their view, typically, was that these issues are best addressed through internal feedback mechanisms, including staff-student discussions

- the views of current students are more relevant than graduates, however recent the latter might be, since there is a need to obtain fairly immediate feedback on programmes

- a need to obtain information in a timely fashion, and a belief that however efficient the management of the National Survey it would take far longer than internal procedures to produce information.

5.18 Third, there is widespread concern that the National Survey would lead to the creation of another league table, and one which might have a significant influence on external perceptions of the institution. Staff are concerned that there would be pressure to improve positions in the league table but, given the scepticism over the validity and usefulness of the National Survey, this might represent a diversion from more beneficial activities.

[11] Information on quality and standards in higher education, Final Report of the Task Group *02/15 Annex B Paragraph 40*

Other stakeholders

5.19 The National Survey is intended primarily as an instrument to inform student choice and we therefore undertook only limited consultations with other stakeholders. However, contact was made with 12 organisations. They were a diverse group and their responses reflected this diversity. Nevertheless, the following themes were common:

- there was general agreement that the National Survey could be useful to applicants to HE. However, almost all felt that results would need to be disaggregated below the institutional level, either to programme or subject, if the information was to be of any value

- most felt that applicants would also be interested in graduates' views on the value of their programme of study to subsequent employment. However, there were some caveats to this. It was pointed out that the skills needed will change over a career. In particular, responses based on first job after graduation might be modified over time. For this reason the National Survey might need to be conducted more than two years after graduation. A further complication is that graduates from different subjects take differing periods to move into employment. The timing in the economic cycle will also influence responses. At the bottom of the cycle, graduates may take jobs for which they are 'overqualified' and the study programme may therefore appear to be of less relevance

- there was less certainty as to whether the National Survey would be useful to employers recruiting graduates. It might help employers to decide which universities to target, although this would also require disaggregation below institutional level. However, several pointed out that most companies which are large enough to target universities have relatively good information on quality from informal sources and contacts. The point was also made that employers recruit individuals, so their personal attributes and capabilities are much more important than the particular programme of study they have followed. This was also one of the findings of the previous study for HEFCE

- finally, one of the professional bodies consulted stated that the National Survey could be useful to them in validating courses. But data would need to be available at the programme level. We suspect, however, that if institutions themselves published information in line with the recommendations in the Cooke Report then the National Survey would become redundant for these purposes.

Conclusions

5.20 We have little doubt that applicants to HE would find a well designed National Survey, which provided information at below the institutional level, useful. This reflects: the importance attached to programmes when deciding where to study; the fact that teaching quality is one characteristic by which programmes are judged; and the lack of independent and credible information currently available. There is little evidence to suggest that the National Survey would contribute directly to quality enhancement. Most HEI staff we interviewed felt that it could add little to internal feedback mechanisms and there is quite widespread resistance to the introduction of such a survey.

Methodological issues

5.21 There are a number of underlying methodological issues which a National Survey will need to address. These are discussed in this section.

Can a National Survey provide valid information on teaching quality?

5.22 This, obviously fundamental, question is difficult to answer in the UK context because national surveys, of the type under consideration, have not so far been undertaken. Assessment of data validity requires a careful analysis of responses and for this, and other, reasons we recommend that the first run of the National Survey is treated as a large scale pilot enabling various tests of the questionnaire structure to be undertaken. However, experience from Australia suggests that a National Survey can provide valid information.

5.23 A survey of Australian graduates has been undertaken annually since 1993. The survey uses the CEQ[12] which derives from work undertaken in England during the 1970s and early 1980s. The CEQ has undergone many modifications since its initial development and subsequent introduction in Australia. It is, so far as we are aware, the only survey instrument for evaluating programmes (as opposed to modules or teachers) which is based on research into educational processes in HE and has been subject to extensive testing and analysis by numerous researchers. For these reasons, we have based our recommendations for the National Survey instrument on the CEQ. The CEQ was discussed in detail in Chapter 3 but its basic structure is as follows:

[12] A number of English HEIs have used, or are considering, the CEQ to survey current students. LTSNs have also used the CEQ in specific subject areas.

- the instrument comprises a number of *scales* each relating to a different aspect of the teaching and learning process, for example Good Teaching, or Clear Goals and Standards. The number of scales, and what they are seeking to measure, has varied over time but in most years there have been five

- for each scale there are a number of individual statements or items. Graduates are asked to indicate their level of agreement on a five point scale, from strongly disagree to strongly agree

- there is a single overall satisfaction item – '*Overall, I was satisfied with the quality of the course'*.

5.24 The research which has been undertaken in relation to the CEQ does not demonstrate conclusively that it is able to measure perceptions of teaching quality. However, it does show that the CEQ is able to satisfy a number of important and necessary conditions. Specifically[13]:

- internal consistency of the scales is generally satisfactory

- composition of the scales and the construct validity of the CEQ are broadly satisfactory

- criterion validity of the CEQ as an index of perceived quality can be tested by examining the correlations between the respondents' scores on the constituent scales and their responses to the item concerned with their overall satisfaction. Typically, all of the scales show statistically significant correlations with ratings of satisfaction

- discriminant validity of the CEQ is shown by the fact that the respondents' scores on the constituent scales vary across different academic disciplines and across different institutions of higher education offering programmes in the same discipline.

On what basis will graduates respond to the National Survey?

5.25 The previous section indicated grounds for believing that the CEQ can provide valid information on teaching quality. However, two more direct concerns over the basis on which graduates would respond to the survey were raised during consultations. The first is the extent to which graduates are able to make valid judgements of quality. Very few graduates

[13] More information on relevant studies is provided in chapter 3.

will be able to compare one programme of study with another so they cannot make comparative assessments. However, they are able to evaluate experiences against expectations and aspirations. Most students we interviewed recognised this but, nevertheless, felt such feedback would have been useful to them.

5.26 The obvious issue which arises is that if expectations differ between graduates then this will influence responses to the National Survey. If expectations are randomly distributed between graduates of different HEIs then there is no problem; a large scale survey would average out extremes. However, if graduates with above average expectations tend to be concentrated in certain universities then these universities would score relatively poorly in a National Survey. We would speculate that students with higher entry qualifications, and therefore greater choice of where to study, are more likely to have higher expectations. If so, universities with higher entry requirements will tend to be undervalued in the National Survey.

5.27 There is some limited evidence that this will not occur. A recent survey[14] found that graduates of pre-1992 universities generally expressed higher levels of satisfaction with their courses than graduates of post-1992 universities, and the latter will, on average, have lower entry requirements than the former. It may, however, be the case that there is systematic variation in satisfaction levels within the two groups.

5.28 However, even if responses to the National Survey are biased in this way, it may not be a problem from the perspective of applicants. National Survey data is only likely to be used to make comparisons between subsets of universities and these subsets are likely to comprise HEIs with similar entry requirements. As such, applicants would have access to information provided by a peer group with a similar range of choice and possibly also expectations. National Survey information could, therefore, still be useful. Comparisons across the sector as a whole could, however, be problematic. There are obvious 'PR' problems for those HEIs which receive lower ratings because their graduates had above average expectations. This also suggests that caution would need to be exercised if the National Survey was to be used for external monitoring purposes.

5.29 The second issue concerns the extent to which graduates are likely to give accurate feedback. Many we consulted, including students, felt that graduates would tend to overstate levels of satisfaction because:

- they will be unwilling to publicly criticise their former teachers

[14]Connor H, Pearson R, Pollard E, Tyers C, Willison R 'Right Choice? a follow-up to 'Making the Right Choice'Universities UK, November 2001.

- it is in their interests to overstate the quality of the teaching they have received, rather than point to shortcomings, since (most) will have received a degree from the HEI.

5.30 It is impossible to know how important these tendencies might be in practice. Some we consulted also pointed to the danger of disgruntled graduates seizing an opportunity to criticise the HEI unfairly. However, we would note that the point made above also applies here. Unless there is a systematic correlation between the likelihood of overestimating quality and specific groups of HEIs then the National Survey could still provide valid comparative information on HEIs and would therefore be useful to applicants. We cannot see any reason why a systematic correlation should exist.

What aspects of the HE experience are relevant to teaching quality?

5.31 As was discussed above, there are many aspects of the HE experience which are relevant to applicants when selecting which HEI to study at. Some, such as social facilities, are clearly unconnected to teaching and learning, but in other cases the distinction is less clear cut. A good example is student accommodation. There is no doubt that students are interested in quality, cost and availability because it will be their residence for a significant period. However, the suitability of accommodation for study purposes – quiet space, computer access etc – also impacts on the quality of the teaching and learning experience. Our view is that the National Survey should not seek to cover this type of information and should instead focus on teaching and learning more narrowly defined. This is essentially for pragmatic reasons:

- as with topics such as method of assessment, they are essentially factual questions and information can be accessed from prospectuses, open days and so on

- covering these topics would add to the length of the survey instrument and, with the aim of maximising response rates, this should be kept as short as possible.

5.32 There are analogous issues with respect to the information needs of different groups of students. In particular:

- disabled students where physical access and learning resources have a direct bearing on teaching quality

- mature students, some of whom may require additional support early on and access to facilities such as crèches

- part-time students, for whom the flexibility of provision and opportunities to integrate with the student body may be especially important.

5.33 Some of those consulted felt that special (additional) questions should be added to the survey instrument for these groups. While we can see merit in this we believe the National Survey should be restricted to a common questionnaire because:

- again, some of these issues are essentially factual questions, although we recognise that information on the quality of provision could also be important

- it would add significantly to the complexity of the National Survey since different groups would have to be identified in advance

- even if specific questions are not asked, it would be possible to analyse responses according to student group. The numbers responding from a given programme/subject area could be too small to report with any confidence but it might be possible to report at the level of the institution. This may be adequate for many of the specific issues the groups face

- the information needs of special groups are probably more appropriately addressed through institutional surveys and this could be a subset of the information which institutions themselves publish.

What level of response is required?

5.34 Any survey may be subject to a number of errors. *Sampling* error may arise because the questionnaire is sent to a sample of the population and that sample does not properly represent the underlying population. However, if the questionnaire is sent to all graduates there would be no danger of sampling error affecting the reliability of the survey. There would, however, be the real possibility of response error (the responses received are unrepresentative simply by chance) or, more worryingly, response bias (those that fail to respond have systematically different views from those that do)[15].

5.35 It would be optimistic to expect a universal response from students but it is important that the response rate is as high as possible. As in any survey, the higher the level of response the

[15] As is discussed below, we recommend that various aspects of the survey are investigated and tested during its first run. This would include the extent to which non-respondents differ systematically from respondents in relevant background characteristics.

more one can have confidence in the representativeness of the data, but when it comes to looking at some of the smaller analysis groups, such as particular subjects within particular institutions, a low response rate may mean that there are too few responses in a particular cell to allow that cell to be analysed.

5.36 Even though sampling error is not a problem with census data, there are real dangers in attempting to draw conclusions from very small cell sizes. This is because the smaller the number of observations within a cell, the less representative it is possible for them to be of the total number of members of the cell. This is more than just a matter of response rates, for the same level of response rate may lead to a large cell justifying some analysis whereas a small one does not. To take a somewhat extreme example, if there are 100 students within a particular institution and 50 take part in the survey then unless there is some noticeable response bias, it is likely that all of the different characteristics within the student body – different age groups, genders, different methods of financing and so on – will be represented within the 50 that actually take part in the survey. If in another institution there are only two students on the particular course, one male and one female, one older and one younger, then even if the response rate is the same, 50%, it would be impossible for all the variability within the student bodies to be represented.

5.37 This means that limits should be set for a minimum cell size below which no analysis should be reported. In sample surveys of a population this is typically set at 100 or possibly 50, but if the National Survey is based on a census, we do not see the need for anything like as large a cell size. Our view is that response rates as small as 10, to a census survey, could still convey meaningful information. Our preference is for a census-based approach but this will add to the costs of the National Survey. We therefore recommend that the first run of the National Survey investigates the extent to which a sample-based approach will generate adequate response rates.

Undertaking the National Survey

Timing

5.38 The key considerations for the survey timing are: when will those surveyed be able to give meaningful responses; and when might response rates be maximised. The first of these obviously depends on the information sought. The Cooke Report suggested that the National Survey might collect information on the value of the programme to subsequent employment, which means the survey would be delayed until some point after graduation. If the survey is to cover both views on teaching quality and value to subsequent employment, its timing will inevitably represent a compromise. The need to ensure that memories of the teaching

experience are fresh implies an early survey. An early survey will also enhance response rates, since contact details for graduates become less reliable over time[16]. The need to collect employment related information implies a survey some period after graduation, so that graduates have had the opportunity to enter employment and evaluate the usefulness of their programmes.

5.39 There is a real dilemma for the National Survey here. During project consultations there was some scepticism as to whether a National Survey could provide useful, subjective, information on the value of the programme. There appear to be a number of issues:

- doubts as to whether graduates can evaluate their programmes from this perspective, at least via a postal questionnaire
- related to the last point, whether graduates could disentangle the impacts of their undergraduate programmes from other influences, including on-the-job learning and in-company training, but also any subsequent training provided by HEIs
- evaluations may be overly influenced by the current post and responsibilities, whereas the real value of the programme might only emerge as graduates assume more senior positions
- as was mentioned above, the occupations which graduates enter, and therefore the relevance of their programmes, will vary over the economic cycle and it may be difficult to adjust for this when reporting the results.

5.40 There are, therefore, real doubts as to whether the National Survey is the appropriate mechanism for collecting this kind of information. If it was to be used for this purpose there is a fairly broad consensus that the National Survey would need to be conducted at least two years after graduation. In our view, such a delay would create problems for the teaching quality assessment purposes of the National Survey. There are logistical problems in contacting graduates even two years after graduation. In addition, if the survey was undertaken two years after graduation then many graduates would be making assessments on teaching which they experienced up to five years previously[17]. Apart from recall issues, this would reduce the value of the information to applicants since there is a perception that quality

[16] Although the total response rate was low, the pilot survey undertaken during the current project generated significantly better response rates from recent graduates than from those graduating in 2000.

[17] The period would be even longer for part-time, sandwich and other graduates taking more than three years to complete their degree.

can change over time (and the content of programmes also changes) and the students we consulted emphasised the importance of up-to-date information.

5.41 We note that the new FDS will be seeking some information on the value of programmes in obtaining subsequent employment (although not on the quality of teaching and learning). We understand that a follow-up sample survey of graduates, at a later date after graduation, is also planned. Given this, and the difficulties discussed above, we recommend that the National Survey should not collect information on employment related issues. If this is accepted, it offers the scope to bring the National Survey forward in time which will considerably simplify its administration. More important, the survey could take place at some stage before graduates 'leave' their institution. The information on contact details for graduates is good at this stage and response rates are therefore likely to be maximised.

5.42 There would be three main options for the timing of the survey:

- *after the degree results are known*. This might be the simplest solution administratively and questionnaires could, for example, be sent with information on graduation procedures. However, there is a possibility that responses would be influenced by the results (graduates with good results evaluating their courses more positively). We note, however, that given the extent of modularisation throughout the sector, many students will have a good insight into their likely degree class before final exams

- *after exams but before results are known*. This might eliminate some of the bias mentioned above, but equally perceptions of how well graduates felt they had performed in the exams could influence responses. It may also be difficult to encourage graduates to turn their attention to a questionnaire immediately after the pressure of exams is over

- *before exams*. This option has some attractions in that graduates could have experienced the entire programme but not be overly influenced by expected outcomes. However, to increase the chances of a good response rate the survey would need to be some time before the pressure of revision and exams starts to mount; this could be as early as the end of term eight, or semester five[18]. The problem with this timing is that it could clash with internal feedback surveys and lower response rates to both. In addition, although programmes will be (largely) completed the respondents would still be students rather than graduates. As such,

[18] This is the period when NOP Research Group undertakes the Student Income Survey for similar reasons.

institutions may feel the National Survey is adding less to the information they receive than it might.

5.43 We have no evidence which would enable us to judge whether, in practice, the answers given to the survey are likely to differ significantly according to the timing of the survey. However, we do believe that response rates are likely to be maximised if the third option above is selected. Contact addresses will be most accurate then and we feel that the survey will be given more attention by students than at a later date. We would therefore recommend that this option be adopted. However, possible differences in response rates, and the in the nature of responses, could be tested during a pilot of the National Survey.

5.44 There is a special timing issue in relation to students who do not complete their programmes. It is important that these are included in the National Survey and any responses they provide should be analysed with those from students in the same cohort who do complete the course, i.e. those graduating when the non-completers would have graduated. We believe the best way of contacting non-completers is to send a questionnaire as soon as it is known they have withdrawn. We recognise that HEIs will often not know if a student has withdrawn until sometime after the event but we believe this is the most effective strategy. The alternative is to mail them at the same time as those who have completed. However, this could be two more years after withdrawal, and contact information is likely to be poor and the incentive to complete the form low.

5.45 A related timing issue is how frequently the survey should be conducted. A biennial survey might be appropriate since this would reduce direct costs and the administrative burden on institutions. The main argument against a biennial survey is that the information will be less up to date. Some of those consulted during the study believed that teaching quality could change significantly over a period as short as two years as a result, for example, of managerial changes or rapid expansion.

Administration

5.46 The initial intention for the National Survey was that it would cover all HE students undertaking programmes up to and including first degree level. The proposed questionnaire, and recommended method of administration, can accommodate this diversity and we do not recommend that the coverage should be restricted.

5.47 We recommend that the National Survey should be a census, rather than a sample survey, and should be administered by post. The main reason for a census approach is the need to report results by institution and also by subject area. We think it likely that the number of responses

in some 'cells' would be unacceptably low if the National Survey was sample based. The survey will, however, be large and the costs of a telephone interviewing would be substantially greater. More important, given the purpose of the National Survey and the nature of the questions, it is not clear that telephone interviewing would add anything to a postal questionnaire.

5.48 We think there would be merit in all aspects of the survey (mail out, analysis and reporting) being contracted to an independent organisation. The main benefits would be the chance to exploit economies of scale and to demonstrate that the survey was independent of the HEIs. However, this is only feasible if HEIs are prepared to divulge contact details on graduates to the contractor. There was some diversity between HEIs we interviewed, but a substantial number believed that the Data Protection Act prevented them from doing so. Several student surveys have encountered problems because of this reluctance.

5.49 Our own view, based on expert advice in relation to similar surveys, is that the Act does not in fact prevent the involvement of outside contractors in this way. There is also experience from other sectors to draw on, for example, the statutory surveys of patient satisfaction which have been conducted over the last few years among hospital patients. These were postal surveys with a random sample of recent patients and a common questionnaire used for all patients. Some NHS Trusts have the resources to conduct the mailing out of questionnaires themselves, but most did not. Sub-contracting this task to a research agency was not a simple matter however, because of the data protection problem of patients' names being released to the research agencies. The solution that was found for this was for a member of staff in the research agency to receive an honorary contract making them effectively a member of staff of the trust itself. In this way they were entitled to have access to the patients' name and address details and were thus able to manage the mail out process within the agency.

5.50 In fact, because of the purposes for which the research is being conducted, it is almost certain that institutions would not be risking prosecution under the Data Protection Act were they simply to release details of their students to research agencies to conduct satisfaction surveys. This is a process which is used on the survey of Post 16 Learner Satisfaction, conducted by NOP Research Group on behalf of the Learning and Skills Council (LSC). After initial concern from the Association of Colleges that their members may be infringing the law by supplying the necessary sample to NOP Research Group, detailed consultation took place between the LSC and the offices of the Information Commissioner, as a result of which the Association of Colleges were assured that their members were unlikely to suffer any consequences from releasing names to research agencies.

5.51 The release of names was more important for the Post 16 Learner Satisfaction survey, for that relied on a sample being drawn from the complete list of all students. This meant that NOP Research Group had to be given details of all learners in order to be able to draw the required stratified sample. If the National Survey is based on a census rather than a sample, access by the contractor to contact details is less important.

5.52 Whatever the legal position in relation to the Data Protection Act, it must be recognised that many HEIs would be reluctant to divulge contact details and, in the absence of a test case, they are unlikely to be persuaded by abstract arguments. Given that many have also expressed a view that the National Survey should not go ahead, we think it would introduce an unnecessary complication to insist that they provide this information. At the very least, this would delay any launch of the National Survey.

5.53 We therefore recommend that, at least initially, the questionnaire would be distributed to graduates by the institutions themselves, with data processing, analysis and reporting undertaken by a single central organisation. We envisage some elements of the process closely resembling the current FDS. However, we would emphasise that the National Survey process would differ in important ways from the FDS and, in particular, that we do not believe it would be possible to combine the two surveys:

- the FDS is conducted six months after graduation. There is no rationale for this time period in relation to the National Survey. It is too short for graduates to evaluate the contribution of their programme to subsequent employment and misses the opportunity to contact graduates before they leave the institution

- the FDS begins with a postal survey to which, we understand, the response rate is around 25%. The 80% response rates achieved by the FDS reflect a telephone follow-up process. However, contact at this stage is often not with the graduate, but instead with someone who knows the destination of the graduate, for example a family member. Clearly this would not be appropriate in relation to a National Survey of teaching quality, and HEIs would not be expected to undertake telephone follow ups of non-respondents. Their involvement would be restricted to mailing out initial questionnaires and subsequent reminders

- there are concerns that combining a survey seeking information on the quality of an institution (National Survey) with one seeking factual information on destinations (FDS) would reduce response rates to the latter.

5.54 Distribution of the questionnaires by the institutions themselves also allows them the opportunity to reduce costs by using internal postal means rather than the Royal Mail.

However, we would not recommend this approach. While many institutions will have systems of inter-departmental post via pigeon holes, it would be unwise to rely on this as a means of getting questionnaires to students unless there is convincing evidence that pigeon holes are used regularly by all students. As this will, we expect, be difficult to prove it will be safer to insist on questionnaires being mailed out to students through the post.

5.55 This will also serve to distance the institution slightly from the survey, which is important if respondents are to feel assured of confidentiality. Indeed, the use of an agency to conduct the mail out would take the process a step further away from the institution and would increase the feeling of confidence among respondents about confidentiality. This is why NOP Research Group, in common with many other agencies, recommend that even companies which can easily distribute internal questionnaires to all staff during staff satisfaction surveys do, in fact, have them sent by post to employees' home addresses to distance the company from the survey itself.

5.56 As was mentioned above, it is important that the response rate is as high as possible. This will inevitably mean the use of reminders although these will increase cost. The same issues about privacy of names and addresses of students will apply equally to any reminders as they do to the initial mail out. The number of reminders needed will depend on the level of response to the initial mailing and initial reminder. At least in the early stages, it will be worth sending out two reminders and measuring the cost effectiveness of this, then reaching a decision whether each survey should be treated on an empirical basis in each institution, with reminders sent out as necessary, or whether to recommend an overall strategy to apply to every survey.

5.57 Whether or not the mailing of questionnaires is carried out by the institution or by a single agency, we consider it essential for responses to be sent back to an independent organisation for data processing and data analysis. Again, this will result in incurring postage costs rather than using departmental drop off boxes, but we cannot stress enough the importance of convincing respondents that the answers they give will be entirely confidential. A cheap survey in which the respondents cannot risk honesty is of no value whatsoever. Some of these extra costs can be offset by cost savings such as from the efficiency of data processing. Data entry is certainly one area where there are savings to be made by economies of scale, and the use of a constant analysis design will also improve efficiency.

5.58 The use of a single organisation in this way will complicate the process of handling reminders. If contact details are not divulged by the HEIs then the organisation will need to inform HEIs of non-respondents, on the basis of the HESA unique identifier, and the HEIs

will need to match the unique identifier numbers with addresses. However, we consider that the process could still be managed effectively.

5.59 One issue which needs to be addressed is the identification of those to be surveyed. For the FDS, HESA sends institutions a list of graduates but this is not available until the December following graduation, which would be too late for the National Survey given the recommendations on timing in the previous section. One possibility would be to ask HEIs to identify those who are about to complete their course or graduate (depending on the timing of the survey), but there are two problems with this:

- we understand that it would be a non-trivial exercise for some institutions to identify those about to graduate. This information obviously exists within the HEI but it may reside on unconnected databases which would require some effort to consolidate[19]. Apart from the burden which would be imposed on these HEIs, there is likely to be variability between HEIs in the ways in which they identify target groups, and uncertainty over the size of the underlying population (and therefore the actual response rates achieved)

- for flexible modular courses it is not always clear when students will graduate.

5.60 For these reasons we recommend that HESA provides institutions with a list of individuals to be surveyed. In the interests of simplicity and consistency we suggest that:

- for programmes of a fixed length, this would be those who were in the penultimate year of their programme during the year previous to the survey. Thus, for example, if a survey was undertaken in 2002-2003 then the target group for three-year programmes would be those in their second year during 2001-2002

- where there is no fixed length, those in the fourth year of their programme in the year prior to the survey[20].

5.61 In both cases, some students would be included who are not about to graduate, for example those repeating a second year. However, they would all have substantial experience on which to base responses, and this approach would ensure consistency while minimising the demands

[19] Some HEIs identified this as a problem during consultations on changing the timing of the FDS

[20] Where such courses represent a significant proportion of an institution's programmes, for example the Open University, an alternative strategy would be for HEFCE to have bilateral discussions on the definition of target groups.

on HEIs. It would also ensure that each questionnaire contains the HESA unique identifier so that additional analysis of the data can be undertaken by HEFCE, if and when required.

The proposed questionnaire

5.62 As is discussed below, we recommend that the first run of the National Survey be treated as a full-scale pilot exercise during which a number of issues are investigated in more detail than has been possible during the current study. One of the key tasks will be to finalise questionnaire design, but the current project has explored this issue and a proposed questionnaire is presented in Appendix E. The questionnaire draws heavily on the Australian CEQ. We consider this to be a good basis for the National Survey for the following reasons:

- as was discussed above, research evidence suggest that the CEQ provides valid indicators of teaching quality

- the structure of the questionnaire has been validated by many studies. The CEQ groups individual questions (*items*) into *scales*. Research has indicated that the items within each scale are consistent indicators of that scale, but also provide additional information rather than simply mirroring another item. In addition, factor analysis indicates that most, but not all, items are relevant to one scale only. One implication is that responses can be aggregated to scales and reported at this level. This is a major benefit since responses can be summarised into a relatively small number of indicators

- the CEQ has been widely used, although only in Australia at the national level, and is subject to continual testing and development. This body of knowledge can be drawn on for the future development of a survey in the UK

- the questionnaire can be administered effectively via a postal survey.

5.63 The CEQ has gone through a number of development stages. Until recently, it comprised five scales, with 23 items, and a single 'overall satisfaction item'. For 2002, it comprised two core scales (Good Teaching and Generic Skills) plus overall satisfaction which all HEIs are to use, and a number of optional scales. The proposed questionnaire contains the core elements plus some additional scales which were identified as providing potentially useful information during discussions with students. The pilot contains an additional section asking graduates to indicate the three best aspects of the course and the three aspects which could be most

improved[21]. In part this provides a useful summary evaluation. However, we also wished to capture information on particular aspects of a course which might be of special interest to applicants. It is very difficult to do this without preparing a questionnaire for each course, and the final section is an attempt to work round this difficulty. The proposed scales and items are presented in Figure 5.1. In the Australian CEQ, respondents are asked to indicate their agreement with each item on a scale of 1 (strongly disagree) to 5 (strongly agree). The same scoring is used in the proposed questionnaire but the intermediate stages are made explicit, (disagree/neither agree nor disagree/agree). We do not recommend that options such as 'don't know' or 'not applicable' are provided. Experience indicates that these categories will attract responses when a positive response would be more accurate. In line with standard practice, the questionnaire contains positive and negative statements in order to encourage respondents to consider each item fully. Finally, the actual questionnaire randomises the order of the items.

Figure 5.1: Proposed questionnaire

1. *Good Teaching*

1.1. Lecturers and tutors motivated me to do my best work
1.2. Lecturers and tutors put a lot of time into commenting on my work
1.3. Lecturers and tutors were good at explaining things
1.4. Lecturers and tutors made the subjects interesting
1.5. Lecturers and tutors made a real effort to understand difficulties I experienced with my work
1.6. Lecturers and tutors normally gave me helpful feedback on my progress

2. *Generic Skills*

2.1. The course developed my problem-solving skills
2.2. The course sharpened my analytical skills
2.3. The course helped me develop my ability to work as a team member
2.4. As a result of my course, I feel confident about tackling unfamiliar problems
2.5. The course improved my skills in written communication
2.6. My course helped me to develop the ability to manage my own work

3. *Clear Goals and Standards*

3.1. It was often hard to discover what was expected of me in this course
3.2. It was always easy to know what standard of work I was expected to achieve
3.3. Staff made it clear from the start what they expected from students

4. *Appropriate Workload*

4.1. For most of the course, the workload was too heavy

[21] A similar section was added by the LTSN for hospitality leisure, sport and tourism to their CEQ.

4.2. I was generally given enough time to understand the things I had to learn

4.3. I felt a lot of pressure, from lecturers and tutors, to do well in the course

4.4. The sheer volume of work meant it could not all be thoroughly comprehended

5. *Support and Advice*

5.1. Over the whole course, I was given sufficient support with my studies by lecturers and tutors

5.2. During my first year, I needed more support and advice with my studies than I received

5.3. I found other students helped me with my studies during the course

5.4. There was good advice available on which course options were best suited to my needs and interests

6. *Learning Resources*

6.1. Availability and access to library resources were appropriate for my needs

6.2. I was not able to access IT resources to the extent I needed to

6.3. Course materials (lecture notes, work sheets, CD-ROMs etc) were useful

6.4. There were insufficient opportunities to apply the theoretical knowledge I acquired to practical situations

6.5. When it was needed, there was sufficient access to specialised equipment or facilities

7. Overall, I was satisfied with the quality of this course

Please tick the three aspects of your course which were best and the three which could be most improved

	Best	**Could be improved**
Acquiring knowledge of the subject		
Acquiring generic skills such as problem solving, team working and communication		
Quality of the teaching staff		
Access to specialised equipment		
Availability and access to learning resources such as the library, course materials and IT		
Advice and support from staff on academic issues		
Opportunities to test theoretical knowledge in a practical situation		
A workload which was appropriate to the time and resources available to students		

5.64 It is important to maintain the integrity of the scales if the questionnaire is to provide meaningful information. We are advised that this requires retention of the core set (1, 2 and 7) and no more than minor changes in wording to items. In particular, it may be dangerous to add or substitute items within scales. We have followed this advice with respect to all scales apart from support and advice (5) and learning resources (6). These are two of the newly introduced options in Australia and the versions in Figure 5.1 reflect the discussions held with students. It will be important to test the robustness of these items and scales during the first run of the National Survey.

5.65 We are proposing that all the components in Figure 5.1 be included in the National Survey. It would still be possible for individual HEIs to add their own questions to the survey and this may be highly attractive to some. However, there may also be some difficulties. It will add to the length of the questionnaire, and may therefore reduce response rates. In addition, there is a possibility of compromising the psychometric properties of the questionnaire. This danger could be minimised by placing the additional questions at the end, but with a postal questionnaire there can be no control over the order in which questions are considered. The impact of optional questions therefore needs to be investigated during the first run of the National Survey.

5.66 Finally, we note that although this project is only concerned with undergraduate courses, the proposed questionnaire could also be applied to taught masters. We have not, however, explored whether there is a demand for such information from those applying to masters programmes. The issues surrounding research degrees are very different and we believe a radically different questionnaire would be required. Indeed, there is evidence that the CEQ is a poor indicator of teaching quality on research degrees.

Reporting

Analysis of responses

5.67 A CEQ type questionnaire will generate responses, on a five point scale, to each of the individual items. These can be averaged to give a score for each of the six scales in one of two ways:

- summation of the percentage responding 'agree' or 'strongly agree' on each item and averaging to obtain the scale score

- calculating means: with 'strongly disagree' assigned a value of -100, 'disagree' -50 through to 'strongly agree' 100 and then averaging as above.

5.68 In both cases, scoring of negative statements – *I did not receive adequate….*can be reversed. In practice there appears to be little to choose between the two types of measure. Although their properties are different, correlation between Australian CEQ scale scores and the corresponding agreement percentages have been in the range of 0.85 to 0.9 since the survey began.

5.69 This procedure will give a simple score for each of the six scales, and the overall satisfaction item which could be reported as it stands. However, the absolute value of this score has no

intrinsic meaning and could be misleading. If, for example, a subject at one HEI scores 30% more than the same subject at another this does not imply that teaching quality at the former is 30% better. One solution might be to rank according to scores but this may also cause problems since small differences in scores could mean large differences in the rankings.

5.70 For these reasons, we recommend that institutions are assigned to a group (on each scale) after an analysis of the distribution of scores on the basis of their range and variation. There is no reason why the groups need be the same for all subjects. If, for example, scores approximated to a uniform distribution with institutions evenly spaced throughout a large range then a large number of groups would be appropriate. If, however, scores clustered around a few points then fewer groups would be needed. Reporting in this way would mean that it was not possible to identify institutions which had improved in an absolute sense but not relatively, but it is relative positions which will be of most interest to applicants. It would also make it impossible to compare UK and Australian HEIs (and other countries which might adopt the CEQ). Some we consulted felt this would be desirable since they considered underlying differences between the UK and Australia would render such comparisons invalid.

Levels at which information should be reported

5.71 As was discussed at the beginning of this chapter, applicants require information at a lower level of aggregation than the institution as a whole. In some ways, programme level data would be the ideal but the volume of information this implies, and the real possibility of very few responses, make this impractical. We recommend that information is reported at the subject level and this raises the questions of which classification and which levels of the chosen classification.

5.72 We recommend that JACS is used:

- much thought and effort has gone into its design and it is becoming established as the classification system. Much of the other data identified for publication by the Cooke Report will also be classified on JACS

- it is used by the Universities and Colleges Admissions Service (UCAS), so that applicants who apply through this route will find it relatively easy to associate relevant National Survey information with the programme(s) they are considering.

5.73 The top two levels of JACS comprise 19 and almost 150 subject groups respectively. During the last round of discussions with students they were shown the 19 subject groups and many felt these were two highly aggregated in some cases. The main problem is perhaps with the

engineering group which includes all engineering subjects (civil, electrical, manufacturing etc). But there are also difficulties with some of the other groups. For this reason we recommend that information be presented at the next level down as well as at the 19 group level. However, this is subject to the condition that a minimum number of responses are received at this level. Following the discussion above, this minimum could be as small as 10 responses. This would result in a substantial volume of information, but the National Survey data will be made available via the web and users can thus drill down from the 19 subjects group level to more detailed levels as and when they require without becoming overwhelmed by information.

5.74 There is a question as to whether averages for the institution as a whole should also be provided. Neither students nor other stakeholders expressed interest in information at this level and we see little point in reporting it. There is also an argument that publication at this level would feed directly into league tables which are considered to be misleading.

Presentation

5.75 The Cooke Report recommended that National Survey information be made available on the web, hosted by Higher Education and Research Opportunities (HERO), and this would be the preference of most students we interviewed. This has the great advantage of allowing users to drill down through varying levels of detail and also to select user-defined tabulations. We envisage three levels from the home page: About the survey; How the information is presented; Student feedback data.

About the survey

5.76 This section would provide an overview of the purposes of the survey and how it was conducted. The headings would be:

- what is it seeking to measure
- how it has been conducted
- what it can tell you
- what it cannot tell you

- other sources of information to help you decide where to study (mainly links to other sites).

How the information is presented

5.77 This section would explain the six scales and the overall satisfaction item. It would also describe the JACS classification system and why information might not be available at the lower level classification for all subjects.

Graduate feedback data

5.78 The data itself would be accessible in the ways described below, but we envisage a common format for all data presentations. This is illustrated in

Figure 5.2. It is assumed that HEIs would be grouped in the way described above, if not the cells would contain actual scores for each of the scales. In addition to the CEQ score, the numbers of responses and the total number graduating would also be given. Responses to the questions on three best aspects of the programme and the three which could be improved (see figure 5.1) would be reported, in a single table, as the percentage which specified a particular aspect. This would also be disaggregated by subject and absolute numbers responding given.

5.79 Two sorts of hyperlinks would be provided:

- from the HEI cell to the HEI website. This could either be the HEI's home page or a special page where the HEI has chosen to provide a commentary on the CEQ scores. Alternatively the link could be to the relevant department

- if the data refers to one of the subjects at the 19 subject group level, a link to the next level down for that subject area. The user would be taken to a page listing the more detailed subjects and selecting one of these would bring up the same HEIs with scores for the lower level subject (providing the minimum number of responses had been generated).

Figure 5.2: Data format

Subject:

	Group to which assigned on basis of CEQ scores (1 highest)								
	Teaching	Generic skills	Goals & standards	Workload	Support & advice	Learning resources	Overall satisfaction	Number of responses	Total number graduating
HEI									
A									
B									
C									

5.80 The user would be able to access data in two main ways. First, by browsing. It should be possible to start with any HEI and then select a specific subject, or to start with a subject area and then select HEIs. Second, by specifying search criteria. The main search criteria which are likely to be of interest are:

- selecting the top (say) five HEIs in a given subject area
- selecting subjects in a specific geographic area
- selecting HEIs according to their entry requirements in specific subject areas. This could be useful to many applicants whose choice will be restricted by their exam grades. The first two requirements are easily handled but this is more problematic since entry requirements relate to programmes rather than subjects. If, however, programmes are mapped onto subject areas then it would be possible to indicate whether there are any programmes offered in a given subject area for which the entry requirements are less than or equal to the search criteria.

5.81 Earlier in this chapter we rejected the idea of different questionnaires for different types of student, for example mature, part-time and disabled. However, this does not mean that responses from specific groups cannot be reported separately. We doubt whether the number of responses would be sufficient to report on a subject basis but they are likely to be sufficient for the institution as a whole. This information could still be useful since many of the issues which specifically affect these groups are likely to be institution-wide rather than subject-specific.

5.82 Finally, we think there would be merit in making (aggregate) responses to individual items (as well as scales) available on the website. There is likely to be academic research interest in the National Survey data and engaging the research community will facilitate its future development.[22]

Management

5.83 The National Survey will need to evolve in a similar way to the Australian CEQ. In particular, the value of specific questions, and the ways in which information is presented,

[22] The hardcopy publication of the CEQ is restricted in the detail it provides but the basic data is available via the web - http://www.avcc.edu.au/students/gradlink/GCCA/. The hardcopy publication only reports results where there was at least a 50% response rate but the web data gives all information irrespective of response rates.

need to be monitored and amended as appropriate. This strategic process needs to be overseen by a Steering Group which we would expect to comprise, inter alia, representatives from HEFCE, QAA, Universities UK, SCOP, NUS and HESA and also the other HE funding councils should they decide to participate in the survey. We would also recommend a technical group be established which would report to the steering group. This group would be concerned with the detailed aspects of the National Survey and would require expertise in large scale survey techniques. Its role would be to advise the steering group on the conduct of the National Survey, as it evolves, and we would envisage it having day-to-day responsibility for managing any contracts with external organisations.

5.84 During the first year of operation there are likely to be special demands on National Survey management. We have drawn on the Australian CEQ because it has been, at least partially, validated through extensive testing and analysis of the responses. We see no reason for believing that the questions will not translate to the English context but this needs to be verified. In addition, new questions have been added and their impact and validity needs to be assessed, which can only be done in the context of a large scale survey.

5.85 For this reason, we recommend the first year of operation be treated as a pilot survey. The coverage need not be reduced, but it would be structured so that various aspects of the survey could be tested. There is no a priori reason why the pilot should not generate robust results which can be reported.

5.86 There are in fact two stages to the pilot:

- developing the questionnaire prior to a large scale survey
- large scale testing, the results of which would feed back into questionnaire development, but during which various survey management issues would be explored.

Initial development of the questionnaire

5.87 The main aim of this stage is to ensure that the statements in the questionnaire are comprehensible and that the questions can be answered. We have a high level of confidence that this is the case with the proposed questionnaire. As already mentioned, it has been used extensively in Australia and more limited testing has also been undertaken during the current project. Nevertheless, the cost of such testing is very low in comparison to that which will be incurred in a large scale survey, and the importance of getting the questions 'right' is such that we recommend further testing.

5.88 Piloting is the traditional means of testing the questionnaire but it is not the only way. Research has shown that expert review can be every bit as valuable as piloting in establishing weaknesses in questionnaires and we suggest expert review be undertaken, by those who are expert in questionnaire design rather than those who are expert in educational assessment.

5.89 One of the best ways of piloting self-completion questionnaires is through cognitive assessment. There are two main forms of cognitive assessment, *full think aloud* and *post interview discussion*. With full think aloud, the respondent is asked to complete the questionnaire but to vocalise all the thoughts that are going through their mind while they are considering how to answer it. With post interview discussion, respondents complete the questionnaire on their own and specially trained interviewers then take them back over key aspects of the questionnaire asking what was going through their mind when they answered particular questions or asking them what they thought the terms or phrases meant to them.

5.90 The former is a far more complex and expensive operation and is probably not necessary in a survey of this kind, instead we suggest that post interview discussions are undertaken. For a survey such as this, some 20 or 30 cognitive interviews would be suitable given that students could be recruited on campus. It would be possible to conduct these interviews in a single day but it would be preferable if they were spread over more than one institution.

Large scale testing

5.91 There are two purposes of large scale testing. The first relates to further development and testing of the questionnaire itself. Merely because respondents understand the concepts does not necessarily mean that the answers they give are measuring the constructs that the questionnaire is supposed to be measuring. Testing this is only possible with large volumes of data as it requires statistical analysis of the patterns of answers, and large numbers are needed to assess statistical significance. There is no right number for conducting analysis of this sort but several hundred responses would be needed to promote any meaningful analysis. The key tests which need to be undertaken are:

- *Reliability* – are consistent responses given to the same (or similar) questions. This can be tested by either asking the same person the same question at different times, or asking a series of different questions at the same time, intended to tap the same attitude. Both approaches have complications. The former could be affected by changes over time in attitudes. The latter could be affected if a respondent identifies there are several questions about the same thing and could either give the same answer to all in order to appear consistent, or could think the survey wouldn't ask the same question twice, and therefore tries to find subtle differences between them

- *validity* which refers to the accuracy with which the question generates information on the construct of interest, in this case quality. It encompasses *correlational* validity – the extent to which a given measure can predict other measures to which it should be related; and *discriminant* validity – the extent to which an approach can differentiate between constructs that are presumed to be different from each other

- the use of factor analysis to see if any items can be dropped from the questionnaire because they are duplicating each other

- the scope for institutions to add their own questions to the National Survey without distorting the results.

5.92 The second purpose of large scale testing is to explore, and refine as appropriate, management processes. The key considerations are:

- following up non-respondents to test whether their views differ systematically from those of respondents. During the pilot, special efforts should be made to elicit responses from non-respondents if at all possible. An analysis of characteristics which might be correlated with perceptions of quality, for example degree class, should also be undertaken

- the response rates which can be achieved and whether there might be scope to restrict the National Survey to sampling, as opposed to a census, in some subject areas

- the extent to which the timing of the survey influenced response rates and whether perceptions of quality also varied according to timing

- the potential for web based surveying. This could be tested by restricting some respondents to on-line entry

- continued exploration of the feasibility of the entire survey being administered by a single organisation. This depends on resolution of the issues relating to the Data Protection Act and would need to be pursued in parallel with other student/graduate surveys. It will entail discussions between HEFCE and HEIs (and others such as Universities UK) but it may also be possible to test attitudes and approaches during the pilot. The National Survey could, for example, offer to relieve individual HEIs of any involvement in the survey provided they agreed to divulge student contact information.

5.93 Large scale testing would require several hundred responses but it need not be a full census. One option would therefore be to run a pilot survey covering a subset of institutions and/or subject areas. This would be significantly cheaper than a census-based pilot, yet would still enable a range of approaches and options to be tested. However, it would delay the introduction of a census-based survey and therefore mean a delay before results covering all institutions could be published. In order to ensure that results are published as soon as possible, a full census survey could be undertaken after the initial development of the questionnaire (paragraph 5.87). This would still be in the nature of a pilot in that the various tests described in this section would be undertaken and the questionnaire, and process, could be modified for subsequent surveys if appropriate. However, if the tests indicated that responses were robust and meaningful then results from the pilot could be published. The disadvantages are that the census-based pilot would be more expensive than a sample-based exercise and there would also be less scope to test options for survey management

National Survey costs

5.94 There are many elements of survey design to be decided so it is only possible to give broadly indicative costs at this stage. There will be certain fixed costs associated with the National Survey, for example the piloting described above, liaising with HEIs, analysis and reporting, and management. However, we estimate that these will be minor in comparison to the variable costs of printing and postage incurred in a large scale postal survey. The main determinant of costs is therefore the number of questionnaires mailed out.

5.95 Table 5:1 shows the number of graduates in 2001, from English HEIs, with a first degree or other undergraduate qualification. There were almost 300,000 but this includes more than 44,000 with combined subjects. A decision will need to be taken on how to treat graduates taking combined subjects. They are too numerous to be ignored and it would not be feasible to report additional categories for combinations of subjects. The Australian CEQ handles this issue by requiring combined subject graduates to complete a separate questionnaire for each subject separately, and we recommend that the National Survey also adopts this procedure. This will mean that the opportunity to get feedback on synergies between the combined subjects is lost, but we see no alternative if these graduates are to be included in the National Survey. This means that some costs, mainly printing and postage, will be double for combined subjects. In addition, the government targets for HE imply significant expansion and we therefore have based cost estimates on a graduate population of 350,000.

Table 5:1: Graduates with undergraduate qualifications (England, 2001)[23]

[23] Source: HESA.

Subject	First degrees	Other undergraduate	Total
Medicine & dentistry	4,700	65	4,765
Subjects allied to Medicine	15,585	22,325	37,910
Bio sciences	14,655	1,070	15,725
Veterinary science	425	10	435
Agriculture and related subjects	1,805	1,100	2,905
Physical sciences	10,880	790	11,670
Mathematical sciences	3,625	245	3,870
Computer science	11,105	4,725	15,830
Engineering & technology	17,000	3,745	20,745
Architecture, building and planning	4,840	1,790	6,630
Social economic & political studies	19,185	4,505	23,690
Law	8,620	585	9,205
Business and administrative studies	26,425	7,695	34,120
Library and information science	3,920	430	4,350
Languages	14,125	2,030	16,155
Humanities	7,935	1,335	9,270
Creative art and design	19,675	2,710	22,385
Education	9,710	4,655	14,365
Combined	29,895	14,435	44,330
Total	**224,120**	**74,240**	**298,360**

5.96 Table 5:2 shows the estimated costs of printing, postage and data processing for the survey. It assumes that a first reminder is sent to 65% of the target group and a second reminder to 55%. The total costs would be £634,000. Some surveys adopt the practice of sending a postcard to the entire target group, soon after the initial mailing, urging them to complete the questionnaire. However, we estimate that this would add in the region of £130,000 to costs and would not be justified by the likely increase in response rates.

Table 5:2: Annual survey costs of National Survey	
Basic set-up of sample etc	£10,000
Initial mail out	£212,000
1st reminder to 65%	£139,000
2nd reminder to 55%	£118,000
Return postage and data processing	£155,000
Total	**£634,000**

5.97 If the actual mail outs were handled by individual HEIs, but data processing contracted to a single organisation, we estimate the costs of data processing (including the return postage

costs for completed questionnaires) would be in the region of £155,000. However, the postage and printing costs would still be incurred.

5.98 As mentioned above, these costs do not include an allowance for full set up and management activities. Also, they do not allow for the staff costs incurred by individual HEIs in organising the mail out. These will chiefly relate to assembling a database of target groups' contact details, in response to information provided by HESA, and mailing out the questionnaire and reminders. They will vary according to size, but also by internal IT systems and organisational structures, and we are not able to provide meaningful estimates on the basis of information currently available.

5.99 These costs could be reduced in two ways. First, the National Survey could be a sample rather than a census survey. Whether this is feasible or not depends on the number of responses which are likely for a given subject and HEI (and also how far responses are disaggregated below the 19 subjects JACS level). This in turn will depend on the distribution of graduates between institutions. As was discussed above, we recommend that the scope for a sample survey is tested during the first run of the National Survey, but we think it likely that some subjects at some HEIs will be large enough to justify a sample. However, if subjects and HEIs are to be differentiated in this way it will inevitably add to the management costs of the survey, especially if the mail out is undertaken by individual HEIs rather than a single organisation.

5.100 The second way to reduce costs is to undertake the National Survey less frequently than annually. We believe there is some merit in this. Students we consulted felt that information needed to be up-to-date but a biennial survey could be adequate and would approximately halve costs. If this approach was adopted then half of subjects could be covered one year and the remainder the next. However, we would note that a biennial survey would provide misleading information on institutions which had been subject to rapid change in the quality of provision.

6 Conclusions and recommendations

Institutional processes for collecting and using student feedback

6.1 Virtually all HEIs possess quite elaborate mechanisms for the collection of student feedback information. While there are considerable variations in detail, all institutions use a range of mechanisms, both quantitative and qualitative. There is considerable variation in the detail of questionnaire design but considerable commonality to the topics covered. A number of institutions have introduced a degree of standardisation to their questionnaires while allowing faculties and departments some discretion to add or indeed modify content. The most common focus of questionnaires is the individual module although a majority of institutions also gather feedback data at other levels.

6.2 Many institutions have some kind of central unit with responsibility for student feedback although in practice much is often devolved to faculties or departments. Even where there is a central unit, analysis of feedback data is often quite limited, with little use of comparative analysis or relating feedback data to other institutional datasets.

6.3 In many institutions, more use could be made of feedback data. This would require additional resource but consideration might be given to collecting less data and analysing it more thoroughly. It may not be necessary to administer a module questionnaire every time the module is offered.

6.4 Greater analysis and more imaginative presentation of feedback data might encourage more use to be made of it, which in itself would increase the commitment of staff and students to the importance of feedback processes. We detected that this commitment was slight in some places.

6.5 One way of making greater use of the data is to enable students to take it into account in choosing modules.

6.6 Whether it is being used for internal or external purposes, it is important to place data in context, to present it in the light of other forms of data on quality and standards and with reference to the distinctive features of the institution. If this contextualisation of information is at the expense of some direct comparability, this is preferable to the publication of data that is misleading or meaningless.

6.7 Although we believe that it would be possible and in principle desirable for institutions to publish summary information, appropriately contextualised, of student feedback information at programme or equivalent level, we do not believe that such a requirement should be made mandatory. There are genuine concerns within institutions about the effects of publication upon the quality enhancement role of feedback and about the resources that would be required to produce good quality publications that would genuinely be useful to prospective students. However, some institutions might well wish to publish if their feedback systems can provide information in a suitable form, if the needs of quality enhancement can be protected, and if the resource can be made available. The purpose of publishing such information would be to increase the prospective student's knowledge of the likely student experience at *that* institution, not to rank order all institutions in the land. Therefore, enunciation of context and use of common themes are more important than direct comparability of numerical data.

6.8 This study has confirmed that HEIs are devoting considerable resource to obtaining feedback from their students. Some of the good practice discovered by the present study will be included in the *Good Practice Guide* to be published by HEFCE later in 2003. However, we believe that there remains a need to do more to share experiences and good practice both within and between institutions, especially with regard to the analysis, presentation and uses of student feedback data. In some institutions, consideration may need to be given to whether more data is being collected than is needed and whether greater effort needs to be devoted to use of existing data than to collecting more of it. At the same time, the overall cost and burden of student feedback should be monitored to ensure that the value is commensurate to the costs.

6.9 Recommendations for institutional processes have been given in each of the sub-sections of Chapter 4. They are gathered together here for ease of reference.

Purposes

6.10 Clarity of purpose is key:

- there is a need for clarity about the differences between student feedback on satisfaction, on learning processes (study methods), or on student objectives and their achievement. The dangers are that questionnaires muddle these different kinds of feedback, although it is not impossible to combine them into a single instrument

- all involved in the collection and use of student feedback data need to be clear about the purposes and intended uses of the data. This is especially important for the students themselves if their commitment to the process is to be maximised. Thus, the

purposes and use to which the information will be put should be stated at the start of questionnaires

- consideration should be given to the range of alternative ways in which purposes can be achieved. For example, questionnaire fatigue among both students and staff is clearly a danger. There may be some potential for reducing the total burden by sampling or by collecting feedback in alternate years or only when other quality indicators have suggested cause for concern

- use of a range of feedback mechanisms will be more effective than reliance on questionnaires. For example, the existence of a well-publicised complaints procedure or a discussion during class may provide more effective ways of checking that a module is going well than asking students to fill out questionnaires

- the needs of users at different levels in the institution should be recognised at the time that data is collected. They will have implications for what is collected and when, and for what forms of aggregation might subsequently be required.

Mechanisms

- the mechanisms used should take account of the form of curriculum organisation, including the length of modules, and the numbers of students enrolled on modules, on programmes, in departments and so on

- reliance should not be placed on any one mechanism for collecting and using feedback

- reliance on informal feedback, while important, is not recommended as a sole mechanism for obtaining student feedback

- mechanisms used will also need to relate to purpose, which should be clearly stated and communicated to staff and students. Although traditionally related to quality and enhancement, additional purposes need to be recognised and accommodated.

Collection

- reliance should not be placed on a single mechanism for gathering feedback

- different mechanisms may be needed for different purposes, levels and contexts

- the purpose of collecting student feedback and how it will be used, including how results/actions will be disseminated to students, should be clearly stated in guidelines to staff and students, and especially at the point when feedback is being requested

- students and staff should be made aware of the benefits of gathering feedback, and the processes involved should be fully explained and understood by all parties. In addition, this will be most effectively done if students and staff find that they are using the results of feedback data e.g. by students in choosing their options, by staff in revising their modules/programmes, by management in planning new programmes, and by the marketing department in promoting the institution

- discussion groups are an alternative to questionnaires and student representation systems (although not necessarily a replacement). They can provide a rich source of information and their uses in certain specific circumstances should be considered.

Specific recommendations on questionnaires

- sufficient time should be allowed for students to complete questionnaires when given out 'in-class', their purpose and use should be fully explained, and students made to believe that their feedback is important and welcome

- if the module is the level at which feedback is collected, consideration should be given to frequency and/or sampling (of modules) to counter questionnaire fatigue

- consideration should be given to capturing student profile data as well as views and opinions to check how far responses vary between types of students

- questionnaires should be standardised (with a set of common questions) within institutions as far as possible, to provide a basis for comparison both within and between institutions. However, it is likely that there will always be some need for special questions to reflect different purposes and contexts, especially at module level

- if questionnaires are not completely standardised, a common core and limits to acceptable variation should be set

- response rates should always be published and where they are below, say, 60% the results should be treated with some caution, especially if not presented alongside other sources of information. This is not to say that information might not still be valuable but much will depend on the degree of local knowledge and the availability

of other information. However, whatever the response rate, it will be important to check the typicality of respondents (e.g. age, entry qualifications and so on)

- information on reliability and validity that can be claimed for the data should be provided to all users

- questionnaire feedback should not be used in isolation, but should take account of the existence of and messages from other forms of feedback

- questionnaires should include open-ended questions to provide students with an opportunity to raise issues not covered by fixed choice questions

- a system should be established to ensure that answers to open-ended questions are at least read by teaching staff, if not processed and analysed

- where questionnaires are administered 'in-class', efforts should be made to obtain responses from those students who are not present

- consideration should be given to using students in the process of distributing and collecting questionnaires 'in-class' to ensure against manipulation and distortion of results and to promote independence and confidence in the system

- where web-based and email systems are used to administer and collect questionnaire responses, every effort should be made to demonstrate to students that their responses are treated confidentially

- above all, the collection of feedback information must take account of its intended uses and the nature of the institutional quality assurance and enhancement procedures.

Specific recommendations on student representation systems

- the importance of the role of student representatives should be recognised by staff at all levels (i.e. not just senior management) and students, and this should be communicated to students

- consideration should be given to involving the Student Union in awareness raising and training in the student representative role

- where training for student representatives is provided by the Student Union, there should be full co-operation between staff at all levels and the Student Union to ensure that students are able to take advantage of the training

- agendas and other papers should be made available to student representatives in advance of meetings and, if necessary, a briefing session held prior to the meeting to discuss issues to be raised

- institutions, students and Student Unions might wish to consider the feasibility of rotating the role between students to share the experience

- time should be made available to student representatives to enable them to gather and feed back issues to the student body.

Analysis and interpretation

- the question of whether to adopt centralised or devolved systems for the analysis of student feedback data will need to reflect institutional structures and circumstances

- nevertheless, it would seem desirable to maintain some central resource, both as a centre of technical expertise and advice, but also to provide a capacity to undertake more sophisticated (especially comparative) analyses and to meet the needs of the institution's central authorities on matters of quality and standards

- analysis of feedback data by a central unit can, providing that the unit's independence is safeguarded, help protect against manipulation or distortion of results

- those undertaking the analysis, whether centrally or locally, should ensure that they are informed about the needs of the users of the data and the purposes of collecting it

- if feedback is to inform and improve the teaching and learning process, interpretation should reside with the teaching staff responsible, although such staff might reasonably be expected to summarise and to comment on feedback data for use elsewhere in the institution

- guidance and criteria should be set for such summaries

- feedback should be interpreted in context and with other sources of information that are available.

Actions and decision-making

- feedback to students is as important as feedback from students. Institutions need to ensure that students are told of the results of feedback and of any actions taken in response to it. When actions cannot be taken, the reasons need to be carefully conveyed to students and, if possible, their further reactions obtained

- virtually all institutions have annual monitoring and review cycles involving key academic committees that can provide an effective means for receiving and acting upon the results of student feedback. Their secretaries need to ensure that feedback data are presented in a digestible way, and chairs need to ensure that they are given due consideration in decision-making processes

- notwithstanding the important role of committees, certain key individuals also play important roles. Probably most important is the individual teacher whose response and actions can deal with many issues raised and convey to students the sense that their feedback is valued by the institution as a whole

- for student feedback to feed into more strategic aspects of decision-making within HEIs, it may need to be analysed and presented in different ways and here the role of a central unit is likely to be important. This is also true of qualitative feedback as recorded in committee minutes and reports. There is a danger that such information can be 'lost' to the decision-making process if not analysed, brought together and placed in context.

Presentation and publication

6.11 Recognising the real concerns that exist in institutions concerning the publication of student feedback data, we believe that external publication of such data should not be made a requirement at the present time. In principle, however, we believe that information derived from feedback data, suitably contextualised, could be published and that it would be useful to existing and prospective students and also to the institutions themselves in suggesting benchmarks and pointing to good practice. In many institutions this would require modification to existing arrangements for student feedback, including safeguards to protect quality enhancement functions. Quite reasonably, institutions will give different priority to taking such steps. Like other published data on quality and standards, any published information derived from student feedback will be subject to periodic audit by QAA.

6.12 However, while we do not recommend a requirement to publish feedback data externally at the present time, we do believe that improvements could be made in the ways that data is published internally.

6.13 Publication might include the following:

- student feedback on individual modules might be published within the institution in order to help inform module choice

- some institutions may wish to consider the publication of information derived from student feedback at programme or equivalent level on the institution's website in order to inform choices of prospective students. (Equivalent levels might be subjects or departments.) Feedback would need to be set in its institutional context and could be complemented by other kinds of information, for example retention rates and employment data

- if information is to be published (internally or externally), readers should be provided with information on response rates, reliability and validity and when the information had been collected. The chief consideration should be that information is meaningful to the reader and is not ambiguous or misleading.

Dissemination to students

6.14 In most institutions, actions need to be taken to improve the feedback to students. These might include:

- to encourage students (and staff) to take the process seriously, face-to-face feeding back of the 'highlights' of results/issues raised should be built into feedback processes

- additionally, results and actions of previous feedback can be added to module handbooks or discussed with students at the start of a module

- the timing of feedback (i.e. collection and reporting) will need to be considered if feedback to students is to include information on actions taken

- opportunity should be provided for student representatives to discuss with and report issues to the students they represent in teaching time

- action sheets from meetings should be prepared so that representatives know who is responsible for following through actions and are updated on progress

- student representatives might play an enhanced role in the administration of student feedback processes, including surveys.

The National Survey

6.15 We have little doubt that applicants to HE would find a well designed National Survey, which provided information at below the institutional level, useful. This reflects: the importance attached to programmes when deciding where to study; the fact that teaching quality is one characteristic by which programmes are judged; and the lack of independent and credible information currently available. There is, however, little evidence to suggest that the National Survey would contribute directly to quality enhancement. Most HEI staff we interviewed felt that it could add little to internal feedback mechanisms and there is quite widespread resistance to the introduction of such a survey.

6.16 There are many aspects of the HE experience which are relevant to applicants when selecting which HEI to study at. Our view is that the National Survey should focus on teaching and learning narrowly defined, essentially for pragmatic reasons:

- many of the other topics of interest are essentially factual questions and information can be accessed from prospectuses, open days and other material

- covering additional topics would add to the length of the survey instrument and, with the aim of maximising response rates, this should be kept as short as possible.

6.17 There are analogous issues with respect to the information needs of different groups of students. In particular:

- disabled students, where physical access and learning resources have a direct bearing on teaching quality

- mature students, some of whom may require additional support early on and access to facilities such as crèches

- part-time students, for whom the flexibility of provision, and opportunities to integrate with the student body, may be especially important.

6.18 Some of those consulted felt that special (additional) questions should be added to the survey instrument for these groups. While we can see merit in this we believe the National Survey should be restricted to a common questionnaire because:

- again, some of these issues are essentially factual questions, although we recognise that information on the quality of provision could also be important

- it would add significantly to the complexity of the National Survey since different groups would have to be identified in advance

- even if specific questions are not asked, it would be possible to analyse responses according to student group. The numbers responding from a given programme/subject area could be too small to report with any confidence but it might be possible to report at the level of the institution. This may be adequate for many of the specific issues the groups face

- the information needs of special groups are probably more appropriately addressed through institutional surveys, and this could be a subset of the information which institutions themselves publish.

6.19 There are real methodological issues surrounding student feedback on teaching quality. We recognise the importance of these, but believe that there is sufficient evidence to suggest that the CEQ, currently used in Australia, has addressed these successfully. We recommend that the National Survey instrument should be based closely on the CEQ.

6.20 The key considerations for the survey timing are: when will graduates be able to give meaningful responses; and when might response rates be maximised. The Cooke Report suggested that the National Survey might collect information on the value of the programme to subsequent employment, which means the survey would be delayed until some point after graduation. If the survey is to cover both views on teaching quality and value to subsequent employment, its timing will inevitably represent a compromise. The need to ensure that memories of the teaching experience are fresh implies an early survey. An early survey will also enhance response rates, since contact details for graduates become less reliable over time. The need to collect employment related information implies a survey some period after graduation so that graduates have had the opportunity to enter employment and evaluate the usefulness of their programmes.

6.21 We recommend that the National Survey should not collect information on employment related issues. If this is accepted, it offers the scope to bring the National Survey forward in

time, simplifying administration and, more important, providing the opportunity to maximise response rates.

6.22 There are three main options for the timing of the survey:

- after the degree results are known
- after exams but before results are known
- before exams.

6.23 The National Survey is intended to cover all HE provision up to and including first degree programmes. We recommend the survey should be census-based and a postal survey. The main reason for a census approach is the need to report results by institution and also by subject area. We think it likely that the number of responses in some 'cells' would be unacceptably low if the National Survey was sample based. The survey will, however, be large and the costs of a telephone interviewing would be substantially greater. More important, given the purpose of the National Survey and the nature of the questions, it is not clear that telephone interviewing would add anything to a postal questionnaire.

6.24 We think there would be merit in all aspects of the survey (mail out, analysis and reporting) being contracted to an independent organisation, but this may not be immediately acceptable to the HEIs because of concerns relating to the Data Protection Act. We therefore recommend that, initially, questionnaires are mailed out by the institutions but data is inputted and analysed by a single contractor. The only additional demand on HEIs would be the mail out of questionnaires and reminders. They would not be expected to follow-up non-respondents by telephone as with the FDS.

6.25 A CEQ type questionnaire will generate responses, on a five point scale, to each of the individual items. These can be averaged to give a score for each of the six scales. The absolute value of this score has no intrinsic meaning and could be misleading. We therefore recommend that institutions are assigned to a group (on each scale) after an analysis of the distribution of scores on the basis of their range and variation. There is no reason why the groups need be the same for all subjects.

6.26 We recommend that information is reported at the subject level using the JACS. If the number of responses is sufficiently large then information should be reported at both the 19 subjects and the next level down. The lower level of reporting will, almost certainly, not be feasible for some subjects at some HEIs.

6.27 There is a question as to whether averages for the institution as a whole should also be provided. Neither students nor other stakeholders expressed interest in information at this level and we see little point in reporting it. There is also an argument that publication at this level would feed directly into league tables which are considered to be misleading.

6.28 The Cooke Report recommended that National Survey information be made available on the web, hosted by HERO, and this would be the preference of most students we interviewed. This has the great advantage of allowing users to drill down through varying levels of detail and also to select user-defined tabulations. In addition to the CEQ score, the numbers of responses and the total number graduating would also be given.

6.29 There should be hyperlinks from the National Survey site to HEI websites, either to the HEI's home page or a special page where the HEI has chosen to provide a commentary on the CEQ scores. Alternatively the link could be to the relevant department. Users should be able to browse the site, by subject or HEI, and also define searches, for example by geographical region and possibly also entry requirements.

6.30 Responses from specific groups of graduates, for example part-time, disabled, mature, could be reported separately. We doubt whether the number of responses would be sufficient to report on a subject basis but they are likely to be sufficient for the institution as a whole. This information could still be useful since many of the issues which specifically affect these groups are likely to be institution-wide rather than subject-specific.

6.31 The National Survey will need to evolve and this strategic process needs to be overseen by a steering group. We would expect it to comprise, inter alia, representatives from HEFCE, QAA, Universities UK, SCOP, NUS and HESA, and also the other HE funding councils should they decide to participate in the survey. A technical group, reporting to the steering group, should also be established.

6.32 We recommend that various aspects of the National Survey should be tested through a pilot exercise. This would comprise:

- initial testing and development of the questionnaire, which could be achieved quickly with a small sample

- large scale testing to further validate the questionnaire and explore a number of issues concerned with the conduct and management of the National Survey. This would require several hundred responses.

6.33 One option would be to run a pilot survey covering a subset of institutions and/or subject areas. This would be significantly cheaper than a census-based pilot, yet would still enable a range of approaches and options to be tested. However, it would delay the introduction of a census-based survey and therefore mean a delay before results covering all institutions could be published. In order to ensure that results are published as soon as possible, a full census survey could be undertaken after the initial development of the questionnaire. This would still be in the nature of a pilot in that the various tests described in this section would be undertaken and the questionnaire, and process, could be modified for subsequent surveys if appropriate. However, if the tests indicated that responses were robust and meaningful then results from the pilot could be published. The disadvantages are that the census-based pilot would be more expensive than a sample-based exercise and there would also be less scope to test options for survey management.

6.34 There will be certain fixed costs associated with the National Survey. However, we estimate that these will be minor in comparison to the variable costs incurred in a large scale postal survey. The main determinant of costs is therefore the number of questionnaires mailed out. On the basis of 350,000 graduates from English HEIs each year we estimate the annual costs of a census-based National Survey would be in the region of £634,000. This would include the costs of:

- printing questionnaires
- mail outs to graduates, including two reminders to non-respondents
- data processing, which we estimate at around £155,000.

6.35 These estimates do not include full set-up or managerial costs, nor the staff costs HEIs would incur in the mail out. These costs could be reduced in two ways. First, the National Survey could be a sample rather than a census survey. Second, the National Survey could be undertaken less frequently than annually. We believe there is some merit in this although the National Survey would not be able report accurately on HEIs where the quality of provision had changed rapidly for some reason.

Relationships between institutional processes and the National Survey

6.36 We have considered possible overlaps and synergies between the National Survey and institutional processes. These might occur in two sorts of ways. First, the National Survey

might substitute for some internal surveys which are currently undertaken. In practice, however, there is comparatively little scope for this for a number of reasons:

- the vast majority of institutional surveys are of current students rather than graduates, since there is a need to collect fairly immediate feedback on specific modules or units. The National Survey could not substitute for this information

- many of the HEIs consulted are keen to retain, or develop, survey instruments which they consider reflect their local circumstances. The National Survey instrument must, however, be consistent across HEIs, although there might be some scope for individual HEIs to add questions to the National Survey.

6.37 The second way is that the National Survey might provide additional information which could inform quality enhancement processes. The National Survey may have a direct impact in this respect in that a poor performance in the National Survey might stimulate an HEI to address teaching quality issues. However, the majority of staff in HEIs were sceptical about the value of the National Survey to them. The view tends to be that the National Survey would be at too general a level to identify specific issues that need to be addressed. There were also concerns over the timeliness of the information provided (although we believe this could be addressed) and, in some cases, the validity of the information provided.

6.38 These views, in part, reflect the fact that the National Survey does not yet exist, and also nervousness over the misleading use of National Survey data in league tables. It may well be the case that, if and when the National Survey became established and its validity tested, HEIs would identify ways in which the information could feed into their quality enhancement procedures.

Abbreviations

AOC	Association of Colleges
CEQ	Course Experience Questionnaire
CHERI	Centre for Higher Education Research and Information
CPQ	Course Perceptions Questionnaire
FDS	First Destination Survey
GCCA	Graduate Careers Council of Australia
HEFCE	Higher Education Funding Council for England
HEI	Higher education institution
HEQC	Higher Education Quality Council
HERO	Higher Education and Research Opportunities (web portal)
HESA	Higher Education Statistics Agency
JACS	Joint Academic Coding System
LSC	Learning and Skills Council
NUS	National Union of Students
OMR	Optical mark reader
QAA	Quality Assurance Agency for Higher Education
RAE	Research Assessment Exercise
SCOP	Standing Conference of Principals
SEEQ	Students' Evaluation of Educational Quality
SET	Students' Evaluation of Teaching
UCAS	Universities and Colleges Admissions Service
UUK	Universities UK