Strathprints Institutional Repository

Buckley, Alex (2013) Engagement for Enhancement : Report of a UK Survey Pilot. [Report] ,

This version is available at http://strathprints.strath.ac.uk/58240/

Engagement for enhancement
Report of a UK survey pilot

The Higher Education Academy

Dr Alex Buckley

Contents

Foreword by Professor Stephanie Marshall

Student surveys are one of the most powerful tools we have for understanding higher education from students' points of view. They can never replace informal conversations between teacher and student, but properly contextualised the results can tell us a great deal. And properly embedded in the processes and practices of enhancement, survey data can help to make real improvements to learning and teaching.

In the UK our surveys have tended to focus on how satisfied students are, or how good they perceive aspects of their course to be. That is understandable, especially given the amount that students are now expected to pay towards their education, but it does miss out what is widely thought to be the biggest factor in their learning outcomes: their level of engagement with their studies.

In the US, the National Survey of Student Engagement (NSSE) has been used by over 1,500 colleges and universities since its inception in 2000. It has an enviable weight of research behind it, both in terms of the survey itself and the use of the data, and makes powerful claims about being a proxy for learning gain, perhaps the key dimension of educational quality. NSSE has been used in a widespread way in a number of countries around the world, such as Australia, New Zealand, South Africa and – most recently – Ireland.

Up until now, the UK has lagged behind, with use of NSSE being restricted to individual institutions working largely in isolation. The project described in this report is the first attempt to gather data from multiple institutions. Not only will this help the participating institutions to benchmark and interpret their results, it will also help us to understand how well the survey works in the UK context. We cannot assume that the NSSE findings from the US (and elsewhere) can be transposed to our HE sector, they need to be evaluated in our own distinct context.

This project has coincided with a review of the National Student Survey. The debate about the role of student engagement in the NSS is timely and important and we at the HEA have an open mind; we see the powerful benefits that come from refocusing the NSS on the amount and quality of effort that student invest in their studies - a key determinant of how much and how well they will learn - but we recognise that there are complex questions about how well engagement sits with the intended purpose of the NSS; a purpose that is itself being debated.

I hope and expect that this report will promote further discussion about the benefits of engagement surveys, particularly for the enhancement of learning and teaching, and help to inform and shape that discussion.

Professor Stephanie Marshall
Chief Executive
The Higher Education Academy

Foreword by Professor Mantz Yorke

Students' engagement with their studies is, self-evidently, vital to their success. What institutions do, pedagogically, influences student engagement – but much of that engagement takes place away from lecture theatres, laboratories and studios. Teaching staff are thus likely to be somewhat in the dark about what students actually *do* as they work through their study programmes, and hence be left uncertain about the best ways to adapt their pedagogic approaches in order to increase the chances of their students' success. Surveys focusing on students' academic behaviour can assist staff in pedagogic innovation.

The *Goals 2000: Educate America Act*, implemented under the Clinton administration in the US, included an emphasis on critical thinking, communication and problem solving among graduates from higher education, and led to the desire for relevant indicators to be produced. Process indicators – indicators of what institutions do to facilitate these developments in students – offered the prospect of providing useful data that had the added advantage of being produced well before students graduated. Importantly, such indicators needed to be contextualised if they were to be of practical use. In other words, institutional types and settings, subject disciplines and year of study would need to be taken into account.

The National Survey of Student Engagement (NSSE), which has run for more than a decade in the US and has spawned analogues elsewhere (notably Australasia), arose in the light of policy discussion regarding goals and their measurement, and focuses explicitly on students' academic behaviour as they work through their programmes. It has been validated as a robust instrument, but this has not precluded further developmental work on it where this has been deemed necessary. The data from NSSE are most meaningful when they are interpreted with reference to the local context, and offer plenty of potential for enhancement-related activity. The institutional commitment to ensuring students receive a high quality experience in higher education in the UK (accentuated by governmental policy statements and changes to fee regimes across the four UK nations) provides a fertile ground for the use of NSSE items (amended as appropriate) in support of enhancement. This publication shows what is currently being done in this respect.

The work reported here is running in parallel with the use of a 'belongingness' and engagement survey as part of the *What Works? Student Retention and Success Programme* that is supported by the Paul Hamlyn Foundation and the Higher Education Academy. The *What Works?* survey is focusing on the *attitudes* of students towards engagement (rather than on their academic behaviour), and the intention is to collect data at four points during their time in higher education. A survey of attitudes allows the use of a single instrument whose results can be compared directly across both subjects and time, since it is less susceptible to variation in academic demand. The survey will allow the *What Works?* programme to assess the extent to which interventions designed to foster student success are effective across the participating institutions, as well as allowing more local, contextually-informed, interpretations of the survey data. Neither of the two methodological approaches is intrinsically better than the other: they address different aspects of engagement.

The National Student Survey is currently under review. Some have suggested that 'engagement' items might be included in a revised version of this survey. Arguments can be made for and against this suggestion, and an informed discussion needs to take place regarding its feasibility. Whatever the outcome of any such discussion, student engagement is of key importance. The work reported here will help the higher education sector better to understand engagement: it will, as a consequence, support institutions in their efforts to foster engagement and to enhance the chances of their students' success.

Professor Mantz Yorke
Visiting Professor
The University of Lancaster

Quick facts about the project

This project was designed to explore the applicability and usefulness of student engagement surveys in the UK. 14 survey questions were used by nine institutions in Spring/Summer 2013, and over 8500 student responses were collected and analysed. The questions were adapted from the National Survey of Student Engagement (NSSE), which is widely used in the United States and across the world, and is known to be an effective tool for evaluating the amount and quality of effort that students invest in their studies. The project marked the first time that NSSE items have been used across multiple institutions in the UK, and it gathered valuable information both about the performance of NSSE in the UK, and about the nature of student engagement in different student groups.

Objectives of the project

1. To evaluate the validity and reliability of adapted elements of the NSSE in the UK.
2. To improve our understanding of the levels of student engagement in the UK.
3. To support the participating institutions in using engagement data for enhancement.
4. To raise awareness of the use of engagement surveys for enhancement.

Scales

The 14 questions were grouped into the following four areas:

Critical thinking
Four questions asked students about the emphasis their coursework has placed on a range of mental activities: analysing, synthesising, evaluating and applying information.

Course challenge
Three questions focused on how hard students have worked and how challenged they have been.

Collaborative learning
Three questions asked students how often they have interacted with other students in a range of ways.

Academic integration
Five questions (one of which also appears in the Collaborative Learning group) explored students' interaction with academic staff, participation in class and discussions with others outside class.

Participating institutions

- University of Bath
- Canterbury Christ Church University
- Cardiff Metropolitan University
- King's College London
- Kingston University
- University of Oxford
- University of South Wales
- University of Warwick
- York St John University

Respondents

8582 responses were gathered from the nine participating institutions, an overall response rate of approximately 17%. Around 60% of the respondents were female and 40% male, and 94% were full-time while 6% were part-time. In terms of domicile, approximately 75% were from the UK, 10% from the rest of the EU and 15% from outside the EU. Around 89% were undergraduates and 11% were taught postgraduates.

Case studies

As well as collecting and analysing survey data, this project has also involved the production of a range of case studies about how institutions have made use of engagement data, and how they are beginning to incorporate engagement surveys into their enhancement processes. These case studies and other supporting material are available from: http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

Plans for 2014

An extended version of this project will run again in Spring/Summer 2014, with the 14 questions modified as necessary following consideration of the findings of this report. For further information and to express an interest in taking part, please visit www.heacademy.ac.uk/surveys

Summary of findings

Validity and reliability

Both quantitative and qualitative analysis suggest the 14 items are broadly valid and reliable indicators of engagement, although some small amendments are recommended.

The relationships between the responses to the 14 different questions support the idea that the four intended dimensions of student engagement were being measured: Critical Thinking, Course Challenge, Collaborative Learning and Academic Integration (validity).

The internal consistency of the four question groups was good, showing that the constituent questions were measuring broadly the same thing (reliability). One exception was the only negatively phrased item, about coming to classes unprepared.

The cognitive interviewing found that students were positive and enthusiastic about the idea of engagement survey questions and broadly supported the face validity of the items. It also revealed a number of issues to consider, including the need to expand the response options, and that students found it difficult to distinguish between the questions in the Critical Thinking group; this finding was supported by the analysis of the survey data. Problems were also found with students' interpretation of questions in the Course Challenge group.

Selected findings

Findings from the survey should be interpreted with caution at this pilot stage and may in some cases indicate the need to revise the survey instrument rather than the existence of real issues for learning and teaching. Nonetheless, the results should stimulate discussion about the engagement of different student groups.

Overall
- Students reported interacting with staff less frequently than engaging in other activities. For the Critical Thinking questions, students felt that their courses emphasised applying and analysing information to a greater extent than synthesising and evaluating information.

Gender, mode, level and year of study
- Male students and full-time students felt that they explained things to their fellow students more often than did female students and part-time students.
- Taught postgraduate students reported a lower frequency of being unprepared for class than undergraduate students.
- Full-time undergraduate students reported a greater frequency of discussions with staff about career plans as they progressed through their degree; this finding was supported by the cognitive interviewing.

Discipline
- There were marked differences between disciplines, likely to be due to different pedagogies and expectations.
- The survey results showed that STEM and health and social care students felt that their courses emphasised evaluation and synthesising to a lesser degree than arts and humanities and social sciences students. For example, over 40% of respondents studying mathematical and computer sciences felt that their course emphasised the evaluation of information very little (at least in the way described in the survey question).
- Conversely, arts and humanities students felt that their courses did not emphasise the application of information. 18% of respondents studying European language, literature and related subjects felt that there was very little focus in their courses on the application of information.

International comparisons
- Results from similar surveys used in other countries suggest that the students who participated in the UK pilot discussed their career plans with staff much less often than students in the US.
- It appears that the students in the UK pilot may have talked to academic staff about ideas from the course more often than students in Australasia.
- Of all the countries and regions included in the comparison, the UK pilot had the lowest proportion of students stating that they had never discussed their academic performance with staff.

1. Introduction

This report contains the findings from a pilot of 14 items (or questions) from the National Survey of Student Engagement (NSSE) widely used in North America. It is the UK's first multi-institutional NSSE pilot, where several institutions have agreed common items and wordings, and shared their data to allow substantive findings to be drawn about how student engagement survey questions function in the UK.

Student engagement surveys have been prominent in other parts of the English-speaking world for many years; their low profile in the UK is perhaps due the dominance here of the National Student Survey (NSS), and similar surveys, that measure students' satisfaction or their perceptions of the quality of provision. There have been vocal advocates of NSSE; Graham Gibbs, in his influential report, *Dimensions of Quality*, described NSSE's impressive claim to be a valid proxy for learning gain. He followed this up at a more practical level in *Implications of 'Dimensions of Quality' in a Market Environment*:

> The National Survey of Student Engagement provides a more valid predictor of learning gains than does the NSS, as well as a clearer indication of the nature of provision that students experience, and current piloting of its use in the UK should be supported and extended. (Gibbs 2012, p.45)

The project described in this report was, in part, inspired by arguments that data about engagement would provide a better indicator of educational quality than existing data sources. Several institutions were either already using, or contemplating the use of, an engagement survey. The role of the Higher Education Academy was to facilitate the co-ordination between the institutions, to undertake the analysis and to produce this report.

This publication provides relatively detailed findings about how well the items and scales (or groups of items) performed in the UK, using both quantitative and qualitative research methods, but the primary aim of the project as a whole was to gather data that would be of practical use. To convey the value of engagement data for institutional efforts to enhance learning and teaching, a set of case studies has been created and is available to download separately. Written by the participating institutions, the case studies cover a range of topics including the importance of working with students as partners, and the relationship between engagement data and student performance. Their primary purpose is to give a sense, beyond the relatively technical findings about validity and reliability, about how collecting information about students' levels of engagement can lead to improvements in learning and teaching. Summaries of those case studies are interspersed throughout this report, to put the findings in a real-world context.

This report happens to coincide with the review of the NSS currently being undertaken, led by the Higher Education Funding Council for England.[1] Given that the inclusion of engagement items in the NSS is being considered as part of the review, we hope this report will provide some evidence about how well some of the NSSE items work in the UK context. However, this report does not seek to address the role engagement data might play in public information and the future of the NSS is outside its scope.

There may be a perception that by moving from surveys that ask students how happy they are about what they have received, to surveys that ask students how much effort they have invested, the responsibility is shifted away from institutions, departments and course teams and onto the students themselves. That was certainly not NSSE's rationale, and is quite the opposite of what we intend. Indeed, it is the institution's responsibility to facilitate and improve engagement, by creating environments and opportunities that allow and encourage students to work hard, to invest emotionally and intellectually in their studies and to interact with their teachers, their course and each other in ways that will benefit their learning.

This project has been challenging, but the findings that have emerged have borne out the effort involved in creating this cross-institutional collaboration. In order to allow more institutions to take part, and to gather more data about student engagement in the UK, we will be running the project again in Spring/Summer 2014. This has just been a first step; with nine institutions contributing, the data has only a limited claim to be representative of the UK as a whole, especially given the lack of participating institutions from Scotland and Northern Ireland. More information is available at http://www.heacademy.ac.uk/surveys, and institutions considering participating in 2014 can express their interest there.

---

[1] http://www.hefce.ac.uk/news/newsarchive/2013/news83583.html [accessed 28 October 2013].

1.1 What this report contains

There are five main elements to this report other than a concluding chapter and this introduction. Chapter 2 describes the background to the project, and the items and scales that we have used. Chapter 3 lays out the findings from the analysis of the survey data, focusing on the relationships between the survey items, the underlying structure of students' responses, and the reliability of the four scales. Chapter 4 provides information about the respondents, and Chapter 5 describes some of the results broken down by selected student characteristics (subject of study, gender, etc). Chapter 6 describes the findings from the qualitative investigation of the items, using the technique of cognitive interviewing; this chapter was contributed by Camille Kandiko and Frederico Matos, researchers from King's College London, and is a summary of a full research report, which is available as a separate document.

As well as this report, various additional documents are available to download from http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement. The full report of the cognitive interviewing can be accessed, as well as a set of case studies from the participating institutions describing the benefits of engagement data for enhancement, and some of their experiences of taking part in the process. Also available is a filterable spreadsheet with more detailed results.

1.2 Statistical note

When analysing survey data, statistical significance is a common test of confidence that a given result does not arise from chance: whether the result observed for the survey sample is also likely to be true for the whole population. Significance levels are reported in a few places in this report, and statistical tests that result in a 'p value' below 0.05 are usually considered significant. Where differences are reported, primarily in Chapter 5, effect size has also been used. Statistical significance is sensitive to sample size, and with several thousand responses analysed (as in this report) differences that are too small to be of practical interest can emerge as significant. Effect size is a way of evaluating whether or not a difference is not only significant, but of substantive importance. Unless otherwise stated, the statistic used to measure effect size is Cramer's V, adjusted for the number of categories. Values of 0.1-0.3 indicate a small effect, 0.3-0.5 a medium effect, and over 0.5 a large effect.

It should be noted that both statistical significance and Cramer's V assume that the results are based on a random sample of the population (in this case, all undergraduate and taught postgraduate students in the UK). As the institutional surveys from which our data were gathered were census surveys (administered to all relevant students) the respondents were not chosen at random but formed a self-selected sample. This has the potential to introduce response bias. Moreover, the participating institutions were not randomly selected either and they are not a representative sample of all UK institutions. For that reason, the significance levels and effect sizes should only be taken as indicative.

Finally, in some places in this report (the correlational analyses, factor analyses and scale reliability tests in Chapter 3) students' responses to multiple items have been averaged to arrive at a mean value between 1 and 4 (as the items all have four response options). This treats the four response options as lying on a continuous scale, and assumes that the gaps between the response options are all the same, which may not be true. For example, the gap between 'never' and 'sometimes' may be larger than the gap between 'often' and 'very often'. Wherever possible the response options have been treated as four distinct categories, but for some of the analyses the use of mean values is required.

1.3 Acknowledgements

What Works? programme, a joint HEA/Paul Hamlyn Foundation project also involving the use of engagement data. His expertise and experience have been very helpful, and I'm grateful to him for writing a foreword to the report. Amber Lambert and Alex McCormick from Indiana University (which administers NSSE) have been very accommodating in arranging the license for the use of the items, and in providing general help and advice. I would like to acknowledge the support of colleagues at the HEA for ideas and practical assistance; Abbi Flint, Jason Leman and particularly Paul Bennett. Paul, Jason, Gosia Turner and Camille Kandiko all provided very valuable comments on drafts of this report.

1.4 About the author

Dr Alex Buckley works in the Student Surveys team at the HEA, and leads on their support for institutions around using undergraduate survey data for enhancement. In 2012 he published *Making it count: Reflecting on the NSS in the process of enhancement*. As well as student surveys, he also works on conceptual and political aspects of student engagement. Before joining the HEA in 2010, Alex taught applied and professional ethics at the University of Leeds, after gaining a PhD in philosophy from the same institution in 2008.

2. Background to the project

2.1 The National Survey of Student Engagement (NSSE)

The National Survey of Student Engagement has been in use in North America continuously since 2000. Administered by the Centre of Postsecondary Research (CPR) at the University of Indiana, it is optional for institutions in the US and Canada, and institutions pay a fee to participate;[2] 1,554 institutions have used the survey since 2000, and in 2013 around 370,000 students responded. Although it was first developed in response to concerns about the inability of the influential US rankings to reflect how much students actually learn (Kuh et al 2001), the focus of the survey is predominantly on enhancement. The CPR encourages institutions to use their results publicly "in ways that serve to increase understanding of college quality and that support institutional improvement efforts"[3], and they use innovative methods to encourage prospective students to take account of NSSE scores (NSSE 2013). However, the survey is designed primarily for the enhancement of learning and teaching and the results are not universally made public.

> The first, and most important [purpose of NSSE], is to provide high-quality, actionable data that institutions can use to improve the undergraduate experience (Kuh 2009, p.9)

Programme-level surveys, designed to gather data about how students get on with their course or programme as a whole, typically focus on students' perceptions of the quality of their educational experience, based on, for example, satisfaction models (such as the Student Satisfaction Approach (Harvey 2003) in the UK, and the Noel Levitz inventory in the US[4]) or models linking student perceptions with approaches to study (the Course Experience Questionnaire in Australia (Ramsden 1991), and its derivative in the UK – now more widely known as a satisfaction survey – the National Student Survey (Richardson et al 2007)).[5] NSSE, in contrast, grew out of decades of (primarily American) work supporting the idea that what is most influential on student learning is *student engagement*, understood as "the amount of time and effort students put into their studies and other educationally purposeful activities…[and] how the institution deploys its resources and organizes the curriculum and other learning opportunities to get students to participate in activities".[6] This broad conception in turn draws on a range of theories of student learning, including the importance of time on task, student involvement (Astin 1984), deep vs surface learning (Marton and Saljo 1976) and the seven 'good practices in undergraduate education' (Chickering and Gamson 1987).

The resulting survey covers a wide range of student behaviours and attitudes. Students are asked to, among other things, report the frequency with which they engage in various tasks, rate the extent to which their course and institution emphasise and encourage various elements, and evaluate how much their experience has led to the development of particular skills. The general aim is to "ask undergraduates about their educationally purposeful experiences" (McCormick et al 2013, p,7).

Up to and including 2012, the primary way in which the results of NSSE were reported was in a set of five 'benchmarks'; aggregates of items, partially derived from empirical testing but also with an eye on conceptual concerns. Those five benchmarks (with example items) are:

- active and collaborative learning (eg 'How often have you worked with other students on course projects or assignments?');
- level of academic challenge (eg 'How much has your coursework emphasized applying facts, theories, or methods to practical problems or new situations?');
- student-faculty interaction (eg 'Have you worked with a faculty member on a research project?');
- enriching educational experiences (eg 'How often have you had discussions with people from a race or ethnicity other than your own?');

---

[2] $7500 dollars for institutions with more than 12000 students: http://nsse.iub.edu/html/pricing.cfm [accessed 30 September 2013].
[3] http://nsse.iub.edu/_/?cid=199 [accessed 30 September 2013].
[4] https://www.noellevitz.com/student-retention-solutions/satisfaction-priorities-assessments/student-satisfaction-inventory [accessed 30 September 2013].
[5] There is a separate body of literature devoted to students' evaluation of individual teachers. See eg Marsh (1987).
[6] http://nsse.iub.edu/html/about.cfm [accessed 30 September 2013].

- supportive campus environment (eg 'To what extent does your institution emphasise providing opportunities to be involved socially?').

However, in 2013 (perhaps following concern about the psychometric properties of the five benchmarks (LaNasa et al 2009)), a new set of ten 'indicators' have been used in place of the five benchmarks:[7]

- higher-order learning
- reflective and integrative learning
- learning strategies
- quantitative reasoning
- collaborative learning
- discussions with diverse others
- student-faculty interaction
- effective teaching practices
- quality of interactions
- supportive environment

Other substantial changes have been made to the survey for the 2013 administration, to respond to a changing HE context and new research findings (McCormick et al 2013).

Extensive research has been done both on the NSSE instrument itself, and using the data that it yields, perhaps due to having its home in a research centre rather than a government department or sector agency.[8] Empirical studies have suggested robust links between NSSE scores and student learning, to the extent that the former can be taken as proxies for the latter: a powerful claim for any large-scale survey.

> Our findings suggest that increases on institutional NSSE scores can be considered as reasonable proxies for student growth and learning across a range of important educational outcomes.
> (Pascarella et al 2010, p.21)[9]

NSSE is not without its critics, who – aside from being concerned about the psychometric properties of the benchmarks – have expressed doubts about the links between the benchmarks and student outcomes (Gordon et al 2008) and about the relationship between the kind of perceived gains measured by NSSE and students' actual gains (Bowman 2009). More generally (and more alarmingly) strong reservations have been raised about students' ability to accurately report information about their behaviour and perceptions (Porter 2011).

Many of the criticisms are well-founded, but the strengths of NSSE in relation to its basis in empirical and theoretical research, its enhancement focus, and the evidence for links between NSSE scores and student learning, mean that it is potentially a powerful tool for understanding and improving student learning.

Those benefits of NSSE have led to interest around the globe, especially in English-speaking sectors. Consisting of a moderately modified version of NSSE for the Australasian context, the Australasian Survey of Student Engagement (the predictable acronym being AUSSE) was first administered in 2007 (Coates 2010). Adapted versions of the survey have also been implemented in China (Ross et al 2011, Zhang 2013), South Africa (Strydom and Mentz 2010) and most recently the Republic of Ireland (Drennan et al 2013).

Although the UK has, until the project reported here, lacked any multi-institutional administration of NSSE, work has been carried out at individual institutions. One of the first to do this was Reading University (Creighton et al 2008); later examples include Warwick University (Taylor et al 2011), Worcester University (Scott 2011), Sheffield Hallam University[10] and York St John University (Payne 2012).[11] The Higher Education Academy has supported this work at a national level, through project funding and events such as the annual Surveys for Enhancement conference.

---

[7] http://nsse.iub.edu/html/developmentEngagementIndicators.cfm [accessed 30 September 2013].
[8] An extensive list of research publications can be found here: http://nsse.iub.edu/html/pubs.cfm [accessed 30 September 2013].
[9] See also Kuh et al (2007).
[10] https://students.shu.ac.uk/feedback/docs/SHSES_results_short.pdf [accessed 28 October 2013]
[11] Two of these institutions, Warwick University and York St John University, participated in the current project. York St John University has used a version of AUSSE rather than NSSE.

One of the complicating features of discussions about NSSE in the UK is the unique and disjunctive nature of the concept of student engagement as it occurs here. In addition to the conception of student engagement similar to that behind NSSE (related to the quality and amount of student effort, energy and motivation), in the UK the phrase is also used to refer to students' participation in decision making, through representative and governance structures (Trowler 2010, QAA 2012). It should be noted that NSSE contains very few items that relate to 'student engagement' in that second sense, and therefore does not fully address the broader idea of student engagement as understood in the UK.

2.2 Information about the project

The project had four key objectives:

1. Evaluate the validity and reliability of adapted elements of NSSE in the UK.
2. Improve our understanding of the levels of student engagement in the UK.
3. Support the participating institutions in using engagement data for enhancement.
4. Raise awareness of the use of engagement surveys for enhancement.

This report most directly addresses objectives (1) and (2): objective (1), by providing the results of quantitative testing (such as factor analysis) and the cognitive interviewing; and objective (2), by providing the results of the survey, broken down by selected student characteristics (eg gender). In addition, the contents of this report will address objective (3) by enabling the participating institutions to learn more about the survey and the results it has yielded, and to compare their results against the aggregate results. We also hope that this report, including both data analysis and examples of how institutions view the importance of engagement data to the enhancement process, will generate greater interest in the potential of student engagement surveys and thus address objective (4). The collection of case studies (available as a separately downloadable document at: http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement) presents a range of institutional perspectives on the value and impact of engagement survey data in the context of enhancement.

The University of Warwick

The benefits of introducing a Student Engagement Survey and running it on an iterative basis are varied. In our case study we focus on our early experiences having carried out a survey of student engagement at the University of Warwick in 2011 and 2013. We illustrate the advantages of gathering and analysing serial data, and indicate how we have been able to quantify any changes that have occurred across the University. The comparisons thus far have allowed us to identify the departments or faculties which indicate higher levels of student engagement. In doing so we can progress to identifying areas of good practice in order to encourage their adoption in departments where students appear less engaged.

Read the full case study at: http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

The following institutions have participated in the pilot project:
- University of Bath;
- Canterbury Christ Church University;
- Cardiff Metropolitan University;
- King's College London;
- Kingston University;
- University of Oxford;
- University of South Wales;
- University of Warwick;
- York St John University.

Oxford Brookes University and Manchester Metropolitan University were also involved in the project. Oxford Brookes University will be administering the items and scales in Spring/Summer 2014 and therefore has not

contributed data to this report. Manchester Metropolitan used alternative engagement items, based on but different from the 14 items described here; their experiences are described in their case study. Additionally, students from a further five institutions participated in the cognitive testing of survey questions (Birkbeck - University of London, Glasgow Caledonian University, University College London, University of East Anglia and University of Manchester).

The steering group for the project included representatives from each of the participating institutions, as well as other individuals with relevant expertise. Full list of steering group members:

- Paul Bennett (Higher Education Academy);
- Alex Buckley (Higher Education Academy);
- Abbi Flint (Higher Education Academy);
- Caroline Gibson (University of Warwick);
- Camille Kandiko (King's College London);
- Mark Langan (Manchester Metropolitan University);
- John Lea (Canterbury Christ Church University);
- Steve May (Kingston University);
- Shaun McGall (University of Bath);
- Denize McIntyre (University of South Wales);
- Berry O'Donovan (Oxford Brookes University);
- Anthony Payne (York St John University);
- Gill Perkins (Canterbury Christ Church University);
- Nicola Poole (Cardiff Metropolitan University);
- Paul Taylor (University of Warwick);
- Gosia Turner (University of Oxford);
- Mantz Yorke (University of Lancaster).

Administration of the items and scales was devolved to institutions; rather than create a single survey instrument that all participants would be obliged to administer, it was decided to develop a smaller set of scales that could be inserted into existing internal surveys, in order to avoid over-surveying students and to minimise the resource required for the project. Seven institutions took this route; one institution administered the full NSSE survey (including the modified items selected by the group) and one institution administered the full AUSSE survey (see below).

Accordingly, both the survey methods and when the surveys were administered differed between the participants. The earliest administration began in February 2013, and the latest ended in June 2013. A mix of paper-based and online questionnaires was used; no institution solely used paper-based questionnaires.

King's College London

As part of an institution-wide curriculum enhancement initiative, King's College London developed the King's Experience Survey. Two competing approaches to the data emerged: enhancement and quality assurance. Student engagement survey data have presented opportunities for enhancement of teaching and learning, and for the student voice to be shared in the context of academic development. We found dissemination at both individual and collective levels provided a balance between collecting meaningful data that is a) used for institutional improvement and pedagogical enhancement, such as tips sheets for new lecturers and data on how students spend their time, and b) for external comparisons and marketing.

Read the full case study at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

The average institutional response rate was 22%, though the response rates varied greatly between institutions (see Chapter 4 for more detail).

The devolved and differing nature of the survey administrations means that caution should be exercised in interpreting the results. Neither the wider questionnaires in which the scales were embedded, the exact point in the year when the survey was administered, nor the publicity and method of administration were the same. These compromises were made in order to facilitate the project at this initial stage, and greater standardisation may be possible in future.

2.3 Selection of items and scales

The project group made two important decisions early in the process. Firstly, it was decided to select a relatively short set of items and scales, to allow participating institutions to include them in their regular internal surveys, rather than necessitating a new survey administration. Secondly, it was decided to employ existing and well-

researched items, both to shorten the development time and to enable international comparisons. A range of existing engagement questionnaires were reviewed, including the following:

- 2012 iteration of NSSE;[12]
- 2013 iteration of NSSE;[13]
- AUSSE;[14]
- The Student Experience in the Research University (SERU) survey;[15]
- The University Experience Survey;[16]

Existing UK adaptations of NSSE (from the universities of Reading, Worcester, Warwick, Sheffield Hallam and King's College London) were also reviewed.

A range of areas were proposed as possible topics to be covered by the items, including the areas finally covered by the four scales. In addition, there was consideration of areas around students as researchers, and student involvement in decision-making (about curriculum design, departmental and university governance etc). The decision to select a small number of pre-existing items meant that those further topics were eliminated for the 2013 iteration of the project; it was agreed however that they were particularly important aspects of student engagement in the UK context. It was proposed that discussions should take place with a view to developing items on those topics for inclusion in the 2014 iteration.

A review of the literature found a range of relevant scales, in addition to the five NSSE benchmarks mentioned above (Kuh 2001, Kuh et al 2001, Nelson Laird et al 2005, Pike 2006a, LaNasa et al 2009, Campbell and Cabrera 2011, Esquivel 2011, Scott 2011).[17] Of those scales, four were selected.

All of the items are derived from items in NSSE, and were used with the permission of the owners of the copyright (the Trustees of Indiana University). What follows is a list of the final scales and items used, and information about their background.

Critical Thinking

During the current academic year, how much has your coursework emphasized the following mental activities? (Response categories: Very much / Quite a bit / Some / Very little)

1. Analysing in depth an idea, experience or line of reasoning

2. Forming a new idea or understanding from various pieces of information

3. Evaluating a point of view, decision, or information source

4. Applying facts, theories, or methods to practical problems or new situations

---

Evidence for the use of these four items as a scale has been found by various researchers, including Nelson Laird et al (2005), LaNasa et al (2009), Wells (2011) and Esquivel (2011). This is also one of the ten 'indicators' used for the 2013 NSSE. The 2013 wording has been used, with the following changes:

- in the item root, the phrase 'academic year' has been substituted for the phrase 'school year' (this change has been applied across all items);
- in item 1, the wording was simplified and clarified from the original NSSE wording, 'Analysing an idea, experience, or line or reasoning in depth by examining its parts'.

---

[12] http://nsse.iub.edu/pdf/survey_instruments/2012/NSSE2012_US_English_Paper.pdf [accessed 24 October 2013]
[13] http://nsse.iub.edu/pdf/survey_instruments/2013/NSSE%20Instrument.pdf [accessed 24 October 2013]
[14] http://www.acer.edu.au/ausse [accessed 24 October 2013]
[15] http://cshe.berkeley.edu/research/seru/index.htm [accessed 24 October 2013]
[16] http://www.ues.edu.au/ [accessed 24 October 2013]
[17] AUSSE is also reported using benchmarks, similar to those of NSSE. The ten engagement indicators introduced for NSSE in 2013, mentioned above, were published after our items and scales were finalised, so were not considered.

<div style="border:1px solid #4a90b8; padding:10px;">

Course Challenge

5. In your experience at your institution during the current academic year, about how often have you worked harder than you thought you could to meet a tutor's/lecturer's standards or expectations?
(Response categories: Very often / Often / Sometimes / Never)

6. During the current academic year, to what extent have your courses challenged you to do your best work?
(Response categories: Very much / Quite a bit / Some / Very little)

7. In your experience at your institution during the current academic year, about how often have you come to taught sessions unprepared? (eg not completed assignments, readings, reports, etc.)
(Response categories: Very often / Often / Sometimes / Never)

</div>

These three items constitute one of Pike's 'scalelets' (Pike 2006a) minus two items. An item asking students to estimate the time spent preparing for class was removed, to maintain consistency of response options. An item asking how much the institution emphasises academic work was removed, as it was felt to be too general and could devalue vocational courses.

- item 5 is drawn from NSSE 2012 (it does not appear in NSSE 2013); the term 'instructor's' has been replaced with 'tutor's/lecturer's';
- items 6 and 7 are drawn from NSSE 2013;
- item 7 has been changed from the NSSE version, which is '…come to class without completing readings or assignments'; the change was made in order to provide a wider range of possible activities.

<div style="border:1px solid #4a90b8; padding:10px;">

Collaborative Learning

In your experience at your institution during the current academic year, about how often have you done each of the following?
(Response categories: Very often / Often / Sometimes / Never)

8. Worked with other students on course projects or assignments

9. Explained course material to one or more students

10. Discussed ideas from your course with others outside of taught sessions (students, family members, co-workers, etc), including by email/online

</div>

This scale is also a Pike 'scalelet' (Pike 2006a). Items 8 and 9 are drawn from NSSE 2013. Item 10 is drawn from NSSE 2012 (it does not appear in NSSE 2013), with the following changes (item 10 also appears in the Academic Integration scale below):

- the phrase 'readings or classes' has been replaced with 'course';
- the phrase 'class' has been replaced with 'taught sessions';
- the phrase 'including by email/online' has been added.

In your experience at your institution during the current academic year, about how often have you done each of the following?
(Response categories: Very often / Often / Sometimes / Never)

11. Asked questions or contributed to course discussions in other ways

12. Discussed your academic performance and/or feedback with teaching staff

13. Discussed ideas from your course with teaching staff outside taught sessions, including by email/online

10. Discussed ideas from your course with others outside of taught sessions (students, family members, co-workers, etc), including by email/online

14. Talked about your career plans with teaching staff or advisors

These items were found to compose a scale by LaNasa et al (2009), with the inclusion of an item asking how often prompt feedback has been received; that item was not used here due to its similarity to an item on the National Student Survey:
- item 11 is drawn from NSSE 2013;
- item 12 is drawn from NSSE 2013, with the replacement of 'a faculty member' with 'teaching staff', and the addition of 'and/or feedback';
- item 13 is drawn from NSSE 2012, with 'readings or classes' replaced by 'course', 'faculty members' replaced by 'teaching staff', and with the addition of 'including by email/online';
- (see above for item 10);[18]
- item 14 is drawn from NSSE 2013, with 'a faculty member' replaced by 'teaching staff or advisors'.

Some of the participating institutions have used items that differ slightly from the wordings described above. For example, York St John University has used the AUSSE wordings, so that item 12 in the Academic Integration scale is worded '…discussed grades or assignments with teaching staff'. Small differences in wordings have been considered equivalent. Where wording differs in a marked way, the data have been excluded.

Manchester Metropolitan University

The Faculty of Science and Engineering at Manchester Metropolitan University ran a pilot study of five 'engagement-style' questionnaire items (ISSe) appended to existing NSS-style questions from the institutional internal student survey (ISS). The ISSe comprised five questionnaire items based on core areas covered by the HEA Engagement Surveys project; contributing to classes, interacting with tutors and peers, reflecting upon feedback and being inspired to work hard. These were required to retain the NSS's five point Likert scale (from Strongly Agree to Strongly Disagree) and were limited to course-level surveys only. A total of 1377 students were surveyed, a response rate of 35%. Preliminary analyses revealed a consistent reduction in ratings of engagement items from first to final levels of undergraduate study and the best predictor of 'Overall Satisfaction' was 'The course inspired me to do my best work'. This was not as good a predictor as existing ISS questionnaire items about ratings of course organisation or teaching.

Read the full case study at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

---

[18] Although item 10 appears in two scales, it only appeared once in the questionnaire.

3. Initial analyses

3.1 Item responses

The following figures show the distribution of item responses. Tables containing the data are available as a separately downloadable appendix at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement
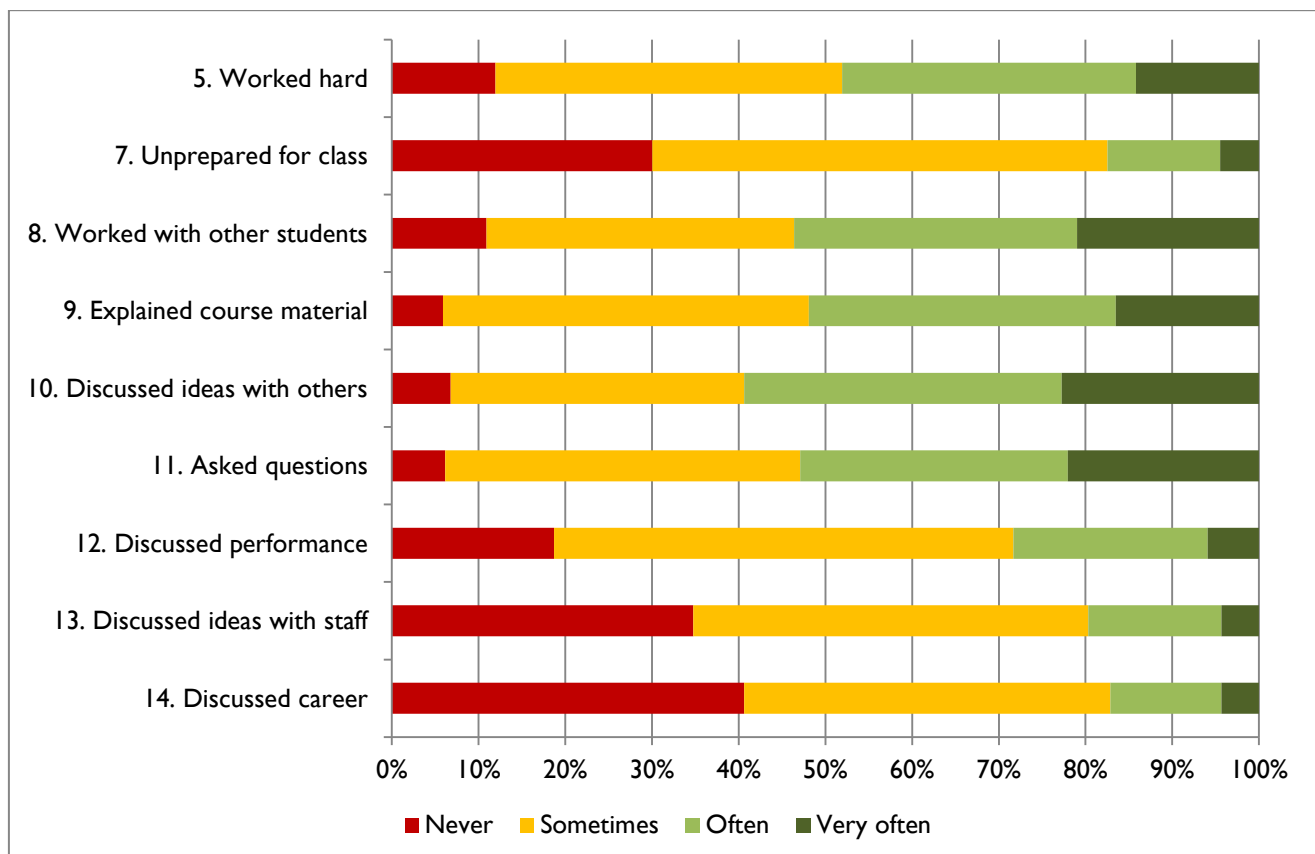
Throughout the rest of the report, items are referred to by abbreviated titles and item numbers.

Figure 1: Responses to items 1-4 and 6.
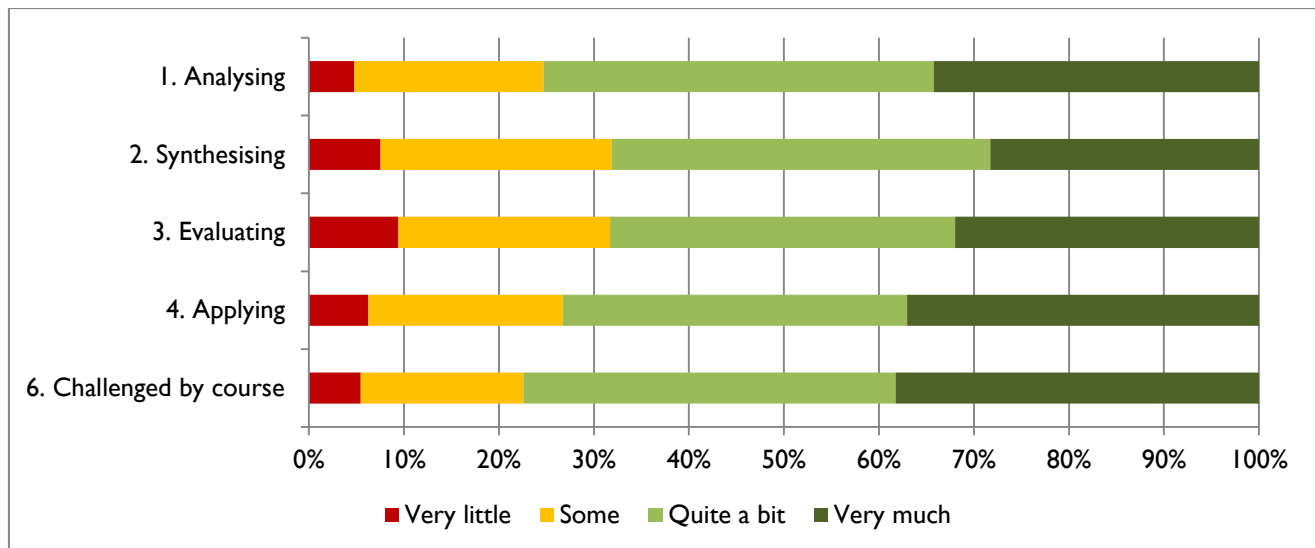
N = 7545 - 8346

Figure 2: Reponses to items 5 and 7-14.

N = 7540 - 8373

Overall, respondents felt that the activities they engaged in least were talking about career plans with staff and discussing ideas from the course with teaching staff outside class. They also felt that they came to taught sessions unprepared relatively infrequently. They felt they engaged most often in asking questions in class, explaining course material to other students and discussing ideas from the course with individuals other than teaching staff. Nearly two in five respondents felt that they were very much challenged to do their best work, while nearly one in ten felt that there was very little emphasis on the evaluation of information on their course.

Detailed breakdowns of the overall results for different student characteristics are available in Chapter 5 and in a separately downloadable appendix

3.2 Relationships between the items

The table below shows correlations between the different items. Correlations range from very small (0.021 between the items on being unprepared for class and evaluating information) to reasonably large (0.583 between the items on analysing and synthesising information), virtually all of them are statistically significant at the 0.01 level. Some of the scale groupings can be seen in the correlation table, for example the high correlations between items 1, 2 and 3 (all from the Critical Thinking scale). Other correlations to be expected include the correlation between items 5 and 6 (both about level of effort) and between 12, 13 and 14 (all about interaction with academic staff).

The correlations with item 7 (on being unprepared for class) are largely negative, reflecting the negative wording of that item ('how often have you come to taught sessions **un**prepared'). More importantly, that item correlates only weakly with the other 13 items, suggesting that it may not belong in this group. The fact it is the only negatively-phrased item may be a partial explanation.

Table 1 Item correlations													
	2.	3	4	5	6	7	8	9	10	11	12	13	14
1. Analysing	.583**	.561**	.392**	.217**	.322**	-.032**	.104**	.119**	.188**	.195**	.161**	.196**	.148**
2. Synthesising		.512**	.401**	.242**	.316**	-.051**	.127**	.126**	.195**	.194**	.195**	.223**	.164**
3. Evaluating			.302**	.210**	.228**	-.021	.112**	.089**	.166**	.236**	.168**	.196**	.146**
4. Applying				.196**	.290**	-.065**	.212**	.197**	.155**	.093**	.137**	.165**	.175**
5. Worked hard					.355**	-.124**	.180**	.221**	.240**	.183**	.261**	.262**	.209**
6. Challenged by course						-.174**	.164**	.121**	.178**	.094**	.196**	.183**	.159**
7. Unprepared for class							.022	.030*	-.044**	-.023*	-.056**	-.037**	-.023*
8. Worked with other students								.403**	.222**	.101**	.143**	.183**	.162**
9. Explained course material									.318**	.229**	.215**	.240**	.222**
10. Discussed ideas with others										.233**	.223**	.311**	.198**
11. Asked questions											.260**	.263**	.213**
12. Discussed performance												.427**	.451**
13. Discussed ideas with staff													.421**

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

	Large correlation (0.50 – 1.0)
	Medium correlation (0.30 – 0.49)
	Small correlation (0.10 – 0.29)

3.3 Identification of dimensions of engagement

Factor analysis is a statistical technique that groups items together based on patterns in the correlations between those items (shown in Table 1). By looking for groups among the correlations, factor analysis can be used to investigate the underlying elements of students' engagement with their courses, as measured by the 14 items. This provides evidence about the 'construct validity' of the 14 items: whether or not they measure the construct being investigated. In the selection of the 14 items (as described in Section 2.3) they were assumed to form four scales focused on four particular dimensions of student engagement: the course emphasis on critical thinking; the extent to which the course was challenging; the amount of collaborative learning; and students' integration into the academic context. Factor analysis allows us to test that assumption. The medium to large correlations in Table 1 provides an initial reason to think that a smaller set of factors should emerge.[19]

The 14 items were found to group into four factors with eigenvalues larger than 1. This suggests that each of the four factors explain an acceptable amount of variance in responses. In other words, the extent to which the score for each item is high or low can be partly explained by four higher level factors that may reflect broader concepts than that measured by any individual item.[20] Those four factors together explained 57% of the total variance, meaning that 57% of the total variation in the results for the 14 items is explained by the four underlying factors.

Table 2 shows the results of the factor analysis; the values indicate the extent to which the scores for the different items 'load' on to (or define) the four factors. The values in the Pattern Matrix indicate the influence the results for each item have on each of the four factors. The values in the Structure Matrix represent correlations between those item results and the factors, and are good indications of the interaction between item scores and factors. Higher values have been highlighted; values of 0.3 or higher on the Pattern Matrix are normally taken to indicate an item may contribute to a factor.

The results in Table 2 suggest that the first factor is defined by the three items on interacting with staff (items 12-14) and the item on asking questions in class (item 11). These are four of the five items assumed to constitute the Academic Integration scale (see Section 2.3). The fifth item, about discussing ideas with others outside class, does not quite reach the 0.3 cut-off value. This will be discussed further in the next section.

The second factor is defined by the four items describing the various mental activities emphasised by the coursework (items 1-4). These are the items assumed to form the Critical Thinking scale. While the item on being challenged by the course (item 6) nearly reaches the 0.3 cut-off value, this is likely to be due to the fact that it shares response options with the Critical Thinking items.

The values for the third factor indicate that it is defined by the items asking about working with other students, explaining material to other students, and discussing ideas from the course with people other than teaching staff (items 8-10). These are the three items assumed to constitute the Collaborative Learning scale.

The values for the fourth factor suggest that it is defined by the items asking about being unprepared for class, being challenged by the course and working beyond expectations (items 5-7). These are the three items assumed to form the Course Challenge scale.

---

[19] Principal Components Analysis (PCA) was used, with Oblimin rotation (which doesn't assume that the different elements being measured are independent). The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.834, well above the recommended value of 0.6. Bartlett's Test of Sphericity was statistically significant, another indicator of the suitability of factor analysis.

[20] The extraction of four factors was supported by looking at the shape of the scree plot. In addition the process of parallel analysis was also used; using http://ires.ku.edu/~smishra/parallelengine.htm a set of random eigenvalues were produced; only the four highest ranked eigenvalues from our analysis were higher than their equivalently-ranked random eigenvalues (Patil et al 2007).

In summary, the factor analysis provides strong evidence for the construct validity of the 14 items; that is, the four hypothesised scales are largely borne out by the patterns of correlations between the items, allowing us to conclude that the 14 items do measure the four distinctive dimensions of student engagement:

- Critical Thinking
- Course Challenge
- Collaborative Learning
- Academic Integration

It should be noted however that in the case of the Academic Integration scale, the results suggest caution in taking all five of the hypothesised items to define a common dimension of student engagement. That dimension appears to be limited to interaction with staff and in classroom settings, rather than interaction with other students, family members, co-workers, etc. This issue will be returned to in the next section.

Table 2 Results of factor analysis

	Pattern Matrix				Structure Matrix			
	Component				Component			
	1	2	3	4	1	2	3	4
12. Discussed performance	.776	.039	-.045	-.111	.766	-.185	.198	-.200
13. Discussed ideas with staff	.714	-.015	.065	-.046	.744	-.238	.294	-.152
14. Discussed career	.740	.047	.006	-.038	.734	-.166	.226	-.128
11. Asked questions	.514	-.163	.035	.166	.548	-.283	.211	.063
1. Analysing	.012	-.853	-.059	.001	.229	-.841	.164	-.153
2. Synthesising	.045	-.803	-.038	-.040	.260	-.813	.188	-.192
3. Evaluating	.092	-.805	-.107	.093	.269	-.786	.115	-.056
4. Applying	-.130	-.542	.258	-.151	.119	-.601	.379	-.272
8. Worked with other students	-.108	.019	.830	.004	.140	-.163	.791	-.095
9. Explained course material	.114	.072	.785	.084	.324	-.146	.789	-.029
10. Discussed ideas with others	.268	-.040	.432	-.053	.418	-.234	.531	-.156
7. Unprepared for class	-.004	-.121	.179	.801	-.022	-.015	.096	.753
6. Challenged by course	-.012	-.294	.117	-.593	.183	-.433	.272	-.663
5. Worked hard	.214	-.095	.207	-.476	.366	-.297	.364	-.551

�	0.60 - 1.0
	0.50 - 0.59
	0.40 - 0.49
	0.30 - 0.39
	0.20 - 0.29

3.4 Analysis of scales

While the factor analysis supports the claim that the 14 items measure the four dimensions of student engagement previously identified (the construct validity) the extent to which they do so reliably also needs to be investigated. The primary method for investigating the reliability of a scale is to evaluate internal consistency using a test known as Cronbach's alpha. An acceptable alpha level for a scale is commonly taken to be 0.7 or higher, but alpha scores are known to be affected by the number of items in a scale; low numbers of items can artificially deflate the alpha values. For scales containing few items (less than around six or seven) it is therefore recommended that the mean

correlation between items also be used as a guide, with optimal values lying between 0.2 and 0.4. Both statistics are referred to below. It should be noted that the results for the negatively-phrased item 7 (being unprepared for class) have been reverse-coded to be consistent with the other items.

3.4.1 Critical Thinking scale

The Cronbach's alpha for the four items in this scale is 0.77, safely above the satisfactory level of 0.7. The mean inter-item correlation is 0.458.

Previous work on this scale has found alpha levels of 0.82 (Nelson Laird et al 2005), 0.83 (Wells 2011) and 0.85 (LaNasa et al 2009).

3.4.2 Course Challenge scale

The Cronbach's alpha for the three items in this scale is 0.456, substantially below the acceptable level of 0.7. The mean inter-item correlation is 0.214, which is within the optimal range but at the lower end. Moreover, the fact that item 7 correlates only weakly with the other two items also suggests caution (see Table 3).

Table 3 Inter-item correlations for Course Challenge			
	6. Challenged by course	7. Unprepared for class (reversed)	5. Worked hard
6. Challenged by course	1.000	.173	.361
7. Unprepared for class (reversed)	.173	1.000	.109
5. Worked hard	.361	.109	1.000

Table 4 shows the relationships between the three items and the scale which they are taken to define. The low item-total correlation for item 7, and the fact that the Cronbach's alpha would increase if it were removed, provides further evidence for the removal or rewording of the item. The fact that it is negatively phrased is relevant; inclusion of negatively-phrased items in a scale is known to reduce internal consistency.

The low item-total correlation for that item has been found in other research (Esquivale 2011). More generally, that item has been found to correlate only minimally with other NSSE items.[21]

All three items will be taken to constitute the Course Challenge scale throughout this report, but these concerns about the reliability of the scale should be borne in mind and will be discussed further in Chapter 7.

Table 4 Item-total statistics for Course Challenge		
	Corrected item-total correlation	Cronbach's alpha if item deleted
6. Challenged by course	.365	.195
7. Unprepared for class (reversed)	.170	.530
5. Worked hard	.315	.293

3.4.3 Collaborative Learning scale

The Cronbach's alpha value for this scale is 0.575; this is below the recommended level of 0.7, but given the shortness of the scale (only three items) it is again appropriate to look at the mean inter-item correlation. In this case that value is 0.323, well within the optimal range of 0.2-0.4. Table 5 shows that all of the three items correlate reasonably with one another.

---

[21] Kuh et al (2001), for example, found no correlations larger than 0.18 between the item about being unprepared for class and any other NSSE item.

Table 5 Inter-item correlations for Collaborative Learning			
	8. Worked with other students	9. Explained course material	10. Discussed ideas with others
8. Worked with other students	1	.404	.222
9. Explained course material	.404	1	.316
10. Discussed ideas with others	.222	.316	1

Pike (2006b) found a slightly higher Cronbach's alpha of 0.72 for an earlier version of this scalelet, but since that time there have been substantial changes to items.

3.4.4 Academic Integration scale

The Cronbach's alpha for the four items (11 to 14) suggested by the Pattern Matrix to constitute this scale is 0.667, slightly below the recommended level of 0.7. The mean inter-item correlation is 0.339, providing evidence for the reliability of the scale.

In the initial selection of the items (see Section 2.3) it was assumed that item 10 (regarding discussing ideas with individuals other than teaching staff) also contributed to this scale, as well as contributing to the Collaborative Learning scale, yet it fell short of the cut-off value in the factor analysis. Adding that item to the scale yields a slightly higher Cronbach's alpha of 0.677, but alpha values are known to be higher for scales with more items, so this is only partly suggestive. The addition of the extra item reduces the mean inter-item correlation to 0.30, a lower value but still within the optimal range of 0.2-0.4.

Tables 6 and 7 show the item-total statistics and inter-item correlations for the scale with all five items. Item 10 appears to fit as well as other items; its lowest correlation (0.198 with item 14) is the lowest in the scale, but its correlation with the total scale score is virtually identical to that of item 11. More tellingly, its removal (or that of item 11) does not increase the Cronbach's alpha, suggesting that it does add to the reliability of scale.

LaNasa et al (2009) found a Cronbach's alpha of 0.758 for a scale consisting of all five items, plus an additional item not used here.

The University of Oxford

In Spring 2013 Oxford piloted the NSSE survey (including the HEA harmonised items) in two colleges surveying over 700 undergraduate students and achieving over 280 responses (38% response rate). Survey responses have been linked with the complete student record giving a unique insight into how engagement can affect the final exam performance. This case study is focused on student gains, expressed as their final exam performance and their self-reported level of academic skills. The relationship between student gains and NSSE items is investigated in order to assess the potential of student engagement surveys for pedagogical research and how it can be used in practice to enhance student experience.

Read the full case study at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

Table 6 Item-total statistics for Academic Integration (with fifth item)		
	Corrected item-total correlation	Cronbach's alpha if item deleted
11. Asked questions	.341	.667
12. Discussed performance	.500	.597
14. Discussed career	.463	.612
13. Discussed ideas with staff	.530	.582
10. Discussed ideas with others	.340	.668

Table 7 Inter-item correlations for Academic Integration					
	11. Asked questions	12. Discussed performance	14. Discussed career	13. Discussed ideas with staff	10. Discussed ideas with others
11. Asked questions	1.000	.260	.213	.264	.231
12. Discussed performance	.260	1.000	.450	.428	.223
14. Discussed career	.213	.450	1.000	.422	.198
13. Discussed ideas with staff	.264	.428	.422	1.000	.311
10. Discussed ideas with others	.231	.223	.198	.311	1.000

3.4.5 Scale results

Table 8 shows a range of statistics for the four scales, including the Cronbach's alpha, the mean inter-item correlations (both discussed above) and the mean values of students' responses to all the items in the scale. The standard deviation is a measure of the spread of results; the results for Academic Integration are reasonably closely grouped (broadly indicating a higher level of agreement among students) while the spread of results for Critical Thinking is wider. The 'skewness' is an indication of how asymmetrical the range of responses are; the positive value for Academic Integration indicates that the results are distributed more on the left of the peak (towards lower responses), while negative values for Critical Thinking and Course Challenge show that the results are more on the right (towards higher responses). The 'kurtosis' is an indication of how flat or peaked the distribution of results is; the positive value of Academic Integration shows that the results for that scale are clustered in a central peak, whereas the negative value for Collaborative Learning indicates a flatter distribution. Both the skewness and kurtosis can be seen visually in Figure 3.

The 'loading' value indicates the extent to which the scale loads on to a single overarching student engagement factor, which explains 51% of the total variance over the 14 items. The highest value is for Academic Integration, suggesting that is the scale with the most influence on that overarching student engagement construct.[22]

---

[22] To obtain these values, the process described in Section 3.3 was run again, with the scale values substituted for the item values.

| Table 8 Scale statistics							
	Mean	Std. Deviation	Skewness	Kurtosis	Cronbach's alpha	Mean inter-item correlation	Factor loadings
Critical Thinking	2.9719	.69630	-.390	-.339	.770	.458	.662
Course Challenge	2.8779	.58524	-.397	.034	.456	.214	.641
Collaborative Learning	2.6728	.64765	.056	-.432	.575	.323	.745
Academic Integration[23]	2.2597	.55347	.530	.276	.677	.300	.801

Table 9 shows the correlations between the four scales. The absence of high correlations provides evidence that the scales are measuring different (but nevertheless related) elements of student engagement. The higher correlation between the Collaborative Learning and Academic Integration scales is likely to be partially explained by the fact that they contain a common item (item 10).

Table 9 Scale correlations				
	Critical Thinking	Course Challenge	Collaborative Learning	Academic Integration
Critical Thinking	1.000	.347	.260	.345
Course Challenge	.347	1.000	.259	.309
Collaborative Learning	.260	.259	1.000	.551
Academic Integration	.345	.309	.551	1.000

---

[23] The five-item version of the scale has been used throughout Section 3.4.5

Figure 3: Scale histograms.[24]
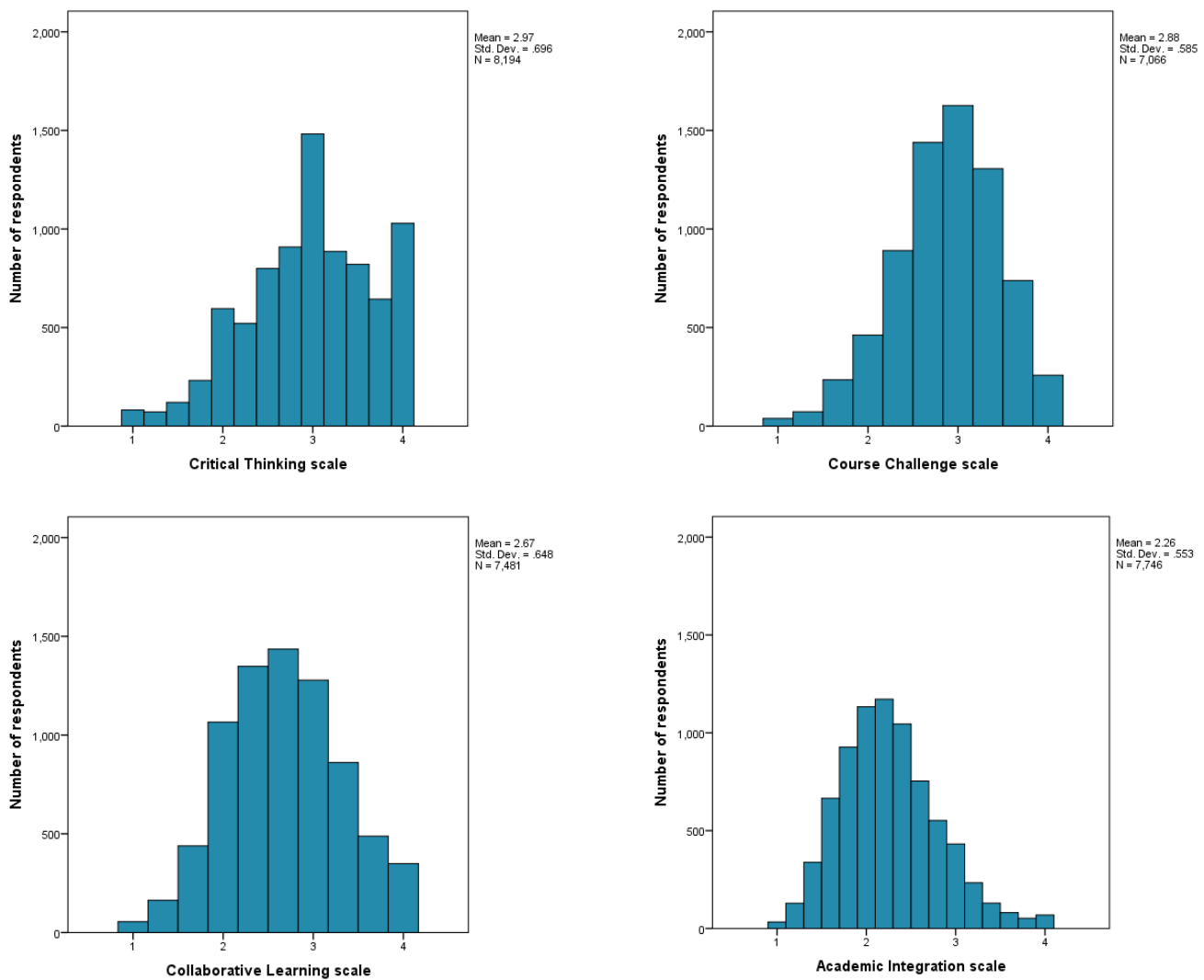
The peaks at the values of 3 and 4 in the histogram for the Critical Thinking scale should be noted. It may suggest that respondents have selected those options without due consideration, or that respondents have found it difficult to differentiate between the items and thus selected the same response for all four items. This will be discussed further in Section 4.3 and in Chapter 6.

3.5 Conclusions

The analyses presented in this chapter allow us to draw conclusions about two important properties of the 14 items. The factor analysis in Section 3.3 sheds light on the 'construct validity' of the 14 items: whether or not they actually measure the four areas of student engagement described in Section 2.3. The evaluation of internal consistency in Sections 3.2 and 3.4 (primarily using Cronbach's alpha) provides evidence about the reliability of the four scales: in particular, evidence about whether their constituent items are measuring the same underlying attribute.

Broadly speaking, the results support the reliability and validity of the 14 items and the four constituent scales. However, the following more detailed findings should be considered:
*   The correlations between the 14 items are reasonable. An exception is item 7 (being unprepared for class) which appears to be largely unrelated to the others;
*   the factor analysis supports taking the 14 items to constitute four scales, measuring the constructs intended (and described in Section 2.3). However, the Academic Integration construct appears to relate only to

---

[24] The horizontal axes shows 1 for 'Never' or 'Very little' to 4 for 'Very often' or 'Very much'.

interaction with academic staff and interaction in class, and not interaction with others such as students and family members;

- the reliability of the Course Challenge scale may be improved by the removal or rewording of item 7 (being unprepared for class);

- the histogram for mean responses to the Critical Thinking scale shows that an unexpected number of students selected 'Very much' for all four items in the scale, raising the possibility that they found it difficult to differentiate between the items.

In summary, the findings in this chapter suggest that two of the four scales (Critical Thinking and Collaborative Learning) appear to function as expected, while the remaining two scales (Course Challenge and Academic Integration) may require modification.

York St John University

York St John University conducts an annual Student Engagement Survey which targets first and second-year undergraduate and postgraduate taught students including those from overseas.

The survey instrument is based on the Australian University Survey of Student Engagement (AUSSE) and is administered electronically during April. Responses are data-matched to student records enabling analysis by demographic, equalities and academic progress variables. The survey has received strong support from staff and the Students' Union.

With minimal promotion, response rates of over 20% have been achieved with over 90% of respondents completing all questions. The results have been benchmarked internationally and are now being used by faculties and central services to inform quality enhancement and to further stimulate student engagement.

Read the full case study at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

4. Profile of respondents

This chapter provides information about the backgrounds and disciplines of the students who responded. It demonstrates that the data represents the responses of a broad range of students, covering all the key student groups.

4.1 Response rates

Of the nine participating institutions, the highest institutional response rate was 40%, while the lowest was 13%. The average institutional response rate was 22%. Overall, 8582 responses were collected from an approximate total sample of 50000, giving an approximate aggregate response rate of 17%.

4.2 Student characteristics

The devolved nature of the survey administration meant that data collection procedures differed between institutions. One of the consequences is that information about student characteristics is not uniformly available. The tables below list the number of students in each category, along with the number of students for whom information is not available ('No information'). The percentage of students in each category refers only to those students to whom information is available (those without information are excluded). The number of institutions who have contributed students to each category is also listed (this information is relevant to Chapter 5).

Table 10 Gender			
	Count	Percentage	Contributing institutions
Female	4693	59.6%	8
Male	3183	40.4%	8
No information	706		3

The set of responses contains considerably more women than men, a common occurrence for survey data. This must be borne in mind when interpreting results, although the differences between genders appear to only be important for item 9 (see Section 5.2).

Table 11 Domicile			
	Count	Percentage	Contributing institutions
UK	5258	74.7%	7
EU	693	9.8%	7
Non-EU	1085	15.4%	8
No information	1546		3

Table 12 Mode			
	Count	Percentage	Contributing institutions
Full-time	6736	93.7%	8
Part-time	450	6.3%	8
No information	1396		3

Part-time students are substantially under-represented in the set of responses.

Table 13 Level			
	Count	Percentage	Contributing institutions
Undergraduate	7600	88.6%	9
Postgraduate Taught	982	11.4%	4

No postgraduate research students are included in the set of responses.

Table 14 Year of study (undergraduate)			
	Count	Percentage	Contributing institutions
1	2926	41.6%	8
2	2040	29.0%	8
3	1208	17.2%	7
4	778	11.1%	6
5	69	1.0%	5
6	12	0.2%	4
No information	567		1

First year students are over-represented in the sample while final year students are under-represented.

Table 15 Year of study (postgraduate taught)			
	Count	Percentage	Contributing institutions
1	575	74.8%	3
2	110	14.3%	3
3	82	10.7%	2
4	2	0.3%	1
No information	213		1

Table 16 Discipline clusters			
	Count	Percentage	Contributing institution
Arts and humanities	1488	18.1%	8
Health and social care	1309	16.0%	8
Social sciences	2407	29.4%	8
STEM	2756	33.6%	9
Combined	240	2.9%	5
No information	382		1

Table 17 Discipline (using JACS code)			
	Count	Percentage	Contributing institution
(A) Medicine and dentistry	624	7.8%	4
(B) Subjects allied to medicine	685	8.5%	7
(C) Biological sciences	819	10.2%	8
(F) Physical sciences	420	5.2%	5
(G) Mathematical and computer sciences	653	8.1%	5
(H) Engineering	543	6.8%	5
(J) Technologies	3	0.0%	1
(K) Architecture building and planning	133	1.7%	2
(L) Social studies	719	9.0%	7
(M) Law	193	2.4%	4
(N) Business and administrative studies	934	11.7%	8
(P) Mass communication and documentation	73	0.9%	5
(Q) Linguistics, classics and related subjects	394	4.9%	7
(R) European language, literature and related subjects	409	5.1%	5
(T) Eastern, Asiatic, African, American and Australasian languages, literature and related subjects	30	0.4%	3
(V) Historical and philosophical studies	343	4.3%	6
(W) Creative arts and design	239	3.0%	7
(X) Education	561	7.0%	6
(Com) Combined	240	3.0%	5
No information	567		1

4.3 Non-response and 'satisficing'

Figure 4 shows, for each of the 14 items, the percentage of respondents who either selected a 'not applicable' option or did not select any option for that item (this is a proportion only of students who responded to at least one item – students who did not respond to any item have not been included in the set of results). This figure gives an indication of whether any of the items are particularly challenging or uninteresting to students. The small scale on the vertical axis should be noted (ranging from 0% to just 4%), nevertheless the relatively high number of non-responses for items 1-5 may suggest that they are more challenging for students to respond to. Item 11 appears to be the most easily comprehensible item.

Figure 4: Levels of non-response.
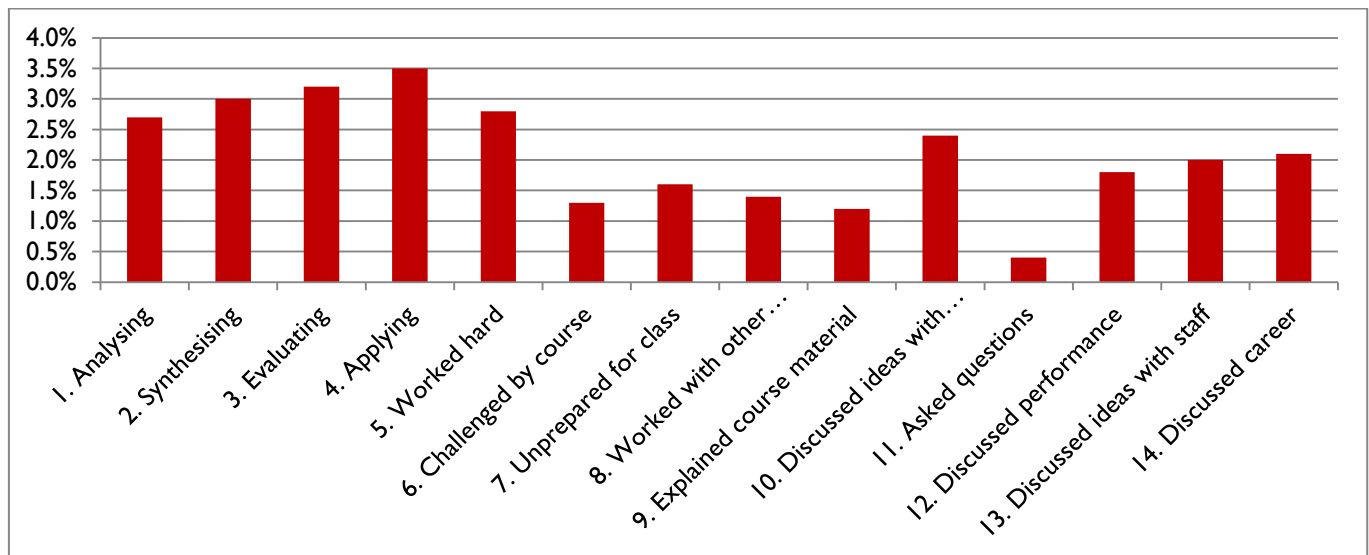
N = 7645 - 8582

It is useful to check how many students are 'satisficing' – responding without due consideration – by seeing how many are selecting the same response option for every item. This is only a rough guide; as items in scales address the same broad topic, identical responses to all the items (excepting the negatively-phrased item) may well be appropriate. Figure 5 shows the percentages of students who selected the same response option for all items. It shows the proportion of students who selected the same option for all 14 items, and then the proportion who selected the same option for each of the four scales. It only includes those students who answered all the relevant items. The proportions show that very few students selected the same response option for all 14 items; this is to be expected given the inclusion of a negatively-phrased item (item 7) and indicates a low level of acquiescence. The figures for the Critical Thinking scale are less encouraging, with a total of 28% of students selecting the same option for all four items. This may suggest that students have trouble differentiating the four items.

Kingston University

The Academic Mentoring Programme (AMP) intervention at Kingston University, focusing on specific undergraduate modules, facilitates second-year students supporting the academic success of first-year students and thereby the engagement of both in university life.

All survey items were converted to both paper-based and online questionnaires and distributed to all students from the modules in question; the results are now being analysed and are expected to reveal whether the AMP has led to significantly greater academic engagement by students.

We intend to utilise the survey in 2013-14 to help determine the effectiveness of a range of interventions.

Read the full case study at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

Figure 5: Same responses to all items.[25]
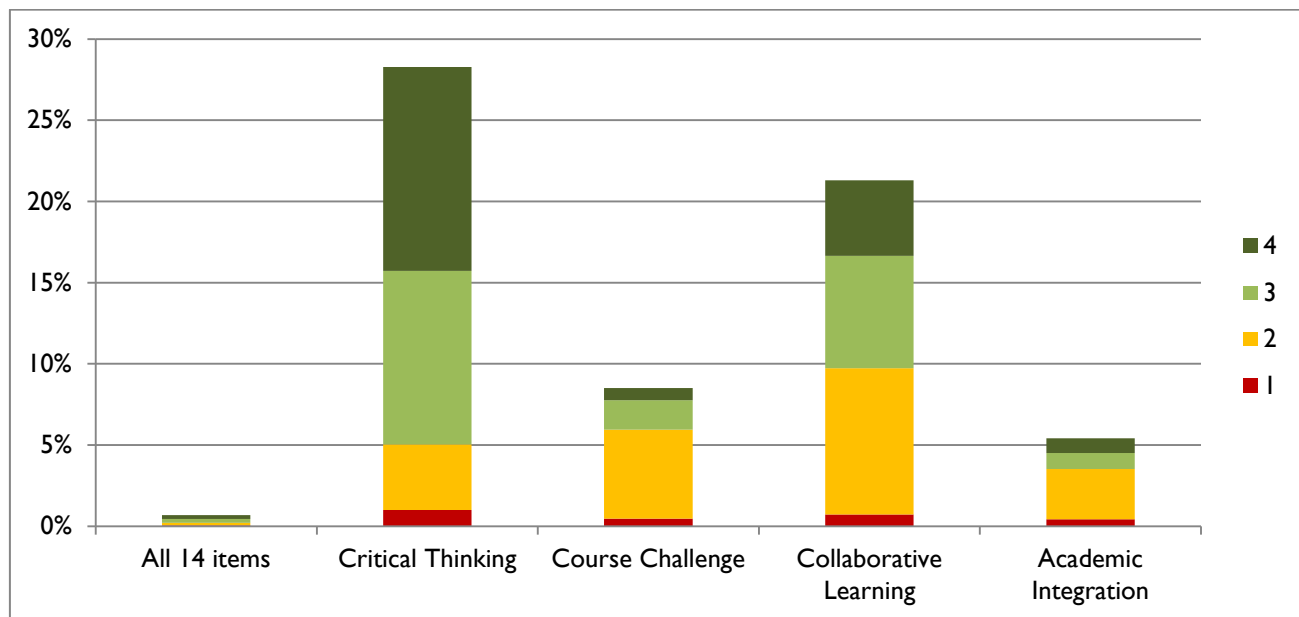
N = 6297- 8194

---

[25] 1 refers to 'Very little' or 'Never', 4 refers to 'Very much' or 'Very often'.

5. Breakdowns by student characteristics

This section contains charts showing the different responses given by different types of students. Comprehensive tables are available in the separately downloadable appendix, charts have only been included here where the results are of particular interest.

To identify substantively important differences, effect size has been used. Effect size relates to the strength of the relationship between categorical variables, such as gender and response option ('very often', 'often' etc) and has been measured using Cramer's V (adjusted for the number of categories); as a rule of thumb, values over 0.1 are considered small effects, values over 0.3 are considered medium effects, and values over 0.5 are considered large effects. A common finding with student surveys is that there is much less variation in experience between broad groups of students (gender, institution, discipline etc) than there is between individual students, even when they are part of the same group. It is therefore unsurprising that few large effect sizes were found; values over 0.1 have therefore been taken to indicate differences of substantive importance unless otherwise stated, but the general lack of medium or large effect sizes should be borne in mind. In all cases where substantively important differences are cited, the differences are statistically significant at the 0.001 level. (For more information see Section 1.3.)

To avoid any risk that individual institutions' results are identifiable, and to ensure that results are reasonably representative of the wider student population, differences are only displayed where at least four institutions contribute to each category, and no institution contributes more than 50% of the respondents. The sole exception to this rule is the comparison between undergraduate and taught postgraduate students, where one institution contributed 56.8% of the taught postgraduate responses.

<div style="border:1px solid #000">

Oxford Brookes University

Oxford Brookes is committed to enhancing its students' experience based on evidence. Consequently, we are keen to get student feedback, not just on their satisfaction but on the impact on student behaviours of an integrated programme of evidence-based, cross-institutional enhancement initiatives. These include: the introduction of five graduate attributes as core outcomes of a Brookes' education; a new framework for academic advice and support; an 'assessment compact' that details effective assessment and feedback practices and responsibilities (staff and student). Subsequent to stringent testing the Brookes Student Engagement Survey will be offered to all non-final-year undergraduates in March 2014.

Read the full case study at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

</div>

Institution-level results are not shown in this report; however, the differences between institutions were analysed, and they were substantively important for all items. At the level of a single discipline (business and administrative studies was used, as that discipline had the highest number of responses and a high number of contributing institutions), differences between institutions were substantively important for all items, and the effect size was medium for six items (including all four of the Critical Thinking items).

5.1 Scale and overall means

Treating the response options as continuous allows us to evaluate the differences between groups of students for the scale means and the overall means (see Section 1.3). Few of the differences are of substantive importance.[26] The difference between institutions on the Course Challenge scale was found to be substantive (for reasons of confidentiality those differences will not be shown). A reasonable effect size was also found between a range of

---

[26] For the comparisons between mean scale scores, ANOVA was used, with eta squared as a test for effect size. The eta squared value for the differences between selected disciplines on the Collaborative Learning scale was 0.05, classified as a small to medium effect.

disciplines for the Collaborative Learning scale; the mean results and other information for that scale for those selected disciplines are shown in Table 18.[27]

| Table 18 Discipline results for Collaborative Learning			
	N	Mean	Std. deviation
Subjects allied to medicine	527	2.67	0.638
Mathematical and computer sciences	641	2.68	0.669
Business and administrative studies	807	2.85	0.626
Linguistics, classics and related subjects	303	2.53	0.579
European language, literature and related subjects	362	2.58	0.588
Historical and philosophical studies	293	2.37	0.571
Creative arts and design	139	2.78	0.607

Respondents studying historical and philosophical studies felt they engaged in collaborative learning less often than students in other disciplines. Students in business and administrative studies generally felt they did this the most often, but there is a reasonable amount of variation from the mean, shown by the standard deviation.

5.2 Gender

There is one item, item 9 (on explaining course material to other students) for which the differences between male and female have a reasonable effect size (0.119). This is shown in Figure 6.

Figure 6: Gender differences for item 9, Explained course material.
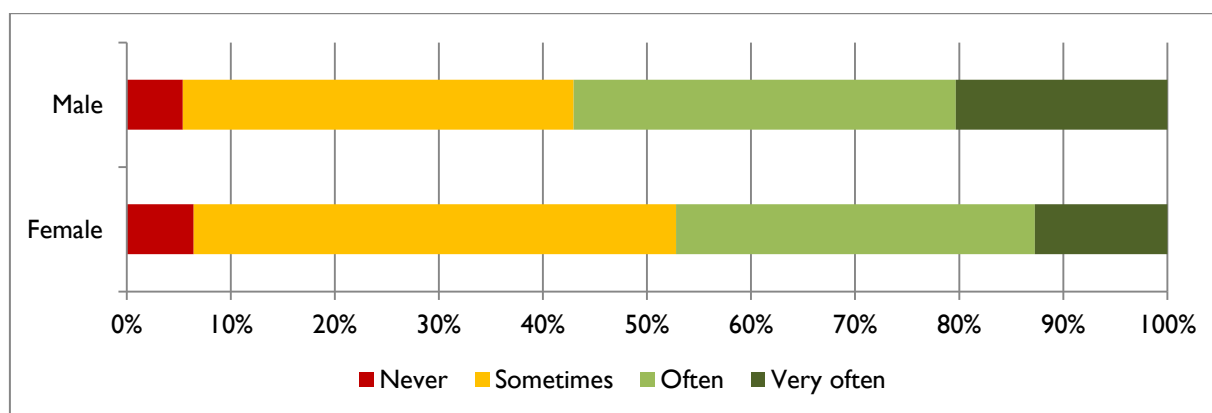

N (male) = 2928, N (female) = 3973

While similar proportions of male and female students felt that they never explained course material to other students or did so often, there is a marked difference between the proportions who selected 'sometimes' or 'very often'. Compared with 13% of female students, 20% of male students felt that they did this very often. Conversely, 46% of female students felt that they did this only sometimes compared with 38% of male students.

5.3 Mode

There are four items where the differences between full-time and part-time students have effect sizes above 0.1 (items 7, 8, 9 and 14) including two of the three items in the Collaborative Learning scale. The largest (with a value of 0.228) is again for item 9, shown in Figure 7.

---

[27] These disciplines were selected because four or more institutions contributed data, and no institution contributed more than 50% of the results.

**Figure 7: Mode differences for item 9, Explained course material.**
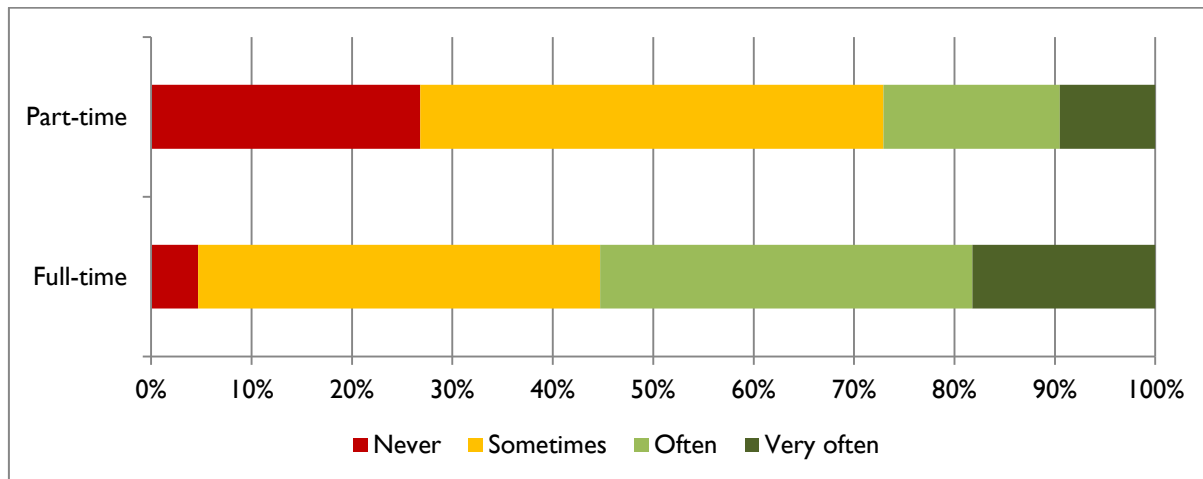
N (full-time) = 5821, N (part-time) = 347

A marked difference can be seen between the proportion of full- and part-time students who felt that they had never, in the current academic year, explained course material to other students (5% of full-time students compared with 27% of part-time students). A number of factors beyond the scope of this report may help to explain this, such as the higher proportion of part-time students who may be distance learners. It should be noted that given students are asked to rate the frequency of activities over their last academic year, one would expect – other things being equal – lower scores for all items from part-time students. This is borne out to some extent by the fact that for all of the four items showing substantive differences, including item 9, the part-time students report less frequent activities than full-time students.

5.4 Level

Comparing the responses of undergraduate and taught postgraduate students, there are two items with effect sizes larger than 0.1 (items 7 and 9). The results for item 7 (with an effect size of 0.104) are shown in Figure 8.

**Figure 8: Level differences for item 7, Unprepared for class.**
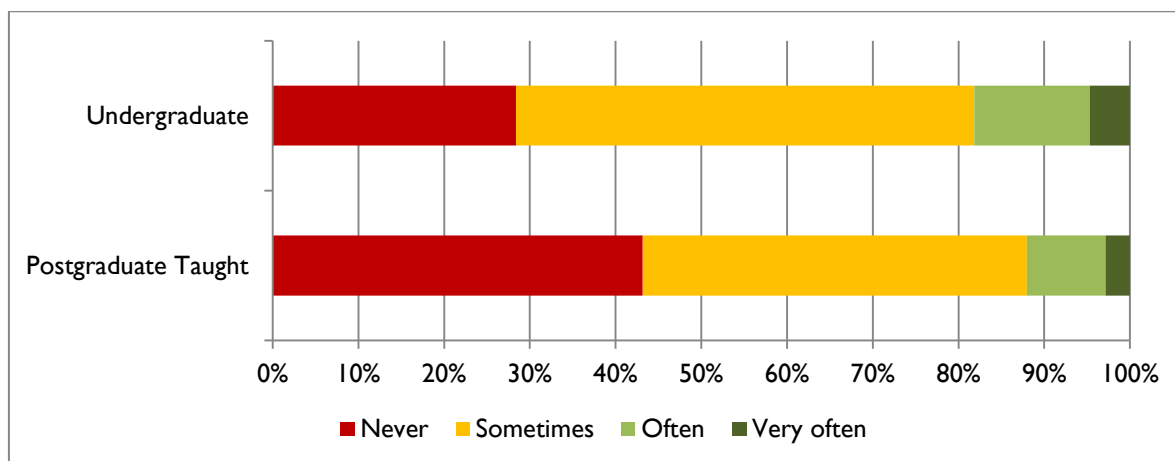
N (undergraduate) = 7178, N (postgraduate taught) = 913

The much higher proportion of postgraduate taught students who felt that they had never arrived unprepared for class (43% compared with 28%) may to some extent be expected, given the extra commitment inherent in the choice to enter postgraduate study. The fact that 5% of undergraduate students state that they are unprepared for class very often is of some concern, but should be understood in the light of the cognitive interviews which found that some students answered 'often' or 'very often' even if they were only partially unprepared (see Chapter 6). For item 9 (explaining course material to other students) taught postgraduate students reported lower frequency of activity than undergraduates.

5.5 Year of full-time undergraduate study

There are three items for which the difference between year of full-time undergraduate study is of substantive importance (items 7, 13 and 14). The largest effect is for item 14 (0.168) which is shown below. Only years one to three were investigated, as less than four institutions contribute results to years four, five and six.

Figure 9: Differences between year of study for item 14, Discussed career with teaching staff.
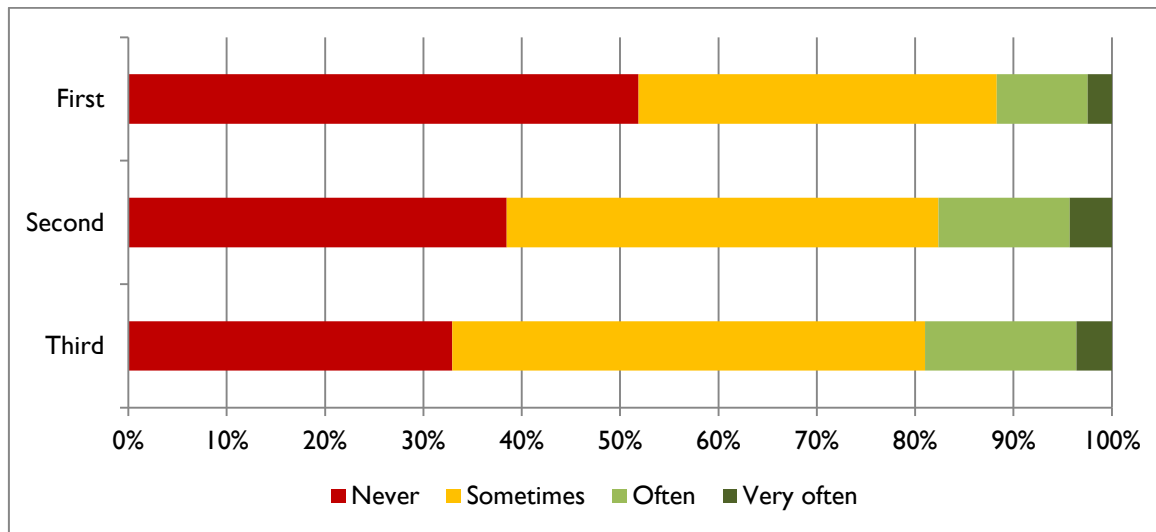
N (first) = 2250, N (second) = 1378, N (third) = 1032

The fact that students at progressive stages of their degree felt that they discussed their career with staff increasingly often is probably unsurprising, given the increasing emphasis on later life as graduation approaches. However, it is also possible that students are, despite the instructions in the survey items, reflecting on more than the last academic year, which would also help to explain the difference. The cognitive interviewing reported in Chapter 6 found further evidence of this. For all three of the items with substantive differences, the frequency of reported activities increases with year of study.

5.6 Discipline

Subject of study is often found to have an effect on students' responses to surveys, and this survey follows that pattern; different pedagogies used in different fields of study are likely to have important impacts on students' participation in different kinds of activity. Substantively important differences (effect sizes greater than 0.1) were found for all 14 items, when selected disciplines were compared.[28] For items 4, 8 and 11 the effect sizes were medium while for item 3, shown below, the effect size was large (0.52).

<div style="border:1px solid #5bb;padding:8px">

Canterbury Christ Church University

In 2011 Canterbury Christ Church University set up a Change Programme to support the delivery of its strategic plan 2011-15 targets. To facilitate a student-University partnership in this work, the Student Ambassadors for Learning and Teaching (SALT) scheme was launched and a new Partners in Learning space was provided.

The SALTs work with staff in developing projects around institutional themes. Data provided by a student engagement survey would help SALTs and staff identify themes that might benefit from project work and address University priorities. Such data would also be integral to evaluating the success of the Partners in Learning space.

Read the full case study at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

</div>

---

[28] These disciplines were selected because four or more institutions contributed data, and no institution contributed more than 50% of the results.

Figure 10: Differences between selected disciplines for item 3, Evaluating.
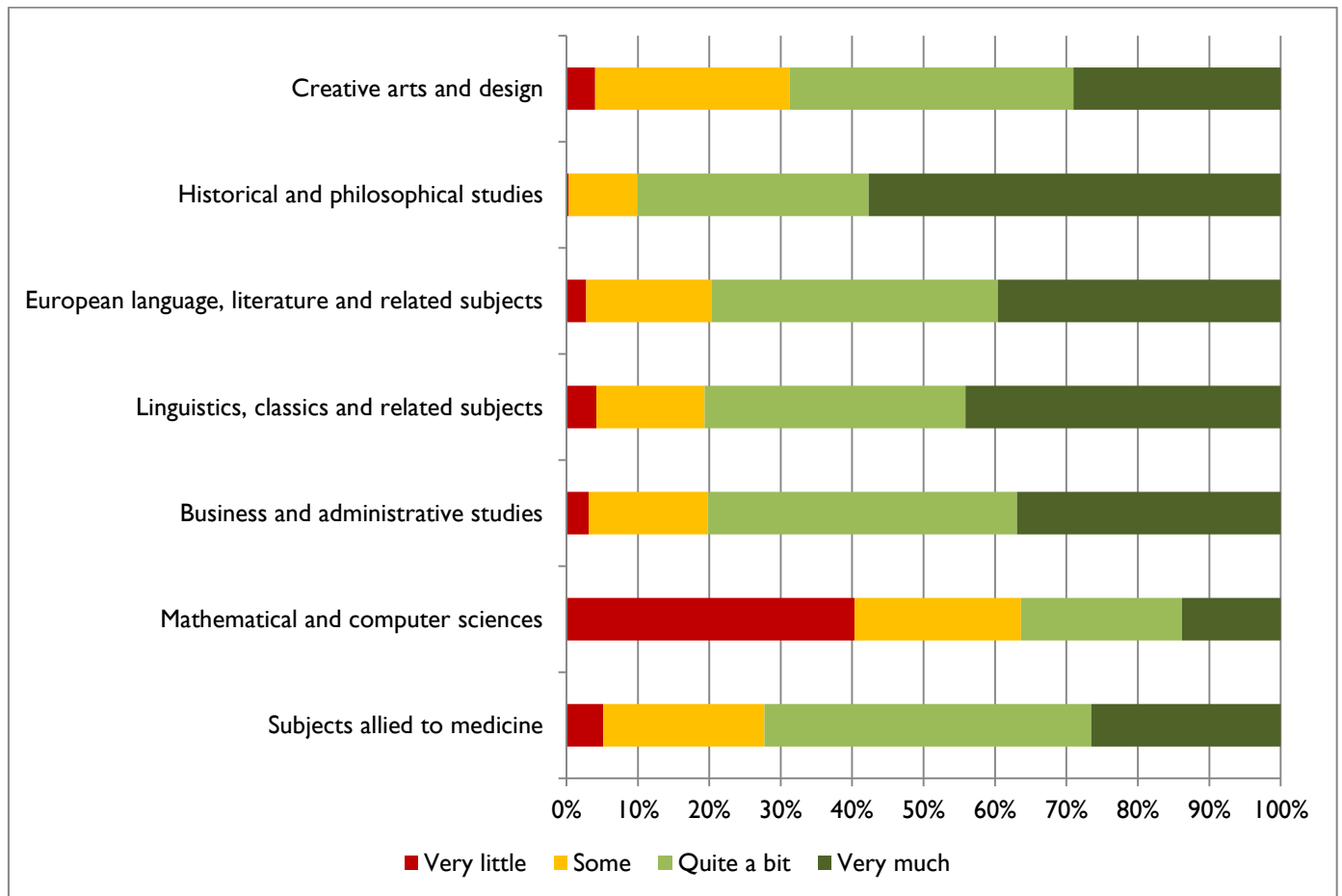
N (creative arts and design) = 224, N (historical and philosophical studies) = 331, N (European language, literature and related subjects) = 399, N (linguistics, classics and related subjects) = 383, N (business and administrative studies) = 905, N (mathematical and computer sciences) = 630, N (subjects allied to medicine) = 664

The large effect size for this item is likely to be caused by the distinctive response of respondents studying mathematical and computer sciences. Of those students 40% felt that there was very little emphasis on the evaluation of points of view, decisions or information sources in their course, compared with (for example) 5.1% for subjects allied to medicine and 0.3% for historical and philosophical studies. This is a surprising finding, and (along with the next section) supports the finding from the cognitive review that students in STEM subjects tended to think that evaluation, at least in the way conceptualised in the survey item, was not an appropriate element of their course (see Chapter 6).

Figure 11: Differences between selected disciplines for Critical Thinking items, items 1-4.
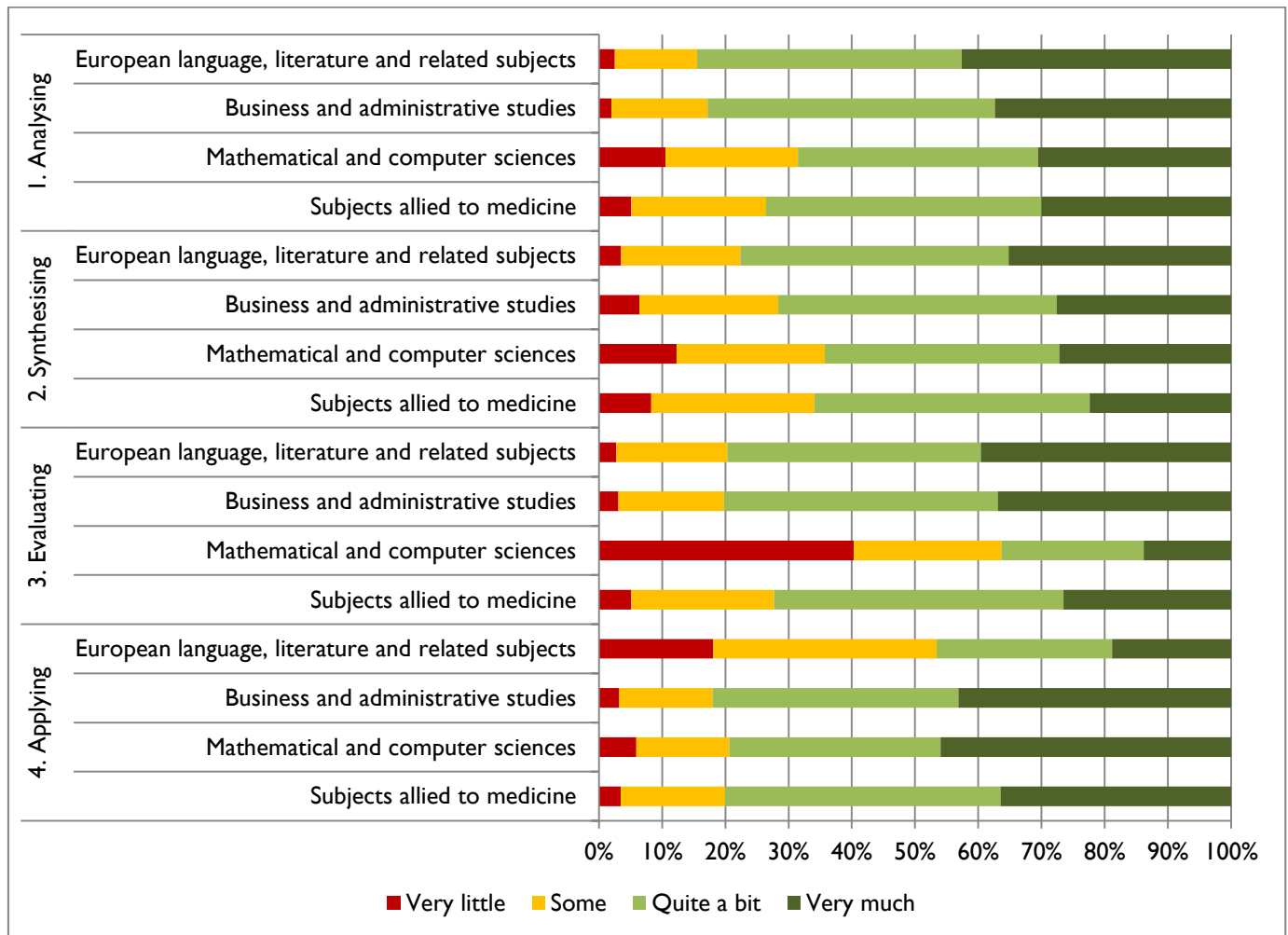
N (European language, literature and related subjects) = 393 – 399, N (business and administrative studies) = 902 – 907, N (mathematical and computer sciences) = 630 – 645, N (subjects allied to medicine) = 664 – 666

Figure 11 shows the results for all four of the items in the Critical Thinking scale, for four selected disciplines, one from each of the four discipline clusters. The results for these four individual disciplines have been shown, not because they are necessarily representative of the four clusters, but because they show the amount of variation that can exist between individual disciplines. The low emphasis on the evaluation of information felt by mathematical and computer sciences students is mirrored to some extent by the low emphasis on the application of information felt by students of European language and literature.[29]

In contrast, the differences for item 8 (about working with other students) displayed in Figure 12, show similarities between the results for mathematical and computer sciences, and European language and literature.[30] Those students report working with other students less frequently than respondents studying subjects allied to medicine and, more markedly, those studying business and administrative studies.

---

[29] The differences between these four disciplines for all four of these items are substantively important: the differences for items 1 and 2 have a small effect, the difference for item 4 has a medium effect, and the difference for item 3 has, as above, a large effect.

[30] The effect size is medium for these differences.

Figure 12: Differences between selected disciplines for item 8, Worked with other students.
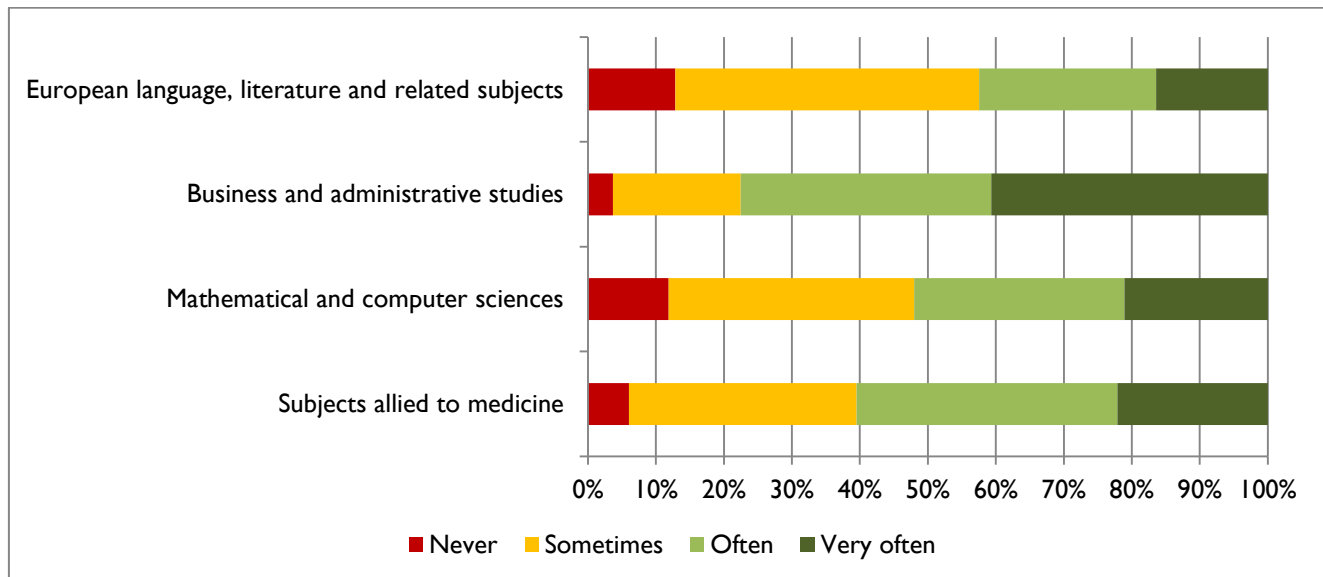
N (European language, literature and related subjects) = 365, N (business and administrative studies) = 814, N (mathematical and computer sciences) = 650, N (subjects allied to medicine) = 529

When the respondents are grouped into five broader discipline clusters, ten of the items demonstrate substantive differences between the discipline clusters, including all four of the items in the Critical Thinking scale (items 1-4, 6-9 and 11-12). Item 3 demonstrated a medium effect size (0.32).

Figure 13: Differences between subject clusters for item 3, Evaluating.
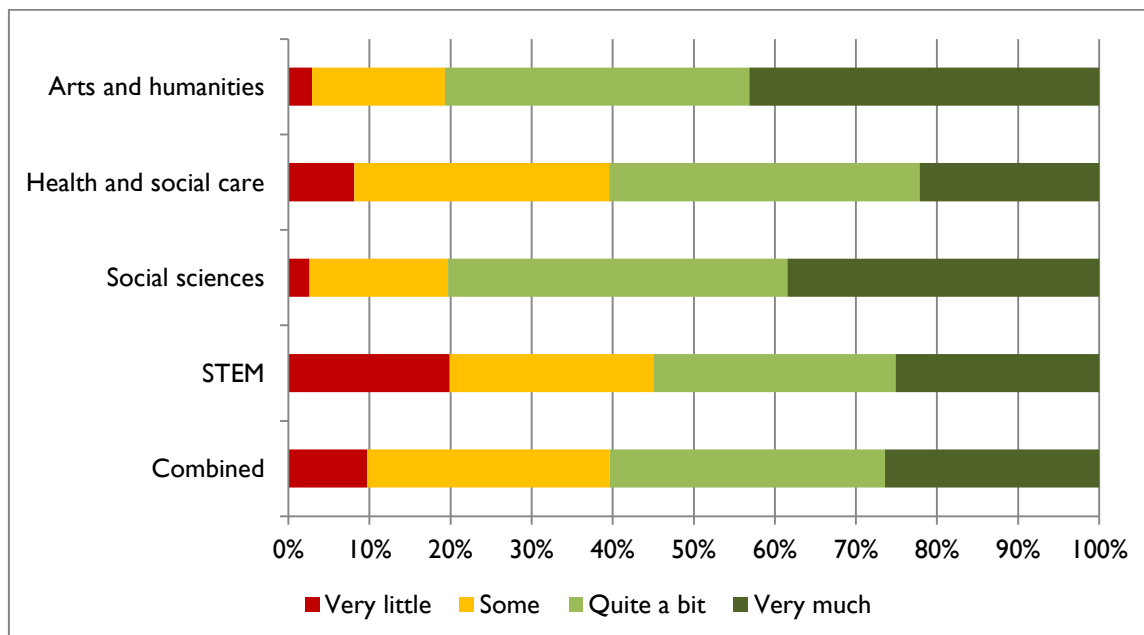
N (arts and humanities) = 1434, N (health and social care) = 1277, N (social sciences) = 2340, N (STEM) = 2680, N (combined) = 227

This chart supports the view that students in STEM subjects felt that evaluation was less appropriate as an activity in their courses. However, it should be borne in mind that mathematical and computer sciences students make up 24% of the STEM students, which significantly contributes to the results for STEM.

5.7 International comparisons

One of the benefits of using items from NSSE without substantial revision is the opportunity to compare results with those from the US. In addition, the close similarities between NSSE, AUSSE and SASSE (the South African Survey of Student Engagement) allow us to compare with results from other countries. In Figure 14, the NSSE

results are all from the 2013 administration, apart from item 10 (missing from the 2013 NSSE) for which the results are from 2012. The AUSSE results are all from the 2012 administration, and the SASSE results are from the 2010 administration. International comparisons for all 14 items are available in the separately downloadable appendix. It should be noted that the results for NSSE, AUSSE and SASSE have a much better claim to be representative of their respective countries or regions, and the results from this project are only indicative of a general UK response.

The largest differences were found for items 5-6, 10-11 and 13-14; the differences for the Academic Integration scale, containing four of those items, are shown below.[31]

Figure 14: Comparisons between UK, US, Australasia and South Africa for Academic Integration scale.
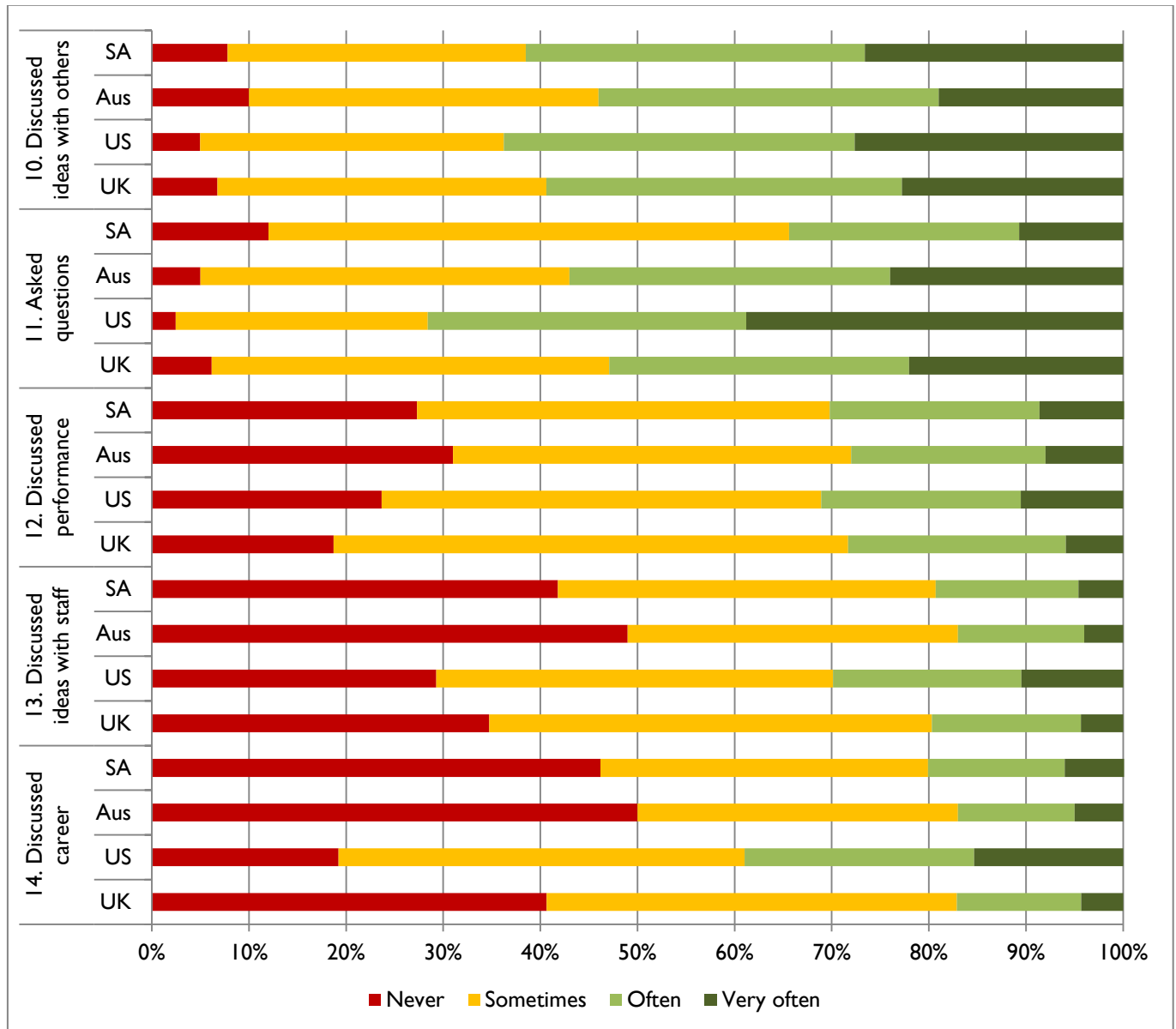
N (UK) = 7540 – 8373, N (US) = 268681 – 334104, N (Aus) = 39790, N (SA) = 9265 – 9420

---

[31] Significance levels and effect sizes have not been calculated for these differences.

For the items in this scale, there is a common pattern of US students reporting greater frequency of these activities. The only exception is item 12 (discussing performance with academic staff) where the UK pilot and the US are broadly similar and, in fact, UK pilot respondents are the least likely of any group to say they have never discussed their performance with staff. The most pronounced differences are for item 14; whereas 15.0% of US students state they have discussed their career plans with staff very often, only 6.1% of South African student, 5.0% of Australasian students and only 4.3% of students from the UK pilot say the same. Around 40% of the UK pilot respondents felt that they had never (during the current academic year) discussed their career plans with academic staff, though this differs by year group (see Figure 10 above) and first years are over-represented in the UK pilot sample. Another marked difference is for item 11 (asking questions in class); nearly 40% of US students report doing this very often, compared to just over 20% of the UK pilot students. The item for which there is the most consistency between the results is item 10 (discussing ideas from the course with individuals other than academic staff).

The University of Bath

The University of Bath, and its Students' Union (SU) has performed exceptionally well over recent years in a range of 'satisfaction' surveys including the NSS, PTES and the Student Barometer. Fundamental to our joint philosophy is the importance of gaining detailed insight into student engagement across all aspects of the student experience. This insight, in turn, informs the University's and SU's partnership to ensure the on-going development and enhancement of the student learning experience. This pilot has contributed significantly to the joint understanding of both the students' experience and their learning styles and behaviours. These pilot 'Student Engagement' scales have been integrated into the University's Programme Evaluation tool, and have influenced our joint decision to re-focus this tool from 'Satisfaction' to wider and more meaningful 'Student Engagement' measures for 2014/15 to supports the University's 'Excellence in Education' Strategy.

Read the full case study at:
http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

6. Summary of the qualitative testing (Camille B. Kandiko and Frederico Matos)

*This chapter, describing the qualitative investigation of the 14 items using the process of cognitive interviewing, was written by **Camille B. Kandiko** and **Frederico Matos** (both of King's College London). The full research report is available at:* http://www.heacademy.ac.uk/resources/detail/nss/engagement_for_enhancement

6.1 Key findings

- Students found the survey questions rigorous and thought they reflected important aspects of their academic student experience. Students overall had a positive response to the survey, particularly for the items in the scales for Academic Integration and Collaborative Learning which focused on what students have done in the current academic year.

- The items in the Course Challenge scale were widely interpreted, indicating that students' responses reflect a range of understandings of the questions. Students found that some of the Critical Thinking scale items appear to be either inter-related or too similar to each other. A final recommended questionnaire draft (included in Section 6.6) appeared to have resolved the majority of issues surrounding interpretation of these items.

- Students prefer shorter questions with fewer examples, and find them less prone to misinterpretation.

- When shown different versions and layouts of the survey students favoured the one that included a wider range of choices in the response scale, but one which did not require students to give a numerical response for how many times they had done a particular activity. In the response categories, for some students 'sometimes' and 'often' meant a couple of times in a period of two years and for others it meant a few times a month.

- Students voiced a concern of how their responses would reflect on their course since they did not want to appear to be criticising it.

6.2 Introduction

This research project took a mixed methods approach, combining an analysis of the literature and related international efforts to modify survey questions, in addition to new primary data collection through individual interviews, to explore student understanding of the survey items used in this project. Established research methods were rigorously applied, providing a wide set of data to validate and refine the student engagement questions. The research was designed to supplement the analysis of the data yielded by the pilot, to evaluate the validity and reliability of the original 14 items and make recommendations for improvements.

Participants

A diverse range of students from a variety of institutional types participated. A total of 43 students were interviewed, with students from: King's College London, University of Warwick, Canterbury Christ Church University, University College London (UCL), University of Oxford, University of Bath, Glasgow Caledonian University, Birkbeck College University of London, University of Manchester and University of East Anglia. Five of these institutions took part in the survey pilot.

There was a mix of first, second, third, fourth and fifth year students interviewed, studying a variety of subjects: American studies, biochemistry, classics, computing, engineering (mechanical, biochemical, electrical power, mechanical & electrical), English language & communication, English literature, environmental management, film studies, history, law, mathematics, medicine, midwifery, music performance, pharmacy, photography, politics, psychology, radiotherapy & oncology, social sciences, social work, and theology & religious studies. There was an equal gender representation and participants included UK, EU, non-EU, mature and part-time students.

6.3 Methodology

The project was based on a review of the literature on student engagement, with a focus on student engagement surveys and the development of the National Survey of Student Engagement (NSSE) (Kuh 2001). There was an

analysis of international examples of adapting US-based NSSE items, particularly the efforts in Canada to amend NSSE items, Australia (through the Australian Survey of Student Engagement, AUSSE) (Coates 2010), South Africa (through the South African Survey of Student Engagement, SASSE) and recent engagement pilot surveys in Ireland. Changes to wording (or not) and any accompanying analysis of validity were examined. Overall, there is a dearth of published validity and reliability testing of student experience surveys.

Stage 1

Individual interviews with students were conducted. Interviews followed in the tradition of Tourangeau (1984):

1. Comprehension of the question (question intent and meaning of terms).
2. Retrieval from memory of relevant information (recall strategy).
3. Decision processes (motivation and sensitivity/social desirability).
4. Response processes (mapping the response).

The 'think-aloud' method (Willis et al 1999) was used, which directs students to 'think aloud' as they respond to the question, with little interference from the interviewer. This was followed by using verbal prompts, such as 'when you answered 'sometimes', how often does that mean?' The research protocol included scripted probes, although spontaneous probes were used as appropriate. Questions were asked with NSSE-based response categories, expanded response categories and without specified categories. Scripts were updated and tested as interviews progressed.

Stage 2

After the initial set of 14 interviews, the original questionnaire used in the pilot (A) and alternative versions (B and C) were modified after initial comments from students, and new versions were created (A1 and B1). After another set of interviews we conducted a focus group with students and the questionnaires were then rewritten and restructured based on the feedback of all interviews conducted up to then (A2). This aimed mainly at simplifying and shortening questions, as well as grouping them differently. This process was conducted one further round, resulting in the recommended questions, order of items and response categories (A3). This final set of questions, recommended for future use, is included at the end of the chapter.

Stage 3

To measure reliability, two methods were used:

- test-retest reliability
- alternate-form reliability

A test-retest method was used with a select group of students, which had them complete survey items at the beginning and end of data collection. Alternate-form reliability was established through asking questions with different response categories or slightly reworded questions. Analysis of the interview data was used to judge the validity and robustness of the items. Analysis was also conducted across the different institutional and student variables for any differences.

6.4 Findings

Overall

Students were very positive about the survey:

> "It's got better answer choices than usual. Normally it is scaled and it's not very clear" (Second year, mechanical engineering.)

> "Looks quite straightforward. Looks great!" (Second year, English literature.)

> "It's clear and clean" (Fourth year, pharmacy)

One note is that even though the questions asked 'in the current academic year' it was often the case that some students answered the question thinking about their overall experience at university rather than the current academic year.

Item analysis

After analysis of the interview data, the 14 items were grouped by the scales (and described below). The Critical Thinking scale included the four items addressing critical thinking and depth of learning. Course Challenge included the two items addressing the challenge of the course and working hard and one about unpreparedness. The items in the Academic Integration and Collaborative Learning scales seemed to work well and students were positive about all the items. Some students, though a clear minority, expressed the need to be given examples of what is exactly being asked in some of the items in this section.

Critical Thinking

The grouping of items from the Critical Thinking scale was somewhat problematic. Firstly it was unclear for some students what 'coursework' actually meant. Students found some of the items were either inter-related or that each item was too similar to the others. Also, some responses pointed to the fact that the difference between 'analysing' and 'evaluating' was not clear, particularly from students outside of the social sciences and humanities. One student commented on the survey overall: "Pretty much bang on with exception of questions [3] and [1] [evaluating and analysing]." (First year, politics). Originally, the items were phrased as follows:

> During the current academic year, how much has your coursework emphasised the following mental activities?
> **Item 1**: 'Analysing in depth an idea, experience or line of reasoning'
> **Item 2**: 'Forming a new idea or understanding from various pieces of information'
> **Item 3**: 'Evaluating a point of view, decision, or information source'
> **Item 4**: 'Applying facts, theories, or methods to practical problems or new situations'

We recommend these be changed as follows:

> During the current academic year, how often has your coursework emphasised the following?
> **Item 1**: 'Analysing ideas or theories in depth'
> **Item 3**: 'Evaluating a point of view, decision, or information source'
> **Item 2**: 'Forming a new idea or understanding from various pieces of information'
> **Item 4**: 'Applying knowledge to new situations'

After these items had been rephrased on modified versions of the survey students seemed to have less doubt about what was being asked in each item. Simpler, more direct and shorter questions appeared to have been preferred by all students, whether they were from the natural sciences or from the humanities. However, it is important to mention that there were noted disciplinary differences in students' understanding of these four items. Many science students stated it was not their role as undergraduate science students to either analyse, evaluate, form new ideas or even to apply knowledge to new situations (though this last item was relatively less controversial than the three previous ones). Students in the humanities and in the social sciences were more confident in replying to these questions. Still, for the majority of students it remained unclear what the differences may be between analysis and evaluation. It may make more sense to collapse the two items into one.

Course Challenge

This scale consisted of three items, one about unpreparedness (using the 'Very often-Never' scale) and two items asking about working hard, and being challenged by coursework (using the 'Very much-Not at all' scale in the final draft). The latter two were the most troublesome questions for students. Students were not clear what the questions were referring to, or what part of the question to respond to.

> **Item 5**: 'Worked harder than you thought you could to meet a tutor's/lecturer's standards or expectations' to be changed to 'Do you feel you have been pushed to work harder than you thought you could?' or deleted.

This question was almost unanimously criticised for being unclear. Very few students were happy to respond to this question without questioning what it actually meant. However, in a couple of institutions, students were told by their personal tutor what grade they were expecting students to get by the end of the year, and this expectation acted as a benchmark for students.

Students were not sure what 'worked harder than you thought you could' necessarily referred to, as students could not know their limits of working. Students were not sure if 'worked harder' referred to more quantity (reading more) or difficulty (more demanding texts and exercises).

Students were often not sure how they would know their tutor's expectations, or what those would be if they did. However, all students except one referred to their own expectations of their effort, rather than the expectations of their tutor. Additionally, students' responses depended on their background and level of preparation for their course, indicating the responses could be quite relative across courses. Also, students' conceptions of expectations varied across institutions, particularly those between research-intensive and other institutions.

> **Item 6**: 'During the current academic year, to what extent have your courses challenged you to do your best work?' to be changed to 'During the current academic year has your course challenged you to do your best work?'

This item was not clear to students, with a few choosing not to answer it at all. Some students responded to this question as 'did you get the grades you'd like to'. Several students were unclear about what 'challenge' meant, usually referring to workload, and others were not sure what their 'best work' could be. One student remarked: 'Being challenged is up to the individual…and possibly not to do with the coursework' (First year, film studies). Ultimately, because of the variety of interpretations, the meaning of the responses may not be clear.

> **Item 7**: 'Come to taught sessions unprepared (eg not completed assignments, readings, reports, etc)' to remain unchanged.

With this item, students indicating they had come unprepared for sessions 'sometimes', 'often' or 'very often', mostly meant they had not prepared at all for individual sessions. However, some students regularly went to class only partially prepared, and thus indicated in their response 'often' or 'very often', even though they had prepared somewhat for sessions. Answers to this item were split between the sciences and engineering, and the arts, humanities and social sciences groups of students. The former were more likely not to have prepared at all for these sessions whereas the latter were more likely to have done some preparation for taught sessions. Science and engineering students thought of 'taught sessions' as lectures, and often said they went to lectures unprepared, although they did prepare for labs and tutorials.

Collaborative Learning

> **Item 8**: 'Worked with other students on course projects or assignments' to remain unchanged.

> **Item 9**: 'Explained course material to one or more students' to remain unchanged.

These questions were very clear. Some courses do not have group projects or assignments but all students seemed to have interpreted this question in the same way.

Item 10: 'Discussed ideas from your course with others outside of taught sessions (students, family members, co-workers, etc.), including by email/online' to be either changed to 'Discussed ideas from your course with others outside of taught sessions (eg students, family members, etc), including by email', or to be deleted.

Many students saw this question has being very similar even a repetition of item 9 'Explained course material to one or more students'. Students in the social sciences and humanities often indicated they had talked to family members about their course. However this was not the case for natural sciences and engineering students, who often felt their course material was too specialised for others outside their degree to understand.

Academic Integration

Overall students' reactions to the items in this scale were positive. After our analysis we recommend the following amendments (for item 10 see above):

Item 11: 'Asked questions or contributed to course discussions in other ways' to be changed to 'Asked questions in class'.

It is important to note that in the majority of cases students in the natural sciences, medical sciences and engineering stated they never asked questions in lectures because it was not the practice of the discipline. They asked questions in labs, seminars or tutorials, which some students reflected in their responses and others did not. For the humanities and the social sciences students the practice of asking questions in lectures was very common.

Item 12: 'Discussed your academic performance and/or feedback with teaching staff' to remain unchanged.

This was something that very few students had done.

Item 13: 'Discussed ideas from your course with teaching staff outside taught sessions, including by email/online' to be changed to 'Discussed ideas from your course with teaching staff outside taught sessions, including by email'.

Students rarely indicated discussion outside of class occurring, particularly in the sciences and engineering, but when students did, it was mainly by email. Also students often felt this was a long-winded item.

Item 14: 'Talked about your career plans with teaching staff or advisors' to remain unchanged.

Students almost never did this. Students in their first or second year did not appear to see this as important. This item was perceived as being more relevant for final-year students.

Response categories

Various response categories were tested, including four categories (e.g. Very often / Often / Sometimes / Never), five categories (e.g. Very often/Often/Sometimes/Rarely/Never), frequency expressed as 'number of times' (e.g. 1-2 times, 3-5 times) and open scales (e.g. Very often <-> Never). Students were least comfortable with open scales and those requiring them to count the number of times they had engaged in something. They were most comfortable and familiar with the non-numeric category scales:

Very often/Often/Sometimes/Rarely/Never, and
Very much/Quite a bit/Some/Very little/Not at all

Students felt that there was a rather large jump between the 'sometimes' and 'never' scale points in the original four-category scale. The addition of a 'rarely' category was also adopted as part of previous cognitive interviewing conducted at King's College London.

However, what 'sometimes' meant to one student differed from what 'sometimes' meant to other students. Moreover, a student could reply 'often' to one question and 'often' to another and what they meant was considerably different, such as four times a year, or twice a month. Perhaps what is most important are the

expectations of regularity and priority when students are surveyed about their experience and their perception of their experience. A student may answer one question with 'sometimes' and by this they mean 'weekly' because they perceive that particular item to be high on their perception of what their student experience should contain. For another item, the same student may reply 'sometimes' meaning they have done the activity once a term. This relates to items students perceive should happen less frequently, such as, for example, discussing career plans or academic performance with staff.

6.5 Recommendations

Overall

Unanimously, students were enthusiastic about the idea of engagement questions. Students felt engagement questions showed that the institution valued students' experiences. Therefore, student engagement survey questions seem to be a valid and valued measure of the student experience.

Scales

Although there was some confusion, students valued the Critical Thinking questions. The recommendation is for these questions to be reordered and shortened, and possibly for the items related to 'analysing' and 'evaluating' to be combined. In light of disciplinary differences, these should be compared primarily within subject groupings.

The Academic Challenge questions should be interpreted with caution, extra emphasis added (as in the final questionnaire) or deleted.

The scales for Academic Integration and Collaborative Learning presented few issues for students. The main recommendation is to add an additional response category.

Response categories

Include a 'rarely' category for the frequency questions: Very often / Often / Sometimes / Rarely / Never. And amend the other scale to: Very much/ Quite a bit/ Some/ Very little/ Not at all.

6.6 Final recommended versions of questions

Critical Thinking

During the <u>current academic year</u>, how often has your coursework <u>emphasised</u> the following?
(Response options: Very often / Often / Sometimes / Rarely / Never)

1. Analysing ideas or theories in depth

2. Evaluating a point of view, decision, or information source

3. Forming a new idea or understanding from various pieces of information

4. Applying knowledge to new situations

Course Challenge

5. Do you feel you have been pushed to work harder than you thought you could?
(Response options: Very much / Quite a bit / Some / Very little / Not at all)

6. Has your course <u>challenged you</u> to do your best work?
(Response options: Very much / Quite a bit / Some / Very little / Not at all)

7. <u>During the current academic year</u>, about how often have you come to taught sessions unprepared
(eg not completed assignments, readings, reports, etc)
(Response options: Very often / Often / Sometimes / Rarely / Never)

Collaborative Learning

<u>During the current academic year</u>, about how often have you done each of the following?
(Response options: Very often / Often / Sometimes / Rarely / Never)

8. Worked with other students on course projects or assignments

9. Explained course material to other students

10. Discussed ideas from your course with others outside of taught sessions (eg students, family members, etc), including by email

Academic Integration

<u>During the current academic year</u>, about how often have you done each of the following?
(Response options: Very often / Often / Sometimes / Rarely / Never)

11. Asked questions in class

12. Discussed your academic performance and/or feedback with teaching staff

13. Discussed ideas from your course with teaching staff outside taught sessions, including by email

10. Discussed ideas from your course with others outside of taught sessions (eg students, family members, etc), including by email

14. Talked about your career plans with teaching staff or advisors

7. Conclusion

The main purpose of this report is to provide information about how well the selected NSSE items and scales work in the UK context, and to see what kind of picture it paints of student engagement in the UK.

Chapter 5 (and there is more detailed data available as a separately downloadable appendix) contains a number of interesting results concerning the differences between different groups of students, including the following:

- male students report a higher frequency of explaining course material to fellow students than female students
- part-time students appear to engage less often in a range of activities, most markedly collaborative learning;
- taught postgraduates report being more diligent than undergraduates in their preparation for class;
- the frequency with which students talk to academics about their career plans appears to increase as they move through their degree;
- students on STEM courses do not seem to feel that their courses emphasise the evaluation of information, students in the arts and humanities feel a similar way about the application of information;
  - for instance, around 40% of mathematical and computer sciences students feel that their coursework emphasises the evaluation of information very little, compared to less than 0.5% of historical and philosophical studies students;
  - and nearly 20% of students on European language and literature courses feel that their coursework emphasises the application of information very little, compared to less than 5% of business and administrative studies students;
- the students who responded to the UK pilot appeared to discuss ideas from their courses with academic staff more than Australasian students, but discuss career plans with staff much less often than US students.
- of all the countries and regions included in the comparison, the UK pilot had the lowest proportion of students stating that they had never discussed their academic performance with staff.

These are interesting results, but should be interpreted with caution given the limited number of institutions taking part. It is also unclear whether the differences in response indicate genuine differences in behaviour, or different expectations; as with all student survey data, triangulation and further investigation are key. It should also be noted that as is common with student surveys, some experiences of students in different disciplines seem to differ quite markedly. This means that (as with the NSS) the most appropriate comparisons, given sufficient data, are likely to be between students studying the same disciplines in different institutions, rather than between students studying different disciplines.

The quantitative and qualitative testing raised a number of concerns that will need to be carefully considered for next year's iteration of the project:

- the Critical Thinking scale emerged as problematic. Both the cognitive interviewing and the analysis of responses suggested that students found it difficult to differentiate between the different items. Those items use relatively sophisticated concepts, and they appear to have been insufficiently distinct in the pilot questionnaire;
- the cognitive interviewing also revealed problems with students' understanding of the items in the Course Challenge scale, and the quantitative analysis found that students' responses to one of those items, the negatively-phrased item about being unprepared for class, do not seem to relate with their responses to other items. This is likely to be a problem with the negative phrasing;
- the factor analysis suggested that the dimension of student engagement relating to Academic Integration is restricted to interaction with academic staff, and interaction in class. Discussions with other people, such as family members or fellow students, did not seem to be part of that dimension of their engagement.

These findings need to be considered in detail before the items and scales are used again. However, there is no evidence to suggest that the NSSE items used failed to work, and in fact substantial evidence that they are reasonably valid and reliable measures of students' engagement in their studies. The project has produced valuable findings about the implementation of NSSE items in the UK.

It is important to acknowledge what this report has not done. We have not undertaken multilevel modelling, to take account of the hierarchical nature of student survey data (students in departments, departments in institutions, etc). Nor have we undertaken an investigation of 'criterion' validity: the extent to which the survey

results are related to outcomes, such as how well students learn. These are both activities that can be pursued in future, as the project will be running again in 2014.

The real test of the project however will be whether or not the data makes a difference to learning and teaching. We know quite a lot about how large-scale survey data can be used for enhancement, through for instance triangulation, and the involvement of students as researchers (Buckley 2012). In the UK most of our experience is with surveys like the NSS, which measure students' perceptions of what they receive. How well those lessons transfer to a survey such as the NSSE is an open and interesting question, though there is a lot of useful information available from other parts of the world.[32] We have included institutional case studies to give a flavour of how the results might be used, but it is important that we gather more information about how engagement surveys fit into enhancement processes and, ultimately, whether and how they impact on student learning.

The emphasis of this project has been on enhancement rather than providing information for external stakeholders about institutional or course performance, and no institutional results have been made public. It is another open question whether external stakeholders in general, and prospective students (and their parents) in particular, would value information about students' engagement with their studies. The claims of surveys like NSSE to indicate 'educational quality' are perhaps stronger than those of surveys like the NSS, but the immediacy and accessibility of engagement data is probably less than is found with – for instance – satisfaction data. It is a strength of engagement surveys that they are based on pedagogical theories that are widely accepted by those working in higher education, but this has the potential to create a barrier for the uninitiated. While it may seem obvious to prospective students that should be satisfied, it may not seem obvious that they should be engaging in collaborative learning, or undertaking the analysis and synthesis of information.

Finally, there is a debate to be had about how 'UK-style' student engagement is measured, in the sense of student representation, student involvement in decision-making at various levels and other related activities. It may seem perverse that a UK pilot of an engagement survey does not address those practices of student representation, partnership and empowerment most often taken to constitute 'student engagement' in the UK. As described in Section 2.3, for this first year it was decided to only use existing (and tested) items and to focus on four core areas, and both of those considerations mitigated against attempting to address student involvement in decision-making. Nevertheless, that is perhaps the next major question for the development of engagement surveys in the UK, and it raises a number of issues. Can we really measure students' sense of empowerment, and if so how? Do we need to decide what level of student involvement in decision-making is appropriate, and is there a consensus about what that level would be? Do we use this kind of survey to investigate student participation in university governance, when only a small fraction of students will have been involved? How can we implement large-scales survey of student engagement in ways that implement and embody the idea that students are partners in the educational process?

The 14 items used in this pilot were drawn from a survey that was first administered in 2000. The UK as a whole (and excepting certain institutions) is late to the idea of engagement surveys, and this has been a first step in the process of catching up. What is needed now is to explore the ways in which such surveys can support the improvement of learning, what we need to do to capture the distinctive nature of student engagement in the UK, and what its impact is on how well students learn.

---

[32] A number of resources about how North American institutions have made use of NSSE data, including an extensive set of case studies, are available from: http://nsse.iub.edu/html/how_institutions_use_NSSE.cfm [accessed 28 October 2013].

References

Astin, A. (1984) Student involvement: A developmental theory for higher education. *Journal of College Student Development.* **25** (4), 297-308.

Bowman, N. (2009) Can 1st-year college students accurately report their learning and development? *American Educational Research Journal.* **47** (2), 466-496.

Buckley, A. (2012) *Making it count: Reflecting on the National Student Survey in the process of enhancement.* York: Higher Education Academy. Available from: http://www.heacademy.ac.uk/assets/documents/nss/Making_it_count.pdf [accessed 24 October 2013].

Campbell, C. and Cabrera, A. (2011) How sound is NSSE? Investigating the psychometric properties of NSSE at a public, research-extensive institution. *The Review of Higher Education.* 35 (1), 77-103.

Chickering, A. and Gamson, Z. (1987) Seven principles for good practice in undergraduate education. *AAHE Bulletin.* March issue: 3-7.

Coates, H. (2010) Development of the Australasian survey of student engagement (AUSSE). *Higher Education.* **60**, 1-17.

Creighton, J., Beasley, S. and Jeffreys, P. (2008) *Reading Student Survey 2008.* Reading: University of Reading. Available from: http://www.reading.ac.uk/web/FILES/cetl-aurs/Reading_Student_Survey_Report.pdf [accessed 30 September 2013].

Esquivel, S. (2011) *The factorial validity of the National Survey of Student Engagement.* Ph.D. Thesis, University of Tennessee, Knoxville. Available from: http://trace.tennessee.edu/cgi/viewcontent.cgi?article=2099&context=utk_graddiss [accessed 8 October 2013].

Gibbs, G. (2010) *Dimensions of quality.* York: Higher Education Academy. Available from: http://www.heacademy.ac.uk/assets/documents/evidence_informed_practice/Dimensions_of_Quality.pdf [accessed 14 October 2013].

Gibbs, G. (2012) *Implications of 'Dimensions of quality' in a market environment.* York: Higher Education Academy. Available from: http://www.heacademy.ac.uk/assets/documents/evidence_informed_practice/HEA_Dimensions_of_Quality_2.pdf [accessed 14 October 2013].

Gordon, J., Ludlum, J. and Hoey, J. (2008) Validating NSSE against student outcomes: Are they related? *Research in Higher Education.* **49**, 19-39.

Harvey, L. (2003) Student Feedback [1]. *Quality in Higher Education.* **9** (1), 3-20.

Kuh, G. D. (2001) *The National Survey of Student Engagement: Conceptual framework and overview of psychometric properties.* Bloomington, IN: Indiana University, Center for Postsecondary Research. Available from: http://nsse.iub.edu/pdf/psychometric_framework_2002.pdf [accessed 8 October 2013].

Kuh, G. (2009) The National Survey of Student Engagement: Conceptual and empirical foundations. *New Directions for Institutional Research.* 141, 5-20.

Kuh, G., Hayek, J., Carini, R., Ouimet, J., Gonyea, R. and Kennedy, J. (2001) *NSSE technical and norms report.* Bloomington, IN: Indiana University Center for Postsecondary Research. Available at http://cpr.iub.edu/uploads/Norms!%20i-44.pdf [accessed 8 October 2013].

Kuh, G., Kinzie, J., Cruce, T., Shoup, R. and Gonyea, R. (2007) *Connecting the dots: Multi-faceted analyses of the relationships between student engagement results from the NSSE, and the institutional practices and conditions that*

foster student success. Bloomington, IN: Indiana University Center for Postsecondary Research. Available from: http://nsse.iub.edu/pdf/connecting_the_dots_report.pdf [accessed 24 October 2013].

LaNasa, S., Cabrera, A. and Transgrud, H. (2009) The Construct validity of student engagement: A confirmatory factor analysis approach. *Research in Higher Education.* 50, 315-332.

McCormick, A., Gonyea, R. and Kinzie, J. (2013) Refreshing engagement: NSSE at 13. *Change: The Magazine of Higher Learning.* **45** (3), 6-15.

Marsh, H. (1987) Students' evaluations of University teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research.* **11** (3), 253-388.

Marton, F. and Saljo, R. (1976) On qualitative differences in learning - 1: Outcome and process. *British Journal of Educational Psychology.* 46, 4-11.

Nelson Laird, T., Shoup, R. and Kuh, G. (2005) Measuring deep approaches to learning using the National Survey of Student Engagement. *In: annual meeting of the Association for Institutional Research: Chicago, May 14-18 2005.* Available from: http://nsse.iub.edu/pdf/conference_presentations/2006/air2006deeplearningfinal.pdf [accessed 8 October 2013].

NSSE (2013) *A Pocket Guide to Choosing a College: Questions to Ask on Your College Visits.* Bloomington IN: Indiana University Center for Postsecondary Research. Available from: http://nsse.iub.edu/pdf/PocketGuide_2013Eng.pdf [accessed 30 September 2013].

Pascarella, E., Seifert, T. and Blaich, C. (2010) How effective are the NSSE benchmarks in predicting important educational outcomes? *Change: The Magazine of Higher Learning.* Jan-Feb, 16-22.

Patil, V., Surendra, N., Mishra S. and Donavan, T. (2007) Parallel analysis engine to aid determining number of factors to retain [Computer software]. Available from: http://smishra.faculty.ku.edu/parallelengine.htm [accessed 7 October 2013]. Developed as part of Patil, V., Surendra, N., Mishra S. and Donavan, T. (2008) Efficient theory development and factor retention criteria: A case for abandoning the 'Eigenvalue greater than one' criterion. *Journal of Business Research.* **61** (2), 162-170.

Payne, A. (2012) How satisfied and engaged are our students? *Yorktalk.* Winter,14. Available from: http://www.yorksj.ac.uk/pdf/YorkTalk%20-%20Winter%202012.pdf [accessed 30 September 2013].

Pike, G. (2006a) The dependability of NSSE scalelets for college- and department-level assessment. *Research in Higher Education.* **47** (2): 177-195.

Pike, G. (2006b) The convergent and discriminant validity of NSSE scalelet scores. *Journal of College Student Development.* **47** (5), 551-564.

Porter, S. (2011) Do college student surveys have any validity? *The Review of Higher Education.* **35** (1), 45-76.

QAA (2012) *UK quality code for higher education: Chapter B5 – Student engagement.* (Gloucester: Quality Assurance Agency, Available from: http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/Quality-Code-Chapter-B5.pdf [accessed 30 September 2013].

Ramsden, P. (1991) A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education.* **16** (2), 129-150.

Richardson, J., Slater, J. and Wilson, J. (2007) The National Student Survey: development, findings and implications. *Studies in Higher Education.* **32** (5), 557-580.

Ross, H., Cen, Y. and Zhou, Z. (2011) Assessing student engagement in China: Responding to local and global discourse on raising educational quality. *Current Issues in Comparative Education.* **14** (1), 24-37.

Scott, I. (2011) *Survey of student activity and engagement.* York: Higher Education Academy. Available from: http://www.heacademy.ac.uk/projects/detail/StudentEngagement_ResearchBid2011_Worcester [accessed 30 September 2013].

Strydom, J. and Mentz, M. (2010) *South African Survey of Student Engagement (SASSE): Focusing the student experience on success through student engagement.* Pretoria: Council on Higher Education. Available from: http://www.che.ac.za/sites/default/files/publications/SASSE_2010.pdf [accessed 30 September 2013].

Taylor, P., Koskela, J. and Lee, G. (2011) *Shaping history.* York: Higher Education Academy. Available from: http://www.heacademy.ac.uk/projects/detail/StudentEngagement_ResearchBid2011_Warwick [accessed 30 September 2013].

Tourangeau, R. (1984) Cognitive sciences and survey methods. In Jabine, T., Straf, M., Tanur, J. and Tourangeau, R. (eds.) *Cognitive aspects of survey methodology: Building a bridge between disciplines.* Washington, DC: National Academy Press. 73-100

Trowler, V. (2010) *Student engagement literature review.* York: Higher Education Academy. Available from: http://www.heacademy.ac.uk/assets/documents/studentengagement/StudentEngagementLiteratureReview.pdf [accessed 30 September 2013].

Wells, P. (2011) *The Higher Education Academy Postgraduate Research Experience Survey: PTES 2011 report.* York: Higher Education Academy. Available from: http://www.heacademy.ac.uk/assets/documents/postgraduate/PTES_report_2011.pdf [accessed 10 October 2013].

Willis, G., DeMaio, T. and Harris-Kojetin, B. (1999) Is the bandwagon headed to the methodological promised land? Evaluation of the validity of cognitive interviewing techniques. In Sirken, M., Herrmann, D., Schechter, S., Schwarz, N., Tanur, J. and Tourangeau, R. (eds.) *Cognition and survey research.* New York: Wiley. 133-154.

Zhang, Z. (2013) An exploratory study of Chinese undergraduate student engagement. *In: the Higher Education Academy's Surveys for Enhancement Conference: 16 May 2013.* Available from: http://www.heacademy.ac.uk/assets/documents/postgraduate/Surveys_conf_13/Zhe_Zhang_presentation.pdf [accessed 30 September 2013].

Contact us

The Higher Education Academy
Innovation Way
York Science Park
Heslington
York
YO10 5BR

+44 (0)1904 717500
enquiries@heacademy.ac.uk

The Higher Education Academy (HEA) is a national body for learning and teaching in higher education. We work with universities and other higher education providers to bring about change in learning and teaching. We do this to improve the experience that students have while they are studying, and to support and develop those who teach them. Our activities focus on rewarding and recognising excellence in teaching, bringing together people and resources to research and share best practice, and by helping to influence, shape and implement policy - locally, nationally, and internationally. The HEA supports staff in higher education throughout their careers, from those who are new to teaching through to senior management. We offer services at a generic learning and teaching level as well as in 28 different disciplines. Through our partnership managers we work directly with HE providers to understand individual circumstances and priorities, and bring together resources to meet them. The HEA has knowledge, experience and expertise in higher education. Our service and product range is broader than any other competitor.

www.heacademy.ac.uk | www.twitter.com/heacademy

To request copies of this report in large print or in a different format, please contact the communications office at the Higher Education Academy: 01904 717500 or pressoffice@heacademy.ac.uk