2013 University Experience Survey National Report

March 2014

Acknowledgements

The UES Consortium, consisting of Graduate Careers Australia (GCA) and the Social Research Centre (SRC), wish to acknowledge with warm gratitude the work of the many people involved in the research process. Without the contribution of the students who took the time to complete the 2013 UES and the invaluable assistance, advice and support provided by the participating institutions and their Survey Managers, these important data would not be available.

Our thanks go also to the (now) Department of Education, which funded this work and provided support throughout the project. Particular thanks go to Phil Aungles, Sam Pietsch, Dr Andrew Taylor, David de Carvalho, Gabrielle Hodgson and Ben McBrien.

The Consortium would also like to acknowledge the advice and support of the UES Project Advisory Group, the members of which included Prof. Shirley Alexander (University of Technology, Sydney), Graeme Bryant (GCA), Robert Dalitz (Universities Australia), David DeBellis (Flinders University), Greg Jakob (from Federation University Australia, representing the Regional Universities Network), Prof. Richard James (The University of Melbourne), Conor King (Innovative Research Universities), Martin Hanlon (University of Technology, Sydney), Caroline Perkins (Regional Universities Network), Dr Andrew Taylor (Department of Education), Jade Tyrrell (National Union of Students), Michael Wells (Tertiary Education Quality and Standards Agency), Sonia Whitely (SRC), and A/Prof. Anne Young (The University of Newcastle).

The following people assisted and advised in the preparation of the Consortium's bid for the project: Prof. Jim Barber (VC, The University of New England), Prof. Don Hine (Behavioural Sciences, University of New England), Prof. Sally Kift (D-VC, James Cook University), Prof. Kerri-Lee Krause (P-VC, University of Western Sydney), Dr Alan McAlpine ((then) President, National Association of Graduate Careers Advisory Services and Queensland University of Technology), Prof. Bev Oliver (D-VC, Deakin University), Prof. Pip Pattison (D-VC, The University of Melbourne), Darren Pennay (SRC), Prof. Paul Ramsden (University of London), Prof. David Sadler (D-VC, University of Tasmania), and Prof. David de Vaus (Dean, Social Sciences, The University of Queensland).

Dennis Trewin (Dennis Trewin Statistical Consulting) provided an independent technical review of the 2013 UES project.

The UES Project Director was Dr Noel Edge (GCA). The Project Team consisted of Graeme Bryant (GCA), Jenna von Carate (SRC), Tim von Carate (SRC), David Carroll (GCA), Graham Challice (SRC), Bruce Guthrie (GCA), Darren Matthews (GCA), Eric Skuja (SRC), and Sonia Whiteley (SRC). Andrew Ward (Senior Statistician, SRC) conducted and reported on the psychometric analysis of the University Experience Questionnaire presented in Appendix F.

For more information on the conduct and results of the 2013 UES, please contact the UES Consortium on ues@graduatecareers.edu.au.

Executive summary

The University Experience Survey (UES) was originally developed to provide a national architecture for collecting feedback on key facets of the higher education student experience and, in doing so, obtain important data on the levels of engagement and satisfaction of current commencing and later-year undergraduate students. The UES was developed and administered to 24 universities as a pilot in 2011. The full-scale UES was subsequently administered as an online and telephone-based survey in 2012 and an online-only survey in 2013.

The UES measures five facets of the student experience: Skills Development, Learner Engagement, Teaching Quality, Student Support and Learning Resources. The UES also contains demographic and contextual items to facilitate data analysis and reporting, and two open-response items that allow students to provide textual feedback on the best aspects of their higher education experience and those most in need of improvement (Section 1.1).

There were several major differences between the 2012 and 2013 UES rounds. First, while the 2012 UES collected data at the student level, the 2013 UES collected data at the course level. In 2013, students completing a combined or double degree were invited to provide feedback on both course elements, which were then treated as two responses in the data file. Second, while the sample frame for the 2012 UES was supplied by participating institutions, the sample frame for the 2013 UES was based on data from the Higher Education Information Management System (HEIMS). This ensured a nationally consistent approach to sampling that had not previously been possible. Finally, while the 2012 UES had been administered as a mixed-mode survey, the 2013 UES was administered entirely online to mitigate potential survey mode effects (Section 1.2).

Extensive consultation with the higher education sector suggested that the results of the 2013 UES should be reported as the percentage of surveyed students who expressed satisfaction with their higher education experience. One consequence of this is that the results presented in the 2013 UES report are not directly comparable to those presented in the 2012 report (Section 1.3).

As was the case in 2012, all 40 Australian universities participated in the 2013 UES. The fieldwork period ran from August to November 2013 (Section 2.1). The "in-scope" population consisted of 341,343 commencing and later-year students (Section 2.2). A stratified sampling approach was employed, with strata defined on the basis of institution and subject area (Section 2.3).

Students were sent one initial email invitation and up to five reminders. A hardcopy letter was sent to non-responding students for whom a postal address was available. Incentives were allocated on an institutional basis, with $1,000 worth of prizes drawn for each institution. Concurrent to these activities, an engagement strategy targeting both students and higher education professionals was conducted to build awareness of the UES and encourage participation (Section 2.6).

The response rate for the 2013 UES was 29.3 per cent, which was up from 24.2 per cent in 2012, even without the aid of telephone follow-up. Institutional response rates ranged from 50.5 per cent to 17.9 per cent. Responses were received from 100,225 students, which equated to 108,940 valid surveys once combined and double degrees were taken into account (Section 3.1).

The sample of secured responses closely matched the in-scope population on most characteristics, but males were notably under-represented (Section 3.2). Post-stratification weighting to correct the gender imbalance in the sample of secured responses did not have any impact on the results at a national level, so it was decided to analyse the data without applying weights (Section 3.3).

Basic national results

Percentage satisfied results for all five UES focus areas and two key questionnaire items are presented in the table below, stratified by stage of studies. Because the focus areas represent different facets of the student experience, it is inadvisable to make comparisons across them. It is interesting, however, to note the aspects with which a smaller proportion of students expressed satisfaction.

	Focus areas					**Questionnaire items**	
	Skills Development	**Learner Engagement**	**Teaching Quality**	**Student Support**	**Learning Resources**	**Quality of entire educational experience**	**Quality of teaching**
Commencing	77	56	81	57	86	82	81
Later year	82	58	76	48	79	76	76
Total	**79**	**57**	**79**	**53**	**83**	**79**	**79**

Percentage satisfied results varied a little on the basis of demographic and contextual characteristics, but considerably on the basis of subject area. The narrowest range in results across subject areas was observed in relation to the student support focus area, with 16 percentage points separating the subject areas with the highest and lowest results. The widest range was observed in relation to learner engagement, with 33 percentage points separating the highest and lowest subject areas. Some notable cases of variation within broad fields of education were also observed, which underscores the fact that broad disciplinary aggregations can hide much useful detail (Section 4.1).

When the results from the 2012 and 2013 UES collections are compared (see table below), the largest difference in terms of focus area results was seen in relation to skills development, with percentage satisfied results of 82 and 79, respectively. Considering the quality of teaching questionnaire item, 83 per cent of students indicated their satisfaction in 2012, compared with 79 per cent in 2013. It is important to note, however, that these differences may be due to changes in the survey methodology. In any case, most differences between years were relatively small at a national level (Section 4.2).

	Focus areas					**Questionnaire items**	
	Skills Development	**Learner Engagement**	**Teaching Quality**	**Student Support**	**Learning Resources**	**Quality of entire educational experience**	**Quality of teaching**
2012	82	58	81	53	82	80	83
2013	79	57	79	53	83	79	79

When considering the individual questionnaire items, it is interesting to observe that many of the highest results relate to learning resources, including library resources and facilities, teaching spaces and online learning materials. Moreover, 79 per cent of students indicated satisfaction with both the quality of teaching and their entire educational experience. Conversely, the lowest results were observed for items in the student support focus area (Section 4.3).

When institutional percentage satisfied results on the quality of teaching and the entire educational experience items are ordered, there is a fairly even increase in results from the bottom of the distribution to near the top, with a few institutions at the top of the distribution notably higher than the majority of institutions (Section 4.4).

Comparing results from the UES to the National Survey of Student Engagement (NSSE), conducted in the USA and Canada, suggests that respondents to the NSSE are more likely to be satisfied with their educational experience than respondents to the UES, especially amongst later-year students. It is critical to note that this result does not necessarily mean that Australian students are less satisfied with their educational experience than their North American counterparts, although this may indeed be the case. It is possible, however, that this result could relate to the nature of universities that participate in NSSE, different distributions across fields of education, or other methodological differences between the two surveys (Section 4.5).

In addition to questions on their higher education experience, students were also asked to indicate whether they had seriously considered leaving their university in 2013. Overall, 17 per cent indicated that they had considered leaving. Commencing students, older students, Aboriginal and Torres Strait Islander students, students with a disability, external/multi-modal students and those who were first in their family to attend university were the most likely to consider early departure, as were those who had achieved low grades to date (Section 4.6).

Recommendations

Based on the experience of conducting the 2013 UES, the following recommendations are made for future iterations of the survey:

Recommendation 1: The in-scope population should comprise commencing and non-commencing students to achieve consistency in the population frame across institutions (Section 2.2.2).

Recommendation 2: The sample frame for individual institutions participating in the UES should be derived using population data from the Higher Education Information Management System (HEIMS) (Section 2.3.1).

Recommendation 3: Early advice should be provided to institutions on future administration of the UES to permit inclusion of non-standard institution-specific items (Section 2.4.2).

Recommendation 4*:* The UES should be administered as an online survey to mitigate potential survey mode effects, though targeted engagement and follow-up of students should be undertaken where required to improve representativeness of survey data (Section 2.5).

Recommendation 5: The analytic unit for the UES should be the course undertaken by the student (Section 2.8.1).

Recommendation 6*:* For purposes of international benchmarking, the CEQ Good Teaching Scale, Generic Skills Scale, Clear Goals and Standards Scale and Overall Satisfaction Item should be administered on the UES to a sample of non-commencing students across all participating institutions (Appendix C).

Recommendation 7: The item, “At university during year x, to what extent have you used university services to support your study?” should be omitted on the basis of psychometric evidence of lack of fit within the Student Support focus area (Appendix F).

Contents

List of tables

List of figures

1. Introduction and overview

1.1 Origin of the University Experience Survey

The University Experience Survey (UES) was developed to provide a national architecture for collecting feedback on key facets of the higher education student experience and, in doing so, obtain important data on the levels of engagement and satisfaction of current students. The UES focuses on aspects of the higher education student experience that are measurable, linked with learning and development outcomes, and for which institutions can reasonably be assumed to have responsibility. The UES was developed in 2011 on behalf of the Australian Government by a consortium led by the Australian Council for Educational Research (ACER) and including the University of Melbourne's Centre for the Study of Higher Education (CSHE) and the Griffith Institute for Higher Education (GIHE). The Consortium designed and validated the University Experience Questionnaire (UEQ) and data collection methodology, and presented recommendations about further development. In 2012, the Consortium was engaged to review and readminister the UES. The UES was refined to be relevant to policy and practice, and to yield robust and useful data that could be used for informing choice and continuous improvement. Linkages were made to facilitate international benchmark comparisons. The 2012 UES was administered as an online and telephone-based instrument.

The UES was designed to measure five facets of the student experience: Skills Development, Learner Engagement, Teaching Quality, Student Support and Learning Resources. These are operationalised by means of summated rating scales, underpinned by 47 individual questionnaire items. These items were supplemented by two open-response items that allow students to provide textual feedback on the best aspects of their higher education experience and those most in need of improvement. The UES also contains two additional sets of items, demographic and contextual, to facilitate data analysis and reporting. A full list of questionnaire items is presented in Appendix A.

1.2 The 2013 UES

In mid-2013, the tender to administer the second iteration of the UES proper was awarded to a consortium consisting of Graduate Careers Australia (GCA) and the Social Research Centre (SRC), hereafter GCA-SRC. Given the tight timeframe for the 2013 UES, with data collection and reporting to be undertaken by the end of the year, the decision was made to administer a largely unchanged instrument. In terms of the instrument, the major difference between the 2012 and 2013 collections was that, while the former collected data at the student level, the latter collected data at the course level.[1] In other words, a student completing a double or combined degree was invited to provide feedback on both course elements in the 2013 UES. As was the case in the 2012 UES, institutions were given the opportunity to add institution-specific items to the UES. These institution-specific items were only presented to students after they had completed and submitted the UEQ component,[2] resulting in a clear demarcation between the two. Moreover, six scales from the Course Experience Questionnaire (CEQ) component of the Australian Graduate Survey (AGS) were administered to students from several institutions in order to investigate the empirical relationship between the UES focus areas and CEQ scales. As with the institutional-specific items, the six CEQ scales were administered only after respondents had completed and submitted the UEQ component.

[1] Throughout this report, "course" is synonymous with "degree program".
[2] "UEQ" is used where necessary in this report to distinguish between the UES proper and additional non-standard elements, such as institution-specific items and CEQ scales.

The data collection methodology of the 2013 UES differed substantially from that of the 2012 survey. While the sample frame for the 2012 UES was based on a "bottom-up" approach, with participating institutions providing extracts from their student data systems to the 2012 UES Consortium to serve as a basis for the sample frame, population data from the Higher Education Information Management System (HEIMS) was used to create the sample frame for the 2013 UES. This ensured a nationally consistent approach to sampling that had not previously been possible.[3] Moreover, while the 2012 UES had been administered as a mixed-mode survey, the 2013 UES was administered entirely online to mitigate potential survey mode effects. Table 1 presents an overview of the main elements of the 2012 and 2013 UES collections.

Table 1. UES project overview, 2012 and 2013

Project element	2012	2013
Number of participating institutions	40	40
Number of "in-scope" students	445,332	341,343
Data collection period	July-October	August-November
Data collection mode	Online and telephone	Online
Overall response rate (%)	24.2 (online 21.1)	29.3
Number of completed surveys (student level)[a]	110,135	100,225
Number of valid surveys (course level)	N/A	108,940
Analytic unit	Student	Course

[a] When the 2013 completion rules are applied to the 2012 data, an additional 6,908 students are excluded, which yields a total of 103,227 completed surveys. These completion rules are defined in Section 2.8.1.

1.3 Reading the results in this report

It is essential that the analysis and reporting of the UES data are conducted in statistically sound and appropriate ways. Since its introduction, the UES data have been reported in two metrics: average scores and percentage satisfied results. Average scores are based on a rescaling of the response scales, with the four-point scales recoded onto a scale that runs 0, 33.3, 66.6 and 100, and five-point scales recoded onto a scale that runs 0, 25, 50, 75 and 100. Scores for each focus area are then computed as the mean of the constituent item scores. Percentage satisfied results reflect the percentage of students who report a focus area score of 55 or greater. This specific value was chosen because it is clearly above the midpoint of the response scale and reflects the maximum percentage of graduates satisfied with their higher education experience. At the individual response level, satisfaction is represented by a binary variable taking the value of one if the student is satisfied with a particular facet of their higher education experience and zero otherwise.

Extensive consultation with the higher education sector indicated a near-universal preference for the reporting of percentage satisfied results over focus area average scores. Percentage satisfied results were seen as being a more understandable measure, especially for less expert users of the UES data, and are straightforward for institutions to replicate and benchmark against. As such, percentage satisfied results are presented throughout this report. In relation to UES focus areas, "percentage satisfied" reflects the percentage of students who give a focus area score of 55 or greater out of a possible 100. In cases where the results on individual UES items are reported, percentage satisfied reflects the percentage of responses in the top two response categories. One consequence of this is that

[3] It is likely that the difference in the number of "in-scope" students in 2012 and 2013 (see Table 1) is due, at least in part, to these differences in sample design between the two surveys. Differences in the definition of what constitutes a later-year student is also likely a contributing factor (see Section 2.2.2).

the results presented in the 2013 UES report are not directly comparable to those presented in the 2012 report. Information on the production of UES focus area average scores and associated percentage satisfied variables is presented in Appendix B.

1.4 Important caveats

While the UES has been trialled and administered twice using best practice techniques, it is critical that certain caveats are borne in mind when interpreting the results in this report, especially in cases when these results are to be used to inform policy decisions. First, it is possible that the results are biased to some extent by the fact that not all members of the target population return a completed response. If non-respondents differ systematically from those who did respond to the UES, the results will not reflect the true experiences of students in the broader higher education student population. If, for example, students who are more engaged with their higher education institution tend to be more likely to respond to the UES than those who are less so, the estimates relating to Learner Engagement may be upwardly biased relative to the true population parameter, or vice-versa.

Post-stratification weighting is a common method employed to ensure that the sample of responses reflects the survey population in terms of key demographic and enrolment characteristics. Exploratory analysis, discussed further in Section 3.3, suggested that corrective weighting does not provide any significant advantage for the 2013 UES. As such, the results presented in this report are based on unweighted data. It is important to note that bias on the basis of unobservable characteristics may still be influencing the results. This type of bias may not be corrected by the application of post-stratification weights in any case. It is difficult to say whether any such bias exists without gathering data from non-respondents, which was not undertaken as part of the 2013 UES. Readers are asked to consider the possible existence of such bias when interpreting the results in this report.

Moreover, it is important to bear in mind that the UES does not encompass all the aspects on which students could evaluate their courses, nor does it specifically measure the relative importance that students place on different aspects of their experience.

Finally, an optimal analysis of UES results should account for the inherent hierarchical structure of the data. Students are nested within subject areas and subject areas are nested within institutions. The experience of students within the same subject area (academic department) may be correlated due to exposure to the same lecturers. Likewise, experience scores for subject areas may be correlated within an institution due to the similar background of the students who attend that institution. To minimise bias, UES data should ideally be analysed using a multilevel model that accounts for this nesting of students within subject areas, within institutions. In order to maintain consistency with past reports and other presentations of UES data, to ensure that the results are clear and meaningful to the widest possible audience, and to enable the results presented herewith to be easily replicated, only single-level analyses are presented in this report.

1.5 Overview of this report

This report presents an overview of the 2013 UES, including the conduct and administration of the survey, and key results based on the national UES data file, which consists of 108,940 responses from 100,225 students representing 40 higher education institutions. All statistics relating to UES focus areas and their constituent items reflect the percentage of students who indicated that they were satisfied with their higher education experience. The UES focus areas relate to Skills Development,

Learner Engagement, Teaching Quality, Student Support and Learning Resources. Selected statistics are presented with 95 per cent confidence intervals to demonstrate the variability of estimates due to sampling variation. Summary statistics on the reasons why students considered leaving their current university are also shown. Recommendations for enhancing future iterations of the UES are presented in boxed text throughout the report. Supplementary analyses and additional materials are presented in appendices and referenced in the body of the report.

2. Methodology

2.1 Institutional participation

All 37 Table A and 3 Table B higher education institutions participated in the 2013 UES. Under the Higher Education Support Act (HESA) 2003, Table A institutions are self-accrediting providers, eligible for funding under the Act. Table B institutions are also self-accrediting, but are not eligible for general Commonwealth funded places. For the purpose of administering the UES, participating institutions were assigned to operational cohorts based on fieldwork commencement date. As a result, the 40 participating institutions were split into 12 operational cohorts; the largest comprising eight institutions and the smallest comprising single institutions. Table 2 lists the institutions in each cohort, along with corresponding fieldwork commencement and completion dates.

Table 2. Operational cohorts for the 2013 UES

Operational cohort	Institution	Commencement date	Completion date
Cohort 1	University of Western Sydney	August 12	October 01
	Southern Cross University		October 16
	Griffith University		October 01
	The Australian National University		October 01
	University of Tasmania		October 01
	University of South Australia		October 16
	University of the Sunshine Coast		October 16
	Charles Sturt University		October 01
Cohort 1a	University of New South Wales	August 14	October 16
	Murdoch University		October 16
	Edith Cowan University		October 01
	Charles Darwin University		October 22
	Central Queensland University		October 01
	University of Southern Queensland		October 16
	James Cook University		October 16
Cohort 1b	The University of Sydney	August 15	October 16
	The University of Melbourne		October 01
	University of Technology, Sydney		October 16
Cohort 1c	Macquarie University	August 19	October 22
Cohort 1d	University of Newcastle	August 20	October 22
	Deakin University		October 22
	Monash University		October 23
Cohort 1e	Flinders University of South Australia	August 21	October 22
	The University of Adelaide		October 16
	University of Canberra		October 16
	MCD University of Divinity		October 16
Cohort 1f	University of Wollongong	August 22	October 29
	Federation University Australia		October 22
	University of New England		October 22
	Australian Catholic University		October 22
Cohort 2	Victoria University	September 2	October 29
	Swinburne University of Technology		October 29
	The University of Queensland		October 29
	La Trobe University		October 29
	Queensland University of Technology		October 29
	The University of Western Australia		October 22
Cohort 3	Bond University	September 25	November 13
Cohort 3a	Curtin University of Technology	September 26	November 19
Cohort 3b	The University of Notre Dame Australia	October 01	November 19
Cohort 4	RMIT University	October 21	November 25

Survey commencement dates were contingent upon three main factors:

- Programming requirements for individual institutions, including the core UEQ setup, sample and reporting link setup, branding (addition of institutional logos), institution-specific questionnaire item setup (where applicable) and associated quality assurance checks.
- Institutional "blackout" dates, including conflicting student surveys, examination periods and semester breaks.
- Institution-specific confidentiality and privacy restraints concerning the release of student information to third parties.

2.2 Survey population

The in-scope survey population for the 2013 UES consisted of commencing and later-year onshore undergraduate students enrolled in Table A and B higher education institutions. The in-scope population consisted of 341,343 students. The definitions used for commencing and later-year students in the 2013 UES are presented in the following subsections.

2.2.1 Commencing students

Commencing students were defined as first-year students who were enrolled in an undergraduate course, studying onshore, commenced study in the relevant target year and enrolled for at least one semester. Records conforming to the agreed definition of a commencing student were extracted from the national HEIMS Submission 1 student file by the Department. Individual institutions were then asked to confirm, where possible, that the selected students were still enrolled.

2.2.2 Later-year students

Later-year students were defined as final-year students who were enrolled in an undergraduate course, studying onshore and generally in their third year of study. Defining later-year students proved to be difficult in practice and a number of different approaches were trialled. The task is relatively easy for full-time students in three-year courses, but more difficult for part-time and external students, those who took leaves of absence and those who transferred from one course to another. To address this, two different options were used in practice. Option 1, based on the ratio of EFTSL completed successfully (E355) and currently in progress (E399) to the total EFTSL for the course (E350) proved the better option for 9 of the 40 institutions.[4] Option 2, which adjusts for attendance mode (E330) and course duration (E350) was the standard solution used for the remaining 31 institutions. Records conforming to the agreed definition of a later-year student were extracted from the HEIMS Submission 1 Student File and individual institutions were asked to confirm, where possible, that the selected students were still enrolled.

Given the difficulty in consistently defining later-year undergraduate students across institutions, a better approach for future UES collections may be to change the in-scope population to commencing and non-commencing students. This can be defined using HEIMS data element 922, which takes the value of 1 for commencing and 2 for non-commencing students. Although this change would remove the capacity of the UES to specifically investigate the experience of students in the final year of their courses, there is no guarantee that the approach trialled in 2013 achieved this end. Changing the in-

[4] The numbers in parentheses refer to HEIMS data elements.

scope population to commencing and non-commencing students defined on the basis of one HEIMS element would at least achieve consistency across institutions. This is crucial if the UES results are to be used for the purposes of cross-institutional benchmarking.

Recommendation 1: The in-scope population should comprise commencing and non-commencing students to achieve consistency in the population frame across institutions.

2.3 Sampling design

2.3.1 Sample frame

The sample frame for the 2013 UES was based on a "top-down" approach using population data from HEIMS to create the sample frames for individual universities. Compared with the "bottom-up" approach utilised for the 2012 UES, whereby institutions provided extracts from their student data systems to the survey administrators to serve as a basis for the sample frame, the approach adopted for the 2013 UES minimised the likelihood of accidental bias being introduced to the sample selection process and ensured a nationally consistent approach to sampling. While it would have been ideal to use validated Submission 2 data for this purpose, this was not possible due to the timeline for data collection. To address this, each institution was asked to verify, where possible, whether or not the selected students were still enrolled.

Recommendation 2: The sample frame for individual institutions participating in the UES should be derived using population data from the Higher Education Information Management System (HEIMS).

2.3.2 Approach to sampling

A stratified sampling approach was employed for the 2013 UES collection, with strata defined on the basis of institution and subject area.[5] The number of students to be sampled in each stratum was calculated using the response assumptions given in the 2012 UES National Report.[6] This means that all students were selected for strata up to 1,333 students, effectively a census of these strata. For strata larger than 1,333 students, a random sample of 1,333 was drawn. For the 2012 UES collection, the target response numbers in these large strata was 200. This value was derived from a desire for error bands of ±5 per cent at a 95 per cent level of confidence.[7]

In a deviation from this blanket approach, the target number of responses in these large strata in 2013 were calculated for each individual stratum as the number of responses required to achieve the target error band, with response data from the 2012 UES collection forming the basis of these calculations. Depending on the stratum, the response target ranged from around 200 to 300 students. The sample selection was validated against population parameters to ensure that appropriate proportions of gender, qualification, mode of attendance, subject area and citizenship characteristics were present in the sample.

[5] Subject area definitions are presented in Appendix H.
[6] Radloff, A., Coates, H., Taylor, R., James, R. & Krause, K. (2012). *2012 University Experience Survey National Report*. Retrieved 15 January, 2014, from http://www.innovation.gov.au/highereducation/Policy/Documents/UES2012NationalReport.pdf
[7] These error bands were calculated on the basis of average scores, not percentage satisfied results.

2.4 Additional questionnaire elements

2.4.1 Course Experience Questionnaire trial

As part of the 2012 UES, a trial of the CEQ was undertaken with a number of volunteer institutions. It was recommended for the 2013 UES that another trial of the CEQ be undertaken to provide more conclusive evidence regarding the role or place of the CEQ in relation to the UES. Six CEQ scales comprising 28 items were administered on the UES as part of this trial:

- Good Teaching Scale (GTS)
- Clear Goals and Standards Scale (CGS)
- Generic Skills Scale (GSS)
- Overall Satisfaction Item (OSI)
- Graduate Qualities Scale (GQS)
- Learning Community Scale (LCS)

During the development of the 2013 UES, it was acknowledged that the wording of the CEQ was aimed at students who had graduated from their course and that the tense of these items may not be appropriate for commencing students in the first year of their studies. As such, it was decided to trial a present tense version of the CEQ. Students identified as commencers in the first year of study in their current course received a present tense version of the CEQ and all other students from the institutions participating in the trial received the original, past tense version of the CEQ.

Fourteen institutions participated in the 2013 CEQ trial. These are listed in Table 3.

Table 3. Institutions that participated in the 2013 CEQ trial

Bond University	Swinburne University of Technology
Charles Darwin University	The University of Canberra
Flinders University	The University of Newcastle
La Trobe University	The University of Western Australia
Monash University	University of Tasmania
MCD University of Divinity	University of the Sunshine Coast
Southern Cross University	University of Western Sydney

An empirical comparison of UES focus areas and CEQ scales is presented in Appendix C. In brief, this analysis suggests that there is considerable overlap between the five UES focus areas and the six CEQ scales administered as part of the trial. This finding raises the question of whether there is any utility in administering CEQ items as a standard component of the UES, since the former appears to be measuring facets of the student experience already addressed by the latter. Given that respondents to the UES are already tasked with completing 47 items on their higher education experience, the addition of potentially redundant items may be seen as increasing respondent load without improving measurement of the UES focus areas to an appreciable degree. An argument can be made, however, that administering selected CEQ scales to currently enrolled students on the UES would allow for additional international benchmarking to be undertaken, as a result of the CEQ and the UK National Student Survey (NSS) having several questions with very similar wording. Given the importance of international benchmarking to the Australian higher education sector, the inclusion of several CEQ scales on the UES is justified on this basis.

2.4.2 Institution-specific items

Institutions were offered the option of including non-standard, institution-specific items as part of the 2013 UES. Three institutions chose to include their own items, which were added to Module 5 of the online survey, administered after the UEQ component (see Section 2.5). It is likely that the limited take-up of institution-specific items was related to the tight timeframe for the 2013 UES. If given a longer lead time, more institutions would likely take up this offer.

Recommendation 3*:* Early advice should be provided to institutions on future administration of the UES to permit inclusion of non-standard institution-specific items.

2.5 Data collection fieldwork

Unlike the 2012 UES, which was administered as a mixed-mode survey, with data collection via online survey and telephone interviewing, the 2013 UES was administered entirely online. This was adopted as a cost-effective approach that would still achieve the necessary response and to mitigate potential survey mode effects associated with combining self-completed and interviewer-administered questionnaires. The online survey was programmed and hosted by GCA-SRC. Students were required to login before completing the survey.

The UES items were organised into a modular structure for ease of administration and to allow for different module rotations to be presented to counteract potential order effects.[8] Table 4 contains the five module rotations that were randomly shown to students. Module 5 applied only to institutions opting to include additional, institution-specific items (see Section 2.4.2), and was not otherwise included in the rotation sequence. Moreover, Module 6 only applied to institutions participating in the CEQ trial (see Section 2.4.1), and was not otherwise included.

Recommendation 4: The UES should be administered as an on-line survey to mitigate potential survey mode effects, though targeted engagement and follow up of students should be undertaken where required to improve representativeness of survey data.

Table 4. 2013 UES module rotations

Version A	Version B	Version C	Version D	Version E
Module 2 UES 1	Module 1 Demographics	Module 1 Demographics	Module 2 UES 1	Module 1 Demographics
Module 3 UES 2	Module 4 UES 3	Module 3 UES 2	Module 3 UES 2	Module 2 UES 1
Module 4 UES 3	Module 3 UES 2	Module 2 UES 1	Module 4 UES 3	Module 3 UES 2
Module 1 Demographics	Module 2 UES 1	Module 4 UES 3	Module 6 CEQ	Module 4 UES 3
Module 5 Institution Q's	Module 5 Institution Q's	Module 5 Institution Q's	Module 1 Demographics	Module 5 Institution Q's
Module 6 CEQ	Module 6 CEQ	Module 6 CEQ	Module 5 Institution Q's	Module 6 CEQ

[8] There is evidence of a small but statistically significant order effect in the 2013 UES. Considering the quality of the entire educational experience item, Version B of the questionnaire was associated with an approximately one percentage point increase in the probability of a satisfied response relative to Version A, while Version E was associated with a two percentage point decrease. This analysis controls for stage of studies, gender, subject area and institution.

2.6 Student engagement strategy

GCA-SRC developed and disseminated a range of promotional materials to build awareness of the UES in the higher education sector and encourage participation amongst the student population, including a new logo for the survey. The development of these materials was informed by feedback from students who completed the 2012 UES. There were two main phases of student engagement for the 2013 UES. The first was an awareness-building campaign focusing on pre-survey engagement, which ensured that students were aware of the survey well in advance of the start of fieldwork. The response maximisation phase commenced after the survey was deployed and centred on scheduled invitation and reminder correspondence encouraging completion of the survey, and an incentive strategy. These are discussed in greater detail below.

2.6.1 Pre-survey engagement

The pre-survey engagement strategy utilised several different media, which were developed by GCA-SRC and, as required, disseminated to participating higher education institutions. Institutions were not obligated to use these promotional materials, but doing so was strongly encouraged. Some developed their own promotional materials to supplement those provided by GCA-SRC.

The following media formed part of the national engagement strategy:

- The UES website (www.ues.edu.au), which was developed as an information portal for all things related to the UES, including information about the survey, FAQs, results from the 2012 UES, contact details of the survey administrators and a login for students to complete the survey.
- Website tiles, which were designed to be placed on institutional websites, learning management systems, news feeds, student association websites, etc. Examples of these website tiles are presented in Appendix D.
- Media releases, which were distributed to media outlets at key milestones, such as the awarding of the contract to undertake the UES and the commencement of fieldwork.
- Institutional communications, including PowerPoint slides for teaching staff to use in lectures, emails to be sent by university management explaining the purpose and importance of the UES, and explanatory text to appear on learning management systems.
- Social media posts, which could be disseminated to students through institutional Facebook and Twitter accounts.

2.6.2 Response maximisation

Students were sent one initial email and up to five reminder emails. The majority of institutions opted to allow GCA-SRC to approach students directly via email, while four institutions chose to send the prescribed invitation and reminder emails themselves.

The email invitations were UES branded and included a hyperlink directly to the online survey as well as manual login and helpdesk details. Students were able to advise of a change to their enrolment status, opt-out of the survey or unsubscribe by reply email. Students who had completed a survey, those who had opted out of the survey and those who had been disqualified from participating on the basis of their enrolment status were removed from each email reminder sample file prior to the email reminders being sent. The standard email reminder schedule was as follows:

- Email reminder 1: 4 business days following initial invitation.
- Email reminder 2: 11 business days following initial invitation.
- Email reminder 3: 16 business days following initial invitation.
- Email reminder 4: 20 business days following initial invitation.

Email reminder 2 was timed to arrive shortly before the prize draw cut-off. Institutions were able to request variations to this schedule to ensure that students were not contacted at inopportune times, such as during non-teaching periods. Email reminders 3 and 4 used tailored text to target specific groups of students with low response rates. The majority of participating institutions also employed a fifth, targeted email reminder before the completion of fieldwork.

A hardcopy letter was sent to non-responding students for whom a postal address was available. The letter was timed to arrive at the same time as email reminder 2; prior to prize draw cut-off.

Incentives were allocated on an institutional basis, with $1,000 worth of prizes in the form of gift vouchers drawn for each institution. The incentives consisted of a major prize to the value of $500 and five runner-up prizes, each to the value of $100.

2.7 Higher education liaison strategy

Given that the success of the UES is contingent on the acceptance and support of the higher education sector, GCA-SRC placed a high priority on successful collaboration with the higher education sector throughout the 2013 UES. As was the case in 2012, the UES was overseen by the Project Advisory Group (PAG), comprising representatives and experts from across the sector.[9] The PAG met regularly during the course of the 2013 UES and advised GCA-SRC on many aspects of the 2013 collection, including timing, questionnaire issues, data collection, analysis, the provision of data to participating institutions, and reporting.

Due to the constricted timeframe for the 2013 UES, liaison with institutions began as soon as the tender was awarded in July 2013. Formal communications were issued to Vice-Chancellors at each Table A and Table B institution asking for their support of the project. Concurrently, an introductory email was sent to survey staff formally asking them to participate in the 2013 collection. Operational details relating to sampling, privacy, additional survey populations and institution-specific items were also obtained through this initial communication with institutional survey staff. A UES institutional administration guide was also prepared and sent to institutions to assist them in undertaking the UES at their institution, such as verifying the sample drawn from HEIMS by GCA-SRC. Promotional materials required to facilitate the pre-survey engagement were distributed to institutional contacts prior to the start of fieldwork. Moreover, the second day of GCA's annual two-day Survey Manager Information Forum (SMIF), held in late-July, was dedicated to the UES.

Regular updates were provided to institutions throughout the data collection phase of the project. The online survey included real-time reporting functionality, which allowed institutions to track their live response rates and report on key demographic variables while the survey was in field.

[9] A list of PAG members is given in the Acknowledgements.

Institutional representatives were invited by GCA-SRC to provide feedback on what they would like to see included in the national and institutional UES reports, and published data tables. Following the completion of data collection fieldwork, a feedback survey was made available to contacts from all participating institutions to obtain information on all aspects of the 2013 UES, which could be used to enhance the conduct and management of the 2014 collection.

2.8 Data processing

2.8.1 Definition of the analytic unit

The analytic unit for the 2012 UES was the student. The data file contained one record for each respondent to the survey. For the 2013 UES, changes to the instrument allowed students in double degrees to respond separately for each course element, which were treated as two separate responses for analytical purposes. As a result of this change, the analytic unit for the 2013 UES is the course. In the 2013 UES, a response was defined as valid and complete if the student had completed units in the course, there was a minimum of one valid UES focus area score, and, in the case of double degrees for which the student had at least one valid UES focus area score for each course, the courses were in two different subject areas. When double degree students had completed units in both components and they were in the same subject area, the first record was selected for analysis. Of the 100,225 students who completed the 2013 UES, 8,715 (8.7 per cent) furnished a valid response for their second course element, resulting in 108,940 valid responses.

Recommendation 5: The analytic unit for the UES should be the course undertaken by the student.

2.8.2 Data cleaning and preparation

To ensure consistency in the cleaning process, records were first merged from all separate institution level files (as collected on the online platform) into one master file. Sample variables were merged from the original population file for checking and to fill any sample data missing from the online collection platform as a result of students prematurely exiting the online questionnaire.

Revised course names entered by students were manually looked up against a master course list for the relevant institution. Where a course name matched multiple course codes, the student was assigned to the course with the highest enrolment where no conflicts between the different courses existed. Where an appropriate course code for the course name supplied by the student could not be found, queries were sent to the Survey Manager of the relevant institution. In cases where the Survey Manager advised that a combined course did not exist for two degrees listed by a student, they were treated as two unrelated concurrent degrees.

Following this process, the scope status of the student (i.e. whether they were enrolled in a degree eligible for the UES) was re-derived based on revised course level data. Students who had switched from an eligible undergraduate course to an ineligible course, such as postgraduate coursework or research, were excluded. All items in the body of the questionnaire were re-filtered to their respective bases to ensure there were no errant responses. After cleaning, normalised UES variables, UES scale variables and consolidated demographic variables were derived. In the case of double degrees, UES scale variables were derived separately for each course. After the data were finalised, the student level file was split to course level.

- Where a student was enrolled in a single degree, the student level record became the course level record.
- Where a student was enrolled in a double degree and had completed units in only one course, the student level record became the course level record.
- Where a student was enrolled in a double degree (including two concurrent unrelated degrees) and had completed units in both courses, two course level records were created: the student level record minus course-specific items completed for the second degree, and the student level record with course-specific items completed for the first degree replaced with those completed for the second degree.

3. Response and representativeness

3.1 Response rates

The response rate to the 2013 UES was 29.3 per cent, up from 24.2 per cent in 2012.[10] As shown in Table 5, there was substantial variation between institutions. Seven institutions exceeded the overall target response rate of 35 per cent and a further ten institutions achieved response rates in excess of 30 per cent. These response rates suggest that higher response rates for the UES are realistic for some if not all of the universities. Thirty institutions improved their response rates from 2012: an achievement considering that no telephone follow-up of non-responding students was undertaken in 2013.

Table 5. UES response rates, 2012 and 2013

University	2012	2013	University	2012	2013
MCD University of Divinity	45.1	50.5	University of the Sunshine Coast	34.5	29.2
The University of Adelaide	32.4	41.4	James Cook University	28.4	29.0
Charles Darwin University	33.4	40.5	University of Technology, Sydney	13.4	28.2
Monash University	26.9	39.7	The University of New South Wales[a]	17.5	27.0
The University of Western Australia	40.0	39.7	University of Western Sydney	29.0	26.6
Central Queensland University	34.7	36.0	Macquarie University	19.1	26.3
Flinders University	21.2	35.2	Curtin University	23.8	26.1
The University of Melbourne	23.9	34.5	The University of Notre Dame Australia[a]	18.6	26.0
The University of Newcastle	31.0	34.0	Swinburne University of Technology	19.2	25.5
University of Tasmania	29.0	33.0	University of Southern Queensland	21.6	25.2
La Trobe University	25.9	33.0	University of South Australia	29.6	25.2
The University of New England[a]	23.3	32.9	Southern Cross University	21.0	24.4
Bond University	19.8	32.8	University of Canberra	27.0	24.4
The University of Queensland[a]	27.0	32.5	Australian Catholic University	16.1	23.7
Charles Sturt University	27.9	32.3	University of Wollongong	20.7	23.5
Murdoch University	28.3	30.6	Griffith University	24.8	23.5
The University of Sydney	27.4	30.3	Federation University Australia	32.0	22.1
Queensland University of Technology	18.2	29.4	RMIT University	16.8	20.8
Edith Cowan University	32.2	29.3	Victoria University	24.2	17.9
The Australian National University	30.6	29.3			
Deakin University	15.9	29.2	**Total**	**24.2**	**29.3**

[a] Institution sent email invitations to their own students in 2013.

A number of lower-performing institutions indicated that they expected to have more difficulty in reaching the targeted response rate given the nature of their students, either because they offer a greater number of courses in the underperforming subject areas (e.g. management and commerce, and engineering), or because the students attending their institution are less likely to engage in survey activity in general. This variation in response rates underscores the need for each institution to have its own student engagement strategy that encourages their students to engage with the UES.

In 2013, four institutions declined to allow SRC to send email invitations to their students, citing privacy concerns based on legal advice (see Table 5). These institutions were responsible for sending their own email invitations and reminders, although copies of each were supplied to SRC so that content and timing could be confirmed prior to distribution. Two of these institutions achieved a response rate in excess of 30 per cent, which suggests that this approach, while not optimal from the

[10] The 2012 response rate figures include both online and telephone responses. When only online responses are considered, the 2012 response rate was 21.1 per cent.

perspective of consistency in survey administration, did not necessarily have a negative impact on student engagement. There was no indication that institutions sending their own emails modified the respondent list in any way and there is every reason to believe that all selected students received all relevant email correspondence.

3.2 Response characteristics

Response rates are arguably less important than the sample representativeness. To investigate this, characteristics of the sample of secured responses are presented alongside parameters of the in-scope population in Table 6 (subgroup) and Table 7 (subject area). To account for the course-level nature of the UES response file, the population file was similarly modified for the purpose of this investigation, in that double degree students were treated as two separate analytic units.

It is evident that many of the characteristics of the sample of secured responses match those of the in-scope population. Stage of studies, indigenous status, mode of attendance and type of attendance are similar in the sample and in-scope population. Language spoken at home and citizenship status are also surprisingly similar, given that students who speak a language other than English at home and international students are traditionally less likely to participate in similar surveys. As was the case in 2012, the largest potential source of non-response bias is in relation to gender, with male students substantially under-represented in the sample of secured responses.

Table 6. 2013 UES response characteristics and population parameters by subgroup

Group	Subgroup	UES sample		In-scope population	
		n	%	n	%
Stage of studies	Commencing	65,267	59.9	238,322	61.2
	Later year	43,673	40.1	150,872	38.8
Gender	Male	36,397	33.4	167,617	43.1
	Female	72,543	66.6	221,577	56.9
Indigenous	Aboriginal or Torres Strait Islander	1,138	1.1	4,526	1.2
	Not Aboriginal or Torres Strait Islander	106,137	98.9	378,226	98.8
Home language	English	84,306	78.2	293,672	76.4
	Other	23,528	21.8	90,622	23.6
Disability	Disability reported	5,340	4.9	16,049	4.1
	No disability reported	103,600	95.1	373,145	95.9
Study mode	Internal	95,099	87.3	342,834	88.1
	External/multi-modal	13,841	12.7	46,266	11.9
International	Domestic student	96,576	88.7	336,522	86.5
	International student	12,364	11.3	52,672	13.5
First in family	First in family	28,262	48.3	110,295	52.2
	Not first in family	30,258	51.7	101,137	47.8
Total		**108,940**	**100**	**389,194**	**100**

The sample also closely matches the in-scope population in terms of subject area (see Table 7). The largest differences between the sample and population were seen in relation to the subject areas of business and management (1.9 percentage points) and humanities (1.1 percentage points), which were the only two differences greater than one percentage point. The similarity of the sample and population in terms of subject area is somewhat surprising given the under-representation of males who, as a group, tend to enrol in different courses than females. This may be attributable to targeted engagement and follow-up of students in under-performing subject areas undertaken during data collection fieldwork. The largest subject areas in the sample are humanities (11.6 per cent), business

management (8.6 per cent), nursing (7.6 per cent), and natural and physical sciences (7.1 per cent). These four subject areas together constitute more than a third of the entire sample.

Table 7. 2013 UES response characteristics and population parameters by subject area

Broad field of education	Subject area	UES sample		In-scope population	
		n	%	n	%
Natural and Physical Sciences	Natural & Physical Sciences	7,701	7.1	25,712	6.6
	Mathematics	328	0.3	1,115	0.3
	Biological Sciences	1,892	1.7	5,777	1.5
	Medical Science & Technology	2,945	2.7	8,382	2.2
IT	Computing & Information Systems	3,203	2.9	11,987	3.1
Engineering and Related Technologies	Engineering – Other	3,558	3.3	13,736	3.5
	Engineering – Process & Resources	595	0.5	2,235	0.6
	Engineering – Mechanical	659	0.6	2,760	0.7
	Engineering – Civil	912	0.8	3,818	1.0
	Engineering – Electrical & Electronic	742	0.7	2,801	0.7
	Engineering – Aerospace	391	0.4	1,560	0.4
Architecture and Building	Architecture & Urban Environments	1,738	1.6	7,155	1.8
	Building & Construction	447	0.4	2,417	0.6
Agriculture and Environmental Studies	Agriculture & Forestry	652	0.6	1,731	0.4
	Environmental Studies	1,375	1.3	3,925	1.0
Health	Health Services & Support	5,693	5.2	20,159	5.2
	Public Health	1,139	1.0	4,264	1.1
	Medicine	2,179	2.0	6,300	1.6
	Nursing	8,279	7.6	27,506	7.1
	Pharmacy	903	0.8	2,680	0.7
	Dentistry	525	0.5	1,465	0.4
	Veterinary Science	494	0.5	1,417	0.4
	Physiotherapy	814	0.7	2,405	0.6
	Occupational Therapy	854	0.8	2,467	0.6
Education	Teacher Education – Other	2,276	2.1	7,400	1.9
	Teacher Education – Early Childhood	1,775	1.6	5,874	1.5
	Teacher Education – Primary & Secondary	4,803	4.4	17,655	4.5
Management and Commerce	Accounting	1,610	1.5	6,929	1.8
	Business Management	9,339	8.6	41,035	10.5
	Sales & Marketing	871	0.8	4,147	1.1
	Management & Commerce – Other	4,852	4.5	20,342	5.2
	Banking & Finance	796	0.7	4,099	1.1
Society and Culture	Political Science	539	0.5	1,685	0.4
	Humanities inc History & Geography	12,598	11.6	41,036	10.5
	Language & Literature	594	0.5	1,884	0.5
	Social Work	2,154	2.0	6,439	1.7
	Psychology	4,592	4.2	13,617	3.5
	Law	3,713	3.4	13,154	3.4
	Justice Studies & Policing	901	0.8	3,980	1.0
	Economics	1,084	1.0	5,100	1.3

	Sport & Recreation	209	0.2	971	0.2
Creative Arts	Art & Design	3,444	3.2	12,529	3.2
	Music & Performing Arts	1,138	1.0	4,024	1.0
	Communication, Media & Journalism	3,550	3.3	13,192	3.4
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	84	0.1	328	0.1
Total		**108,940**	**100**	**389,194**	**100**

3.3 Weighting

In the 2012 UES, weighting was undertaken to ensure that reported results were representative of the overall population. Given the serious under-representation of males in the 2013 UES sample, post-stratification weights by gender, subject area and stage of studies were computed separately for each institution.[11] This resulted in a total of 3,716 non-zero weighting strata.[12] Weights ranged in size from 1.0 to 42.5, with larger weights associated with lower stratum response rates.

Table 8. Comparison of raw and weighted percentage satisfied results by subgroup

Group	Subgroup	Quality of entire educational experience (item)		Quality of teaching (item)		Teaching Quality (focus area)	
		Raw	Weighted	Raw	Weighted	Raw	Weighted
Stage of studies	Commencing	82	81	81	81	81	81
	Later year	76	75	76	75	76	75
Gender	Male	77	77	76	76	77	77
	Female	80	80	80	80	80	80
Indigenous	Aboriginal or Torres Strait Islander	80	80	81	80	80	79
	Not Aboriginal or Torres Strait Islander	79	79	79	78	79	79
Home language	English	81	80	81	80	80	80
	Other	74	74	74	73	75	74
Disability	Disability reported	76	76	79	78	78	77
	No disability reported	79	79	79	78	79	79
Study mode	Internal	80	79	79	79	79	79
	External/multi-modal	77	77	77	77	78	77
International	Domestic student	80	80	80	79	80	79
	International student	72	71	72	71	74	73
First in family	First in family	82	81	82	82	82	81
	Not first in family	82	81	81	81	81	81
Total		**79**	**79**	**79**	**78**	**79**	**79**

Because the costs of weighting can include increased variance in estimates, difficulty in the calculation of standard errors and additional complexity when analysing the data, raw (unweighted) and weighted percentage satisfied results were compared to establish the utility of weighting the UES data. Two questionnaire items and one focus area were selected for this analysis: the quality of the

[11] For each institution, the post-stratification weights equal the in-scope population frequency of each stratum, defined on the basis of gender, subject area and stage of studies, divided by the frequency of the corresponding stratum in the sample of responses. When weights are applied, the weighted total of the sample approximates the total of the population.

[12] When calculating the weights, 10 cases in the response file were found to belong to strata that had no corresponding strata in the population file. Because weights could not be calculated for these strata, the cases were excluded from the analysis presented in Tables 8 and 9.

entire educational experience and quality of teaching items, and the teaching quality focus area. These were selected because they relate to what is arguably the core focus of the UES – the quality of teaching and the entire educational experience – and because the two items currently constitute the UES statistics reported on the *MyUniversity* website. The results are presented in Table 8 (subgroup) and Table 9 (subject area).[13]

It is evident from Tables 8 and 9 that post-stratification weighting as undertaken does not notably affect the results at a national level, which suggests that the under-representation of males in the sample of secured responses has not introduced any serious bias. This is likely related to the fact that the sample of secured responses reflects the in-scope population on most characteristics and subject area in particular. As a result, it was decided to analyse the data without applying weights. All results presented in this report, aside from those in Tables 8 and 9, are based on unweighted data.

Table 9. Comparison of raw and weighted percentage satisfied results by subject area

Broad field of education	Subject area	Quality of entire educational experience (item)		Quality of teaching (item)		Teaching Quality (focus area)	
		Raw	Weighted	Raw	Weighted	Raw	Weighted
Natural and Physical Sciences	Natural & Physical Sciences	83	82	84	83	83	82
	Mathematics	83	83	80	81	80	81
	Biological Sciences	84	84	85	85	84	84
	Medical Science & Technology	83	84	83	83	82	82
IT	Computing & Information Systems	75	75	72	72	74	74
Engineering and Related Technologies	Engineering – Other	73	73	69	69	71	71
	Engineering – Process & Resources	71	72	70	70	73	72
	Engineering – Mechanical	74	73	69	67	70	68
	Engineering – Civil	74	73	69	68	71	70
	Engineering – Electrical & Electronic	76	76	72	71	72	71
	Engineering – Aerospace	72	71	70	71	71	73
Architecture and Building	Architecture & Urban Environments	75	75	72	73	75	76
	Building & Construction	72	72	70	70	70	69
Agriculture and Environmental Studies	Agriculture & Forestry	80	80	78	77	77	76
	Environmental Studies	82	81	82	81	84	84
Health	Health Services & Support	81	81	82	82	81	81
	Public Health	84	84	83	83	83	83
	Medicine	78	78	74	74	75	75
	Nursing	75	75	75	75	75	75
	Pharmacy	80	79	80	79	80	79
	Dentistry	73	74	67	67	72	71
	Veterinary Science	84	85	84	85	84	85
	Physiotherapy	84	85	86	87	86	87
	Occupational Therapy	85	84	84	84	82	82
Education	Teacher Education – Other	80	79	80	80	78	78
	Teacher Education – Early Childhood	82	81	82	82	82	81

[13] This analysis was conducted using the *Weight Cases* procedure in SPSS, which gives cases different weights by simulated replication for statistical analysis. As such, the value of the weighting variable should indicate the number of observations represented by single cases in the data file.

	Teacher Education – Primary & Secondary	81	81	81	81	80	79
Management and Commerce	Accounting	75	75	74	74	76	76
	Business Management	77	76	75	75	76	75
	Sales & Marketing	76	75	75	74	75	75
	Management & Commerce – Other	75	75	73	73	75	75
	Banking & Finance	76	76	69	70	71	72
Society and Culture	Political Science	79	79	82	82	82	81
	Humanities inc History & Geography	81	81	84	84	84	84
	Language & Literature	84	83	86	86	88	88
	Social Work	80	80	81	82	82	83
	Psychology	85	85	87	87	86	85
	Law	80	80	82	82	81	81
	Justice Studies & Policing	76	74	75	73	75	72
	Economics	71	71	69	69	71	71
	Sport & Recreation	80	79	80	81	78	80
Creative Arts	Art & Design	79	78	79	79	80	80
	Music & Performing Arts	82	82	83	83	82	83
	Communication, Media & Journalism	82	82	83	83	82	82
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	84	82	82	83	83	84
Total		**79**	**79**	**79**	**78**	**79**	**79**

3.4 Precision of estimates

Because the 2013 UES data constitute a sample of the in-scope student population, it is reasonable to use statistical methods to analyse the sample of secured responses. To gauge the variability of the estimated results due to sampling variation, Tables 10 and 11 present percentage satisfied results for the quality of the entire educational experience and the quality of teaching items by subgroup and subject area, respectively, with 95 per cent confidence intervals around the point estimates. These confidence intervals have been calculated as 1.96 times the standard error. Because the student population is finite, and because the sample of secured UES responses constitutes more than a quarter of this population, standard errors have been adjusted by a finite population correction. This correction reduces the size of the confidence intervals surrounding the estimates. The calculation of these confidence intervals is explained in Appendix E.

As expected in this large national sample, the confidence intervals are generally narrow. At a national level, for example, the one-sided width of the 95 per cent confidence interval is around 0.2 percentage points for both items (see bottom row of Table 10), although the confidence intervals tend to be wider when the sample is subdivided. The subject area with the widest confidence interval was tourism, hospitality and personal services, with one-sided widths of 6.8 and 7.1 percentage points observed in relation to the entire experience and teaching quality items, respectively. This is not surprising, given that the point estimates are based on a small number of observations, even at the national level. This was the only subject area with a confidence interval wider than ±5 percentage points at the national level, although greater variability would likely be observed if this same exercise was performed on the data of a single institution. Regardless, this analysis has given evidence that the results presented in this report are likely to be close to the unknown population parameters.

Table 10. Percentage satisfied results by subgroup with 95 per cent confidence intervals

Group	Subgroup	Quality of entire educational experience[a]	Quality of teaching[a]
Stage of studies	Commencing	81.5 (81.2,81.7)	81.4 (81.1,81.7)
	Later year	75.7 (75.3,76.0)	75.5 (75.1,75.8)
Gender	Male	76.6 (76.2,77.0)	76.2 (75.8,76.6)
	Female	80.4 (80.2,80.7)	80.4 (80.2,80.7)
Indigenous	Aboriginal or Torres Strait Islander	80.4 (78.4,82.4)	81.2 (79.2,83.1)
	Not Aboriginal or Torres Strait Islander	79.2 (79.0,79.4)	79.0 (78.8,79.2)
Home language	English	80.7 (80.4,80.9)	80.5 (80.3,80.7)
	Other	73.9 (73.5,74.4)	73.9 (73.4,74.4)
Disability	Disability reported	76.4 (75.4,77.3)	79.0 (78.1,79.9)
	No disability reported	79.3 (79.1,79.5)	79.0 (78.8,79.2)
Study mode	Internal	79.5 (79.2,79.7)	79.3 (79.1,79.5)
	External/multi-modal	77.0 (76.4,77.6)	77.3 (76.7,77.9)
International	Domestic student	80.1 (79.9,80.3)	80.0 (79.7,80.2)
	International student	71.6 (70.9,72.3)	71.7 (71.0,72.4)
First in family	First in family	82.0 (81.7,82.4)	82.3 (81.9,82.7)
	Not first in family	81.8 (81.5,82.2)	81.2 (80.8,81.6)
Total		**79.1 (78.9,79.4)**	**79.0 (78.8,79.2)**

[a] Results are presented as estimate (lower confidence limit, upper confidence limit).

Table 11. Percentage satisfied results by subject area with 95 per cent confidence intervals

Broad field of education	Subject area	Quality of entire educational experience[a]	Quality of teaching[a]
Natural and Physical Sciences	Natural & Physical Sciences	82.7 (82.0,83.4)	84.0 (83.3,84.6)
	Mathematics	82.5 (79.0,85.9)	79.8 (76.2,83.5)
	Biological Sciences	84.3 (83.0,85.7)	85.1 (83.8,86.4)
	Medical Science & Technology	83.3 (82.2,84.4)	82.9 (81.8,84.0)
IT	Computing & Information Systems	74.7 (73.5,76.0)	72.2 (70.8,73.5)
Engineering and Related Technologies	Engineering – Other	73.0 (71.7,74.2)	68.7 (67.4,70.0)
	Engineering – Process & Resources	71.3 (68.2,74.4)	69.6 (66.4,72.8)
	Engineering – Mechanical	73.6 (70.6,76.5)	68.5 (65.4,71.6)
	Engineering – Civil	74.3 (71.8,76.7)	68.8 (66.1,71.4)
	Engineering – Electrical & Electronic	76.3 (73.7,79.0)	71.7 (68.9,74.4)
	Engineering – Aerospace	71.6 (67.7,75.5)	69.5 (65.5,73.4)
Architecture and Building	Architecture & Urban Environments	74.6 (72.8,76.4)	72.3 (70.5,74.2)
	Building & Construction	71.9 (68.1,75.7)	70.4 (66.6,74.2)
Agriculture and Environmental Studies	Agriculture & Forestry	80.0 (77.6,82.4)	77.8 (75.3,80.4)
	Environmental Studies	82.1 (80.4,83.7)	81.9 (80.2,83.5)
Health	Health Services & Support	80.8 (79.9,81.7)	82.0 (81.2,82.9)
	Public Health	84.4 (82.5,86.2)	83.2 (81.3,85.0)
	Medicine	77.7 (76.2,79.1)	73.7 (72.2,75.2)
	Nursing	75.3 (74.6,76.1)	75.4 (74.6,76.2)
	Pharmacy	79.6 (77.5,81.7)	79.7 (77.6,81.9)
	Dentistry	73.0 (69.9,76.0)	66.7 (63.5,70.0)
	Veterinary Science	84.1 (81.5,86.8)	83.5 (80.9,86.2)
	Physiotherapy	84.3 (82.3,86.3)	86.0 (84.0,87.9)
	Occupational Therapy	84.6 (82.6,86.6)	83.6 (81.6,85.6)
Education	Teacher Education – Other	79.5 (78.1,80.9)	80.2 (78.9,81.6)
	Teacher Education – Early Childhood	81.8 (80.3,83.3)	81.9 (80.4,83.4)
	Teacher Education – Primary & Secondary	81.3 (80.4,82.3)	80.7 (79.8,81.7)
Management and Commerce	Accounting	75.4 (73.6,77.3)	73.7 (71.8,75.6)
	Business Management	76.7 (75.9,77.5)	75.3 (74.5,76.1)
	Sales & Marketing	76.3 (73.8,78.8)	75.3 (72.8,77.9)
	Management & Commerce – Other	75.2 (74.1,76.2)	73.1 (72.0,74.2)
	Banking & Finance	76.0 (73.4,78.7)	69.0 (66.1,71.9)
Society and Culture	Political Science	79.1 (76.3,82.0)	81.8 (79.1,84.5)
	Humanities inc History & Geography	81.1 (80.6,81.7)	84.0 (83.5,84.6)

	Language & Literature	83.9 (81.5,86.4)	85.7 (83.4,88.0)
	Social Work	80.0 (78.6,81.4)	81.2 (79.8,82.5)
	Psychology	84.7 (83.8,85.5)	86.8 (86.0,87.6)
	Law	79.6 (78.5,80.7)	81.7 (80.7,82.8)
	Justice Studies & Policing	76.0 (73.5,78.5)	75.3 (72.9,77.8)
	Economics	71.2 (68.8,73.6)	69.3 (66.8,71.7)
	Sport & Recreation	79.8 (75.0,84.6)	80.3 (75.5,85.1)
Creative Arts	Art & Design	79.0 (77.8,80.1)	78.9 (77.7,80.0)
	Music & Performing Arts	81.6 (79.7,83.5)	82.5 (80.6,84.4)
	Communication, Media & Journalism	81.9 (80.8,83.0)	82.5 (81.5,83.6)
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	84.3 (77.6,91.1)	82.1 (75.1,89.2)
Total		**79.1 (78.9,79.4)**	**79.0 (78.8,79.2)**

[a] Results are presented as estimate (lower confidence limit, upper confidence limit).

4. Key results from the 2013 UES

4.1 The university experience of specific groups

Percentage satisfied results for all five focus areas are presented in Table 12, stratified by a number of important demographic and contextual characteristics, with overall results presented in the bottom row. It is critical to note that the results presented in this section are based on a series of separate analyses and thus do not reflect any interactions between any of the characteristics. This approach was adopted in the interest of parsimony of reporting and explanation.

Considering first the overall results, there is much variation in percentage satisfied results. These ranged from 83 per cent in relation to the learning resources focus area, to 53 per cent for the student support focus area. Encouragingly, a relatively large proportion of students indicated satisfaction with their skills development and the quality of teaching provided by their institution (each with 79 per cent). In terms of their level of engagement with their institution, however, 57 per cent of students expressed satisfaction. Because the five UES focus areas represent different facets of the student experience, it is inadvisable to make comparisons across them. It is interesting, however, to note the aspects with which a smaller proportion of students expressed satisfaction.

Table 12. Percentage satisfied results by subgroup

Group	Subgroup	Skills Development	Learner Engagement	Teaching Quality	Student Support	Learning Resources
Stage of studies	Commencing	77	56	81	57	86
	Later year	82	58	76	48	79
Gender	Male	76	57	77	52	82
	Female	81	57	80	54	84
Age group	Under 25	79	60	79	52	84
	25 to 29	77	49	76	54	80
	30 to 39	78	42	80	58	82
	40 and over	78	40	82	61	82
Indigenous	Aboriginal or Torres Strait Islander	81	55	80	61	86
	Not Aboriginal or Torres Strait Islander	79	57	79	53	83
Home language	English	79	58	80	53	84
	Other	78	55	75	55	82
Disability	Disability reported	76	54	78	60	80
	No disability reported	79	57	79	53	84
Study mode	Internal	79	60	79	53	84
	External/multi-modal	78	36	78	53	81
International	Domestic student	79	58	80	53	84
	International student	77	51	74	58	81
First in family	First in family	78	55	82	58	87
	Not first in family	76	59	81	56	87
Previous university experience	Current university	79	57	79	52	82
	Another university	77	51	79	55	81
	New to higher education	79	59	79	53	84
Total		**79**	**57**	**79**	**53**	**83**

Later-year students were more likely to be satisfied with their skill development compared with those who had recently commenced their studies, and were marginally more likely to be satisfied with their level of engagement. They were, on the other hand, less likely than commencing students to indicate their satisfaction with the teaching quality, student support and learning resources provided by their

institution. Student support is of particular concern, with fewer than half of all later-year students and 57 per cent of commencing students expressing satisfaction with this facet.

Considering male and female students, most of the differences in percentage satisfied results were marginal, with female students generally more likely to be satisfied with their educational experience than male students. A difference between males and females of five percentage points was observed in relation to the skills development focus area; however this result may be influenced by differences in the courses undertaken by males and females.

There is a clear negative association between age and learner engagement, with young students (aged under 25) much more likely to be satisfied with their level of engagement than students in the three older age groups, and students aged 30 and over in particular. This result is consistent with the fact that older students are more likely to study either externally or by mixed-mode delivery, which are study modes characterised by relatively low levels of student engagement as compared with internal delivery. Older students are also presumably more likely to be balancing their studies with their work and family lives, which would further limit their learner engagement opportunities (as measured by the UES). Interestingly, though, older students were more likely to express satisfaction with the student support provided by their institution.

Aboriginal and Torres Strait Islander students, while constituting only 1.1 per cent of the sample (see Table 6), were much more likely to be satisfied with the student support provided by their institution than their non-Indigenous classmates. Differences between Indigenous and non-Indigenous students in relation to the other four UES focus areas were smaller in magnitude and, given the width of the confidence intervals associated with the percentage agreement results for Aboriginal and Torres Strait Islander students (see Table 10), may not be statistically significant.

With the exception of student support, students who spoke English as their main language at home were slightly more likely to be satisfied with most aspects of the educational experience. Aside from teaching quality, which saw five percentage points separating the two language groups, differences tended to be fairly small. A similar pattern is observed in relation to domestic students, who were more likely than international students to be satisfied with every aspect of their educational experience, with the exception of student support. These differences were fairly small in relation to skills development and learning resources, but were larger in relation to learner engagement and teaching quality.

Students who reported having a disability were much more likely to be satisfied with student support, with seven percentage points separating them from students who did not report any disability. The opposite is observed in relation to the four other focus areas; however in all cases these differences were relatively minor.

Few noteworthy differences were observed based on whether the student was the first in their family to attend university, with the largest difference being that students who were the first in their family to attend university were less likely to be satisfied with their level of engagement. Considering whether students had previous university experience, it is interesting to note that students who had previously been enrolled at another university were less likely to be satisfied with their level of engagement, especially in relation to students new to higher education. There were no other notable differences on the basis of this characteristic.

Looking now at subject area (see Table 13), there is considerable variation in percentage satisfied results both across and within subject areas. The narrowest range of results across subject areas is seen in relation to student support, with 16 percentage points separating the subject areas (language and literature with the highest results, and architecture and urban environments, and economics with the lowest), followed the by teaching quality and skills development (both 19 percentage points), and learning resources (20 percentage points) focus areas. The widest range is observed for learner engagement, with 33 percentage points separating the two subject areas with the highest and lowest results (physiotherapy, and justice studies and policing, respectively). There are also some notable cases of variation within broad fields of education. Within the society and culture broad field of education, for example, economics students were less likely to express satisfaction in relation to skills development, teaching quality and student support than students in other subject areas.[14] Another example can be seen for teacher education – early childhood students in relation to learner engagement. This result underscores the fact that broad disciplinary aggregations hide much of the detail that is relevant to schools, faculties and academic departments.

While confidence intervals are not shown in Table 13, it is important to interpret the results with respect to the remarks made in Section 3.4 concerning the precision of estimates in the UES. It is possible that some of the differences in this table, especially those seen in relation to subject areas containing small numbers of observations, may not be statistically significant.

Table 13. Percentage satisfied results by subject area

Broad field of education	Subject area	Skills Development	Learner Engagement	Teaching Quality	Student Support	Learning Resources
Natural and Physical Sciences	Natural & Physical Sciences	77	58	83	54	87
	Mathematics	73	53	80	58	87
	Biological Sciences	82	62	84	57	87
	Medical Science & Technology	80	63	82	57	87
IT	Computing & Information Systems	72	58	74	54	81
Engineering and Related Technologies	Engineering – Other	76	63	71	51	81
	Engineering – Process & Resources	79	69	73	53	82
	Engineering – Mechanical	76	61	70	49	78
	Engineering – Civil	78	66	71	48	81
	Engineering – Electrical & Electronic	73	63	72	51	80
	Engineering – Aerospace	77	63	71	50	82
Architecture and Building	Architecture & Urban Environments	77	61	75	45	71
	Building & Construction	72	53	70	47	81
Agriculture and Environmental Studies	Agriculture & Forestry	73	56	77	54	85
	Environmental Studies	79	61	84	58	84
Health	Health Services & Support	80	61	81	54	86
	Public Health	84	61	83	55	87

[14] A plausible explanation for this may be related to the fact that, while economics is rightly classified as a social science, economics departments are often located in business schools. An examination of Table 13 suggests that the university experience of economics students is closer to that of management and commerce students than that of society and culture students.

	Medicine	85	73	75	51	78
	Nursing	84	55	75	56	84
	Pharmacy	84	65	80	55	85
	Dentistry	82	62	72	50	76
	Veterinary Science	85	71	84	50	81
	Physiotherapy	87	77	86	57	90
	Occupational Therapy	85	70	82	54	90
Education	Teacher Education – Other	81	54	78	54	85
	Teacher Education – Early Childhood	84	49	82	56	84
	Teacher Education – Primary & Secondary	82	60	80	52	84
Management and Commerce	Accounting	74	45	76	55	82
	Business Management	77	54	76	54	83
	Sales & Marketing	78	57	75	49	81
	Management & Commerce – Other	75	52	75	55	82
	Banking & Finance	71	45	71	54	83
Society and Culture	Political Science	76	52	82	51	81
	Humanities inc History & Geography	78	52	84	53	83
	Language & Literature	77	51	88	61	85
	Social Work	83	48	82	55	80
	Psychology	82	55	86	58	86
	Law	83	54	81	52	84
	Justice Studies & Policing	74	44	75	50	81
	Economics	68	50	71	45	81
	Sport & Recreation	85	60	78	58	89
Creative Arts	Art & Design	78	60	80	50	79
	Music & Performing Arts	78	71	82	52	78
	Communication, Media & Journalism	81	61	82	50	85
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	81	58	83	55	91
Total		**79**	**57**	**79**	**53**	**83**

4.2 University experience perceptions over time

Table 14 compares results from the 2013 UES with those from 2012. When reading the results in this table there are several critical issues to bear in mind. First, while confidence intervals are not shown in Table 14, it is important again to consider the precision of the estimates, especially in relation to subject areas based on small numbers of observations. Some apparent differences may not be statistically significant. Second, as noted previously, the 2012 and 2013 collections differed in a number of key respects, including the use of telephone interviewing in 2012, the new course-level focus of the 2013 UES, and differences between the two collections in terms of what constitutes a completed survey and a later-year student. These methodological differences may explain at least some of the differences in results between the 2012 and 2013 collections. Finally, while the 2012 UES report presented average scores for each focus area, the 2013 report presents percentage satisfied results. To allow for valid comparisons to be made between the two years, percentage satisfied results were generated from the 2012 data file using the approach detailed in Appendix B. As a result, these figures will not reconcile with those presented in the 2012 UES National Report.

At the national level, the largest difference in percentage satisfied results between years was observed in relation to the skills development focus area, with 3 percentage points separating 2012 (82 per cent) and 2013 (79 per cent). In relation to the teaching quality focus area, 2 percentage points separated 2012 (81 per cent) and 2013 (79 per cent). The difference between years in relation to the learner engagement and learning resources focus areas was only 1 percentage point, with the latter being the only focus area with a higher result in 2013 than 2012. No difference was observed in relation to the student support focus area. Given the large number of observations at the national level, these differences are likely to be statistically significant; however, as noted previously, they may be due to methodological differences between the 2012 and 2013 collections. In any case, the differences at the national level are relatively small.

Table 14. Percentage satisfied results by subject area, 2012 and 2013

Broad field of education	Subject area	2012[ab]					2013[a]				
		SD	LE	TQ	SS	LR	SD	LE	TQ	SS	LR
Natural and Physical Sciences	Natural & Physical Sciences	82	60	85	56	86	77	58	83	54	87
	Mathematics	78	55	81	59	86	73	53	80	58	87
	Biological Sciences	84	60	87	58	88	82	62	84	57	87
	Medical Science & Technology	84	62	84	56	87	80	63	82	57	87
IT	Computing & Information Systems	76	56	76	54	83	72	58	74	54	81
Engineering and Related Technologies	Engineering – Other	79	61	74	48	82	76	63	71	51	81
	Engineering – Process & Resources	80	64	76	48	80	79	69	73	53	82
	Engineering – Mechanical	80	60	72	43	77	76	61	70	49	78
	Engineering – Civil	82	66	74	45	81	78	66	71	48	81
	Engineering – Electrical & Electronic	79	59	76	49	82	73	63	72	51	80
	Engineering – Aerospace	83	63	75	46	79	77	63	71	50	82
Architecture and Building	Architecture & Urban Environments	80	63	77	47	69	77	61	75	45	71
	Building & Construction	72	41	65	41	75	72	53	70	47	81
Agriculture and Environmental Studies	Agriculture & Forestry	78	51	78	49	81	73	56	77	54	85
	Environmental Studies	83	58	83	56	83	79	61	84	58	84
Health	Health Services & Support	83	61	83	53	84	80	61	81	54	86
	Public Health	84	57	84	58	90	84	61	83	55	87
	Medicine	85	73	77	44	77	85	73	75	51	78
	Nursing	84	57	76	56	81	84	55	75	56	84
	Pharmacy	84	63	79	54	83	84	65	80	55	85
	Dentistry	86	59	79	46	69	82	62	72	50	76
	Veterinary Science	87	70	86	46	80	85	71	84	50	81
	Physiotherapy	88	75	90	56	82	87	77	86	57	90
	Occupational Therapy	88	71	87	51	85	85	70	82	54	90
Education	Teacher Education – Other	84	56	81	57	85	81	54	78	54	85
	Teacher Education – Early Childhood	86	54	83	54	81	84	49	82	56	84
	Teacher Education – Primary & Secondary	85	61	80	51	82	82	60	80	52	84
Management and Commerce	Accounting	76	47	77	58	82	74	45	76	55	82
	Business Management	80	53	77	52	82	77	54	76	54	83
	Sales & Marketing	83	59	78	50	80	78	57	75	49	81
	Management & Commerce – Other	80	54	78	55	84	75	52	75	55	82

	Banking & Finance	77	51	73	53	82	71	45	71	54	83
Society and Culture	Political Science	84	61	85	54	81	76	52	82	51	81
	Humanities inc History & Geography	83	53	86	55	83	78	52	84	53	83
	Language & Literature	79	57	86	55	82	77	51	88	61	85
	Social Work	85	52	82	55	79	83	48	82	55	80
	Psychology	86	54	87	56	87	82	55	86	58	86
	Law	82	50	80	52	81	83	54	81	52	84
	Justice Studies & Policing	83	49	82	55	88	74	44	75	50	81
	Economics	78	55	74	50	80	68	50	71	45	81
	Sport & Recreation	83	58	85	49	83	85	60	78	58	89
Creative Arts	Art & Design	80	59	80	47	77	78	60	80	50	79
	Music & Performing Arts	81	71	85	48	77	78	71	82	52	78
	Communication, Media & Journalism	85	62	85	53	85	81	61	82	50	85
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	87	57	87	62	86	81	58	83	55	91
Total		**82**	**58**	**81**	**53**	**82**	**79**	**57**	**79**	**53**	**83**

[a] SD = Skills Development, LE = Learner Engagement, TQ = Teaching Quality, SS = Student Support, LR = Learning Resources.

[b] Due to changes in methodology, care should be taken when comparing 2012 and 2013 results (see Section 4.2). Moreover, the 2012 results presented in Table 14 will not reconcile with those in the 2012 UES National Report due to the change in reporting metric from average scores to percentage satisfied (see Section 1.3).

4.3 Results on individual questionnaire items

Table 15 presents percentage satisfied results for the 47 individual survey items underpinning the five UES focus areas, stratified by stage of studies. Results from the 2012 UES are presented to facilitate comparisons over time. When reading the results in Table 15, the previously discussed caveats on comparing the 2012 and 2013 UES collections should be borne in mind (see Section 4.2).

In relation to these individual items, percentage satisfied relates to the percentage of responses in the top two response categories. It is interesting to observe that many of the highest percentage satisfied results relate to the items constituting the learning resources focus area, with the quality of library resources and facilities especially highly rated (87 per cent), along with the quality of teaching spaces and online learning materials (both with 84 per cent). It is also reassuring to see a large percentage of responses expressing satisfaction with the quality of teaching and the entire educational experience (both with 79 per cent), which were also amongst the highest-rated items. The lowest results were observed in relation to items in the student support focus area, with few students indicating that they used university services to support their studies (16 per cent) or that they received appropriate English language skill support (18 per cent). It could also be concerning that only 26 per cent expressed the view that their institution offered support relevant to their circumstances.

As expected, some of the largest differences in percentage satisfied results between commencing and later-year students were observed in relation to the skills development focus area, specifically written communication skills, spoken communication skills and the ability to solve complex problems (each with 11 percentage points). Commencing students, on the other hand, were much more likely than later-year students to indicate satisfaction with the support they received to settle into study, with 12 percentage points separating them. Given that this experience would still be fresh in the minds of commencing students, this is hardly a surprising result. In general, commencing students were more

likely to indicate satisfaction with the items relating to teaching quality, student support and learning resources, whereas later-year students were more likely to indicate satisfaction with the items relating to skills development and learner engagement.

Table 15. Percentage satisfied results for UEQ items by stage of studies, 2012 and 2013

Focus area	Item	2012[ab]			2013[a]		
		C	LY	Total	C	LY	Total
Skills Development	Developed critical and analytical thinking	66	75	71	66	73	69
	Developed ability to solve complex problems	58	69	64	53	64	58
	Developed ability to work effectively with others	61	66	64	57	64	60
	Developed confidence to learn independently	72	77	75	67	75	70
	Developed written communication skills	60	70	65	57	68	61
	Developed spoken communication skills	52	62	57	47	58	52
	Developed knowledge of field studying	81	83	82	76	78	77
	Developed work-related knowledge and skills	64	63	63	60	62	60
Learner Engagement	Felt prepared for your study	54	55	55	56	61	58
	Had a sense of belonging to your university	49	42	46	52	47	50
	Participated in discussions online or face-to-face	56	59	57	54	59	56
	Worked with other students as part of your study	68	70	69	60	64	62
	Interacted with students outside study requirements	54	54	54	45	47	46
	Interacted with students who are very different from you	54	51	53	54	52	53
	Been given opportunities to interact with local students	51	49	50	56	54	55
Teaching Quality	Study well structured and focused	71	65	68	68	61	65
	Study relevant to education as a whole	74	72	73	71	68	70
	Teachers engaged you actively in learning	66	62	64	62	59	61
	Teachers demonstrated concern for student learning	61	58	60	59	56	57
	Teachers provided clear explanations on coursework and assessment	68	64	66	63	61	62
	Teachers stimulated you intellectually	68	65	67	68	65	67
	Teachers commented on your work in ways that help you learn	52	50	51	49	49	49
	Teachers seemed helpful and approachable	71	67	69	70	68	69
	Teachers set assessment tasks that challenge you to learn	77	71	74	77	72	75
	Quality of teaching	86	80	83	81	76	79
	Quality of entire educational experience	83	77	80	82	76	79
Student Support	Experienced efficient enrolment and admissions processes	62	56	59	68	63	66
	Induction/orientation activities relevant and helpful	46	37	42	51	42	48
	Received support from university to settle into study	43	29	36	52	40	47
	Administrative staff or systems: available	67	61	64	63	56	61
	Administrative staff or systems: helpful	62	55	59	61	53	58
	Careers advisors: available	49	46	47	47	42	45
	Careers advisors: helpful	48	44	45	47	41	44
	Academic or learning advisors: available	60	56	58	61	57	59
	Academic or learning advisors: helpful	62	58	60	64	59	62
	Support services: available	56	52	54	54	49	52
	Support services: helpful	55	52	54	54	51	53
	Used university services to support study	24	20	22	16	15	16
	Offered support relevant to circumstances	24	20	22	28	24	26
	Received appropriate English language skill support	39	34	36	20	16	18
Learning Resources	Quality of teaching spaces	89	82	85	86	80	84
	Quality of student spaces and common areas	79	71	75	78	70	75
	Quality of online learning materials	85	80	83	86	82	84
	Quality of computing/IT resources	81	75	78	83	77	80
	Quality of assigned books, notes and resources	81	76	78	81	76	79
	Quality of laboratory or studio equipment	85	77	81	85	77	82
	Quality of library resources and facilities	89	85	87	88	84	87

[a] C = Commencing, LY = Later year.

[b] Due to changes in methodology, care should be taken when comparing 2012 and 2013 results (see Section 4.2).

Table 15 also demonstrates the extent of the variation in percentage satisfied results between items in the same focus area. The smallest variation is observed in relation to the learning resources focus area, with 12 percentage points separating the lowest and highest results. Similarly, a range of 16 percentage points was observed in relation to the learner engagement focus area. Conversely, 50 percentage points separated the lowest and highest percentage satisfied results in the student support

focus area. In general, however, there was more variation in percentage satisfied results between the items in different focus areas than in the same focus area.

Several large differences were observed between years. In particular, students who completed the 2012 UES were more likely than those who completed the 2013 survey to indicate that they received appropriate English language skill support (18 percentage points), whereas the opposite was true in relation to support provided by the university to settle into study (11 percentage points). Given the previously discussed methodological differences between the 2012 and 2013 collections, these results should be interpreted with caution.

An analysis of the psychometric properties of the UEQ, which uses a Rasch measurement model to assess how well the items in each focus area work together to measure a common trait, is presented in Appendix F. Response category percentages are presented in Appendix G.

4.4 The university experience of students from different institutions

Percentage satisfied results on the entire educational experience and teaching quality items are given in Figures 1 and 2, respectively, for students from different higher education institutions. While this analysis is useful in terms of measuring differences in quality between institutions in the Australian higher education sector, it is important to note that this analysis does not account for differences in course offerings between institutions and the composition of the student bodies. To avoid creating a simplistic "league table" of higher education institutions, university names have been replaced with randomly-assigned numerical identifiers in Figures 1 and 2.[15]

Because of the relatively small number of students at the institutional level, 95 per cent confidence intervals have been included in these figures. A wider confidence interval implies that there is more variability in results. If the confidence intervals for two institutions overlap, this suggests that there may be no statistically significant difference between the results. If the confidence intervals do not overlap, then any difference between results is likely to be statistically significant.

When institutional percentage satisfied results are ordered for the two selected items, there is a fairly even increase from the bottom of the distribution to near the top, with a few institutions at the top of the distribution notably higher than the majority of institutions. Looking at Figure 1, which reports percentage satisfied results on the quality of the entire educational experience item, the majority of institutions in the lower third of the distribution are significantly different to those in the higher third of the distribution, when confidence intervals are considered. While there does not appear to be many significant differences between institutions in the middle of the distribution, there are institutions at both ends of the distribution that are significantly different to those in the middle.

A similar picture emerges from Figure 2, although the increase in results across the distribution is steeper. There is a strong correlation in the ranking of institutions in both figures; however, given that the items on which these two figures are based constitute part of the teaching quality focus area, this is not an unexpected result. The steeper increase in results means there is a larger difference between the highest and lowest percentage satisfied results in Figure 2 than Figure 1 (24 percentage points and 16 percentage points, respectively), although this result should be interpreted in the context of the wide confidence interval on the highest-ranked institution in Figure 2. The steeper slope observed in Figure

[15] For example, "U01" represents the same institution in Figures 1 and 2.

2 also means that there are, in general, greater statistically significant differences for the quality of teaching item than for the quality of entire educational experience item.

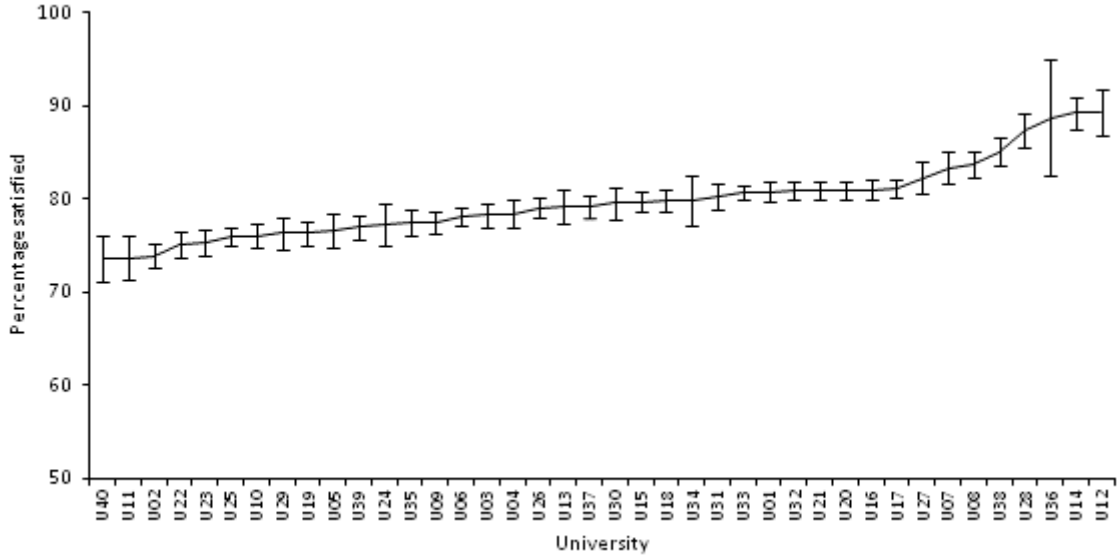


Figure 1. Percentage satisfied results on the quality of entire educational experience

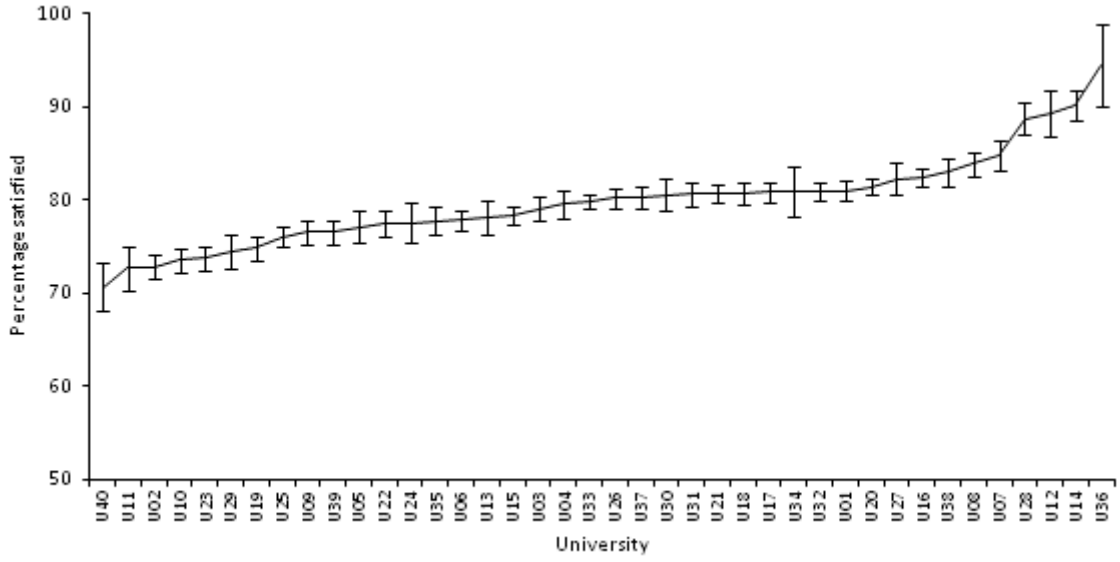


Figure 2. Percentage satisfied results on the quality of teaching

4.5 International comparisons

A consideration when developing the UES was to ensure the ability to use the data for benchmarking against similar student satisfaction surveys conducted in other national contexts. The "overall satisfaction" question on the National Survey of Student Engagement (NSSE), for example, is highly similar to the quality of the entire educational experience item on the UES.[16] NSSE collects

[16] "How would you evaluate your entire educational experience at this institution?"

information on student participation in programs and activities that institutions provide for their personal development. It is administered widely in the USA and Canada, with 371,284 students from 621 colleges and universities completing the 2013 NSSE.[17]

Figure 3 presents the percentage of surveyed students who rated their entire educational experience positively. Data from the 2011 UES should be treated with caution, as this was a pilot administration in which only 24 universities participated. The caveats noted in Section 4.3 concerning changes to the UES collection methodology should also be considered in relation to this figure. It is also critical to note that, while the 2012 and 2013 UES collections included every Australian university, NSSE is only administered to a subset of universities and colleges in the USA and Canada, which number more than 2,700 in total. If the institutions that participate in NSSE differ from those that do not, the results will not necessarily reflect an unbiased estimate of student satisfaction at the overall sector level. If, for example, the NSSE is administered to students of "better" institutions, the results will be biased upward. Therefore, as more years of UES data are gathered using a consistent data collection methodology, comparing movements over time within sectors (Australia and USA/Canada) could be more valid than comparing the two sectors directly.

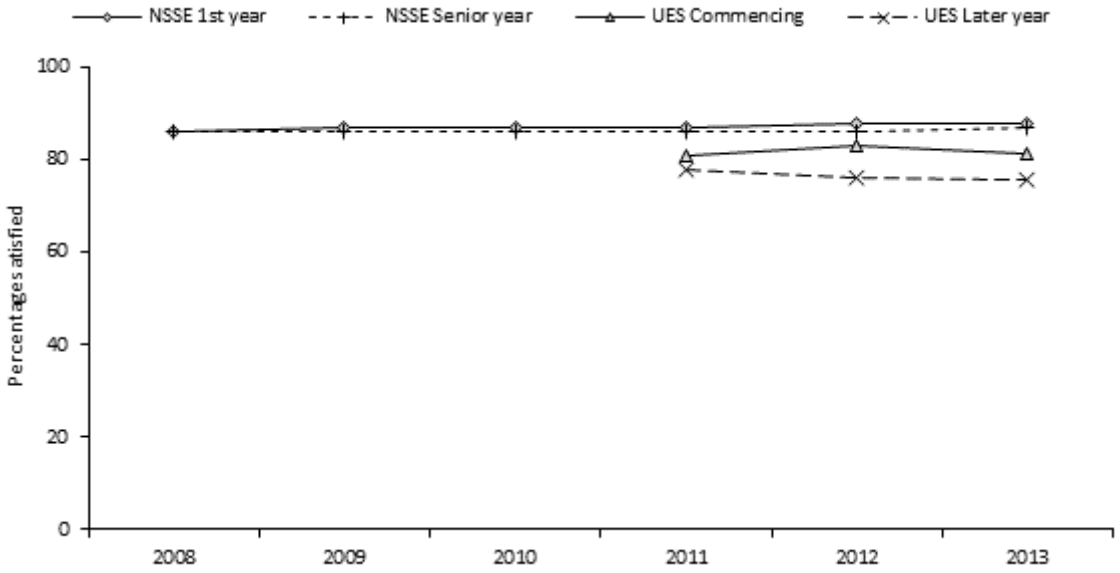


Figure 3. Entire educational experience rated positively, UES and NSSE, 2008 to 2013

Bearing these caveats in mind, Figure 3 shows that respondents to the NSSE are more likely to be satisfied with their educational experience than respondents to the UES, especially amongst later-year students. It is also interesting to note that the percentage satisfied results of NSSE first- and senior-year students are much closer together than those of commencing and later-year students from the UES. The reason for this is not clear, but could relate to non-random participation in NSSE, in terms of both students and institutions, fundamental differences between the Australian and North American higher education sectors, or other methodological differences between the two surveys.

[17] Indiana University. (2014). *About NSSE*. Retrieved 15 January, 2014, from http://nsse.iub.edu/html/about.cfm

4.6 Early departure

In addition to the items asking students to rate their level of satisfaction with different aspects of their educational experience, students were also asked to indicate whether they had seriously considered leaving their university during 2013. The results of this question are presented by student subgroup in Table 16. Overall, 17 per cent indicated that they had considered leaving.

As might be expected, commencing students were more likely than later-year students to consider leaving their university; however the difference between these two groups was only four percentage points. This unusually small difference may be due to the fact that many commencing students who considered leaving university had already done so by the time the UES was conducted in August, well into semester 2, and would not appear in the data.

Young students aged under 25 were less likely than their older classmates to have considered leaving their university. Aboriginal and Torres Strait Islander students were notably more likely to consider early departure than non-Indigenous students. This is interesting in light of the fact that Indigenous students were also much more likely to be satisfied with the support provided by their university (see Table 12). It could be that the support provided by their institution allowed them to remain enrolled in their courses, since the students who completed the UES were those who did not ultimately leave their university, at least at the time the UES was administered.

Students who spoke English as their main language at home were more likely to consider leaving their university than those who spoke a language other than English at home. A similar pattern is observed in relation to domestic and international students.

Table 16. Percentage of students considering early departure by subgroup

Group	Subgroup	Per cent considering departure	Group	Subgroup	Per cent considering departure
Stage of studies	Commencing	19	Disability	Disability reported	23
	Later year	15		No disability reported	17
Gender	Male	16	Study mode	Internal	17
	Female	17		External/multi-modal	19
Age group	Under 25	16	International	Domestic student	17
	25 to 29	20		International student	14
	30 to 39	20	First in family	First in family	20
	40 and over	21		Not first in family	17
Indigenous	ATSI	24	Previous university experience	Current university	18
	Not ATSI	17		Another university	17
Home language	English	18		New to higher education	17
	Other	15	**Total**		**17**

Students who reported having a disability were more likely to have considered leaving their university than students who did not report having a disability. As was the case with Aboriginal and Torres Strait Islander students, students with a disability were more likely to express satisfaction with the level of support provided by their university (see Table 12). Students completing their studies externally or by mixed-mode delivery were more likely than internal students to consider early departure, but, at only two percentage points, the difference is not large. Students who were the first in their family to attend university were more likely than their peers to have considered leaving their university. This result is logical, considering that these students would generally know less about what to expect at university

than those with a family history of higher education. No substantial differences in departure intentions were observed in relation to gender or previous university experience.

The percentage of students considering leaving their university in 2013 is plotted against average grades in Figure 4. The expected relationship is observed, with students achieving lower grades much more likely to consider early departure than students achieving high grades. This is most apparent for students achieving a grade of less than 50 per cent, of whom more than 40 per cent considered early departure in 2013.

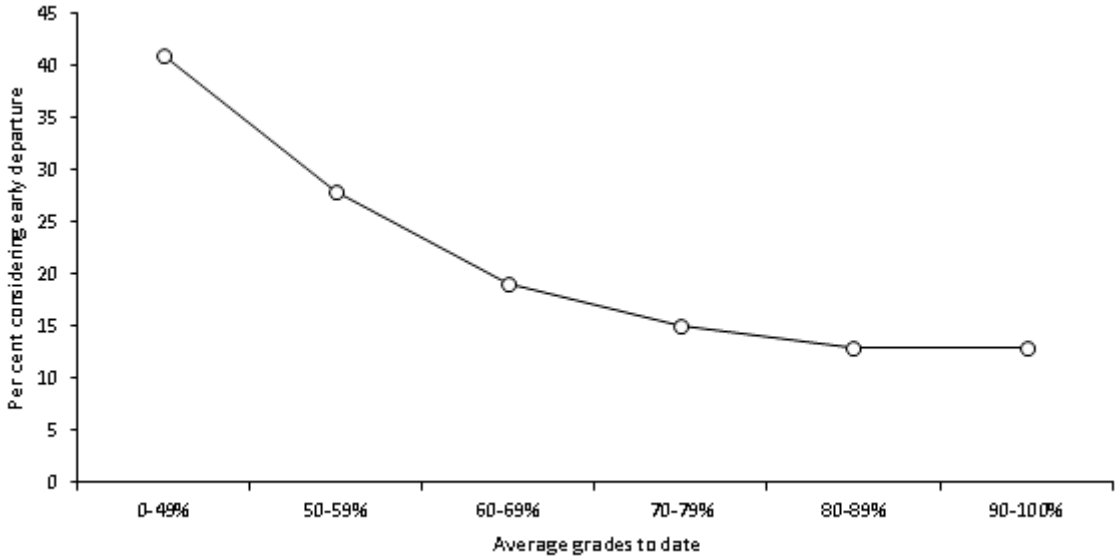


Figure 4. Percentage of students considering early departure by average grades to date

Students who expressed a serious consideration of leaving their university in 2013 were then asked to indicate, from a list of 30 possible reasons, why they considered doing so. These are summarised in Table 17. Students could select as many reasons as applied, so the percentages do not total 100. It is evident from the table that some of the most common reasons relate to situational factors, such as health or stress (31 per cent), study/life balance (29 per cent), difficulties relating to workload and finances (both 28 per cent), the need to do paid work and unspecified personal reasons (both 23 per cent). The fact that these reasons were indicated by such a large percentage of students underscores the importance of student support in terms of allowing students to continue with their studies.

Encouragingly for institutions, the most common (arguably) institutional factor indicated by students was that their expectations had not been met (24 per cent). Other institutional factors were indicated much less frequently (e.g. academic support, administrative support, institutional reputation). Several dispositional factors were also relatively common, including boredom or lack of interest (21 per cent), career prospects, a change in direction and a need to take a break (each 20 per cent).

Table 17. Selected reasons for considering early departure

Departure reason	Per cent of those considering departure	Departure reason	Per cent of those considering departure
Health or stress	31	Other	13
Study / life balance	29	Gap year / deferral	12
Workload difficulties	29	Commuting difficulties	11
Financial difficulties	28	Academic exchange	10
Expectations not met	24	Fee difficulties	9
Need to do paid work	24	Other opportunities	9
Personal reasons	24	Social reasons	8
Boredom/lack of interest	21	Administrative support	7
Career prospects	20	Travel or tourism	7
Change of direction	20	Institution reputation	7
Need a break	20	Standards too high	6
Family responsibilities	18	Graduating	5
Academic support	15	Moving residence	5
Paid work responsibilities	15	Government assistance	4
Quality concerns	15	Received other offer	3

Appendix A: University Experience Questionnaire (UEQ)

Table 18. Skill Development items

Stem	Item	Response scale
To what extent has your course developed your:	critical thinking skills?	Not at all / Very little / Some / Quite a bit / Very much
	ability to solve complex problems?	
	ability to work with others?	
	confidence to learn independently?	
	written communication skills?	
	spoken communication skills?	
	knowledge of the field(s) you are studying?	
	development of work-related knowledge and skills?	

Table 19. Learner Engagement items

Stem	Item	Response scale
At university during 2013, to what extent have you:	felt prepared for your study?	Not at all / Very little / Some / Quite a bit / Very much / Not applicable
	had a sense of belonging to your university?	
In 2013, how frequently have you:	participated in discussions online or face-to-face?	Never / Sometimes / Often / Very often
	worked with other students as part of your study?	
	interacted with students outside study requirements?	
	interacted with students who are very different from you?	
At university during 2013, to what extent have you:	been given opportunities to interact with local students?	Not at all / Very little / Some / Quite a bit / Very much / Not applicable

Table 20. Teaching Quality items

Stem	Item	Response scale
Thinking about your university course:	overall how would you rate the quality of your entire educational experience this year?	Poor / Fair / Good / Excellent
	how would you rate the quality of the teaching you have experienced?	
During 2013, to what extent have the lecturers, tutors and demonstrators:	engaged you actively in learning?	Not at all / Very little / Some / Quite a bit / Very much
	demonstrated concern for student learning?	
	provided clear explanations on coursework and assessment?	
	stimulated you intellectually?	
	commented on your work in ways that help you learn?	
	seemed helpful and approachable?	
	set assessment tasks that challenge you to learn?	
In 2013, to what extent has your university course been delivered in a way that is:	well structured and focused?	Not at all / Very little / Some / Quite a bit / Very much
	relevant to your education as a whole?	

Table 21. Student Support items

Stem	Item	Response scale
At university during 2013, to what extent have you:	received support from your university to settle into study?	Not at all / Very little / Some / Quite a bit / Very much / Not applicable
	experienced efficient enrolment and admissions processes?	
	felt induction/orientation activities were relevant and helpful?	
During 2013, to what extent have you found administrative staff or systems (e.g. online administrative services, frontline staff, enrolment systems) to be:	available?	Had no Contact / Not at all / Very little / Some / Quite a bit / Very much / Not applicable
	helpful?	
During 2013, to what extent have you found careers advisors to be:	available?	Had no contact / Not at all / Very little / Some / Quite a bit / Very much / Not applicable
	helpful?	
During 2013, to what extent have you found academic or learning advisors to be:	available?	Had no contact / Not at all / Very little / Some / Quite a bit / Very much / Not applicable
	helpful?	
During 2013, to what extent have you found support services such as counsellors, financial/legal advisors and health services to be:	available?	Had no contact / Not at all / Very little / Some / Quite a bit / Very much / Not applicable
	helpful?	
At university during 2013, to what extent have you:	used university services (e.g. phone hotlines, online support, learning skills service, careers service, childcare, health service) to support your study?	Not at all / Very little / Some / Quite a bit / Very much
	been offered support relevant to your circumstances?	Not at all / Very little / Some / Quite a bit / Very much / Not applicable
	received appropriate English language skill support?	

Table 22. Learning Resources items

Stem	Item	Response scale
Thinking of this year, overall how would you rate the following learning resources provided for your university course?	Teaching spaces (e.g. lecture theatres, tutorial rooms, laboratories)	Poor / Fair / Good / Excellent / Not applicable
	Student spaces and common areas	
	Online learning materials	
	Computing/IT resources	
	Assigned books, notes and resources	
	Laboratory or studio equipment	
	Library resources and facilities	

Table 23. Open-response items

Stem	Item	Response scale
What have been the best aspects of your course?	What have been the best aspects of your course?	Open response
What aspects of your course most need improvement?	What aspects of your course most need improvement?	

Appendix B: Production of scores

A series of steps are taken to produce the focus area percentage satisfied results used in this report. A selection of the SPSS syntax used to produce these scores is presented below.

To begin, all UEQ items are rescaled into the conventional reporting metric. Four-point scales are recoded onto a scale that runs from 0, 33.3, 66.6 and 100, and five-point scales recoded onto a scale that runs from 0, 25, 50, 75 and 100. These rescaled items are denoted with an "r" suffix. The SPSS syntax to recode the UEQ items to the conventional reporting metric is shown in Figure 5.

```
RECODE qlovledu (1=0) (2=33.3) (3=66.6) (4=100) (ELSE=SYSMIS) INTO qlovledur.
RECODE partidiscus (1=0) (2=33.3) (3=66.6) (4=100) (ELSE=SYSMIS) INTO partidiscusr.
…
RECODE qllibres (1=0) (2=33.3) (3=66.6) (4=100) (ELSE=SYSMIS) INTO qllibresr.

RECODE supsettle (1=0) (2=25) (3=50) (4=75) (5=100) (ELSE = SYSMIS) INTO supsettler.
RECODE effenrolm (1=0) (2=25) (3=50) (4=75) (5=100) (ELSE = SYSMIS) INTO effenrolmr.
…
RECODE englang (1=0) (2=25) (3=50) (4=75) (5=100) (ELSE = SYSMIS) INTO englangr.
```

Figure 5. SPSS syntax to recode UEQ items into the conventional reporting metric

Scores for each focus area are then computed as the mean of the constituent item scores. A focus area score is only computed for respondents who have a valid item score for at least six skill development items, five learner engagement items, eight teaching quality items, six student support items and five learning resources items respectively. The SPSS syntax used to generate focus area average scores is shown in Figure 6. The recoded item scores are not retained in the analysis file.

Because the reporting metric for the 2013 UES is percentage satisfied (see Section 1.3), satisfaction variables must be created for each focus area. Percentage satisfied results reflect the percentage of students who achieve a threshold focus area score of 55 or greater. At the individual response level, satisfaction is represented by a binary variable taking the value of one if the student is satisfied with a particular facet of their higher education experience and zero otherwise. The SPSS syntax used to generate these satisfaction variables is presented in Figure 7.

```
COMPUTE DEVELOPMENT=MEAN.6(expthinkr, expprbslvr, expwrkothr, expconfindr, expwritingr,
expspeakr, expknowlr, expwrkskillr).
COMPUTE ENGAGEMENT=MEAN.5(opplocr, sensebelongr, feelpreparedr, partidiscusr, workothersr,
interactothr, interactdiffr).
COMPUTE TEACHING=MEAN.8(qlteachr, qlovledur, stdstrucr, stdrelevr, tchactivengr, tchconlrnr,
tchclexpecr, tchstimintr, tchfeedbckr, tchhelpappr, tchasschlngr).
COMPUTE SUPPORT=MEAN.6(englangr, offsupr, indorienr, supsettler, uniservicesr, admavailr, admhelpr,
caravailr, carhelpr, acdavailr, acdhelpr, supavailr, suphelpr, effenrolmr).
COMPUTE RESOURCES=MEAN.5(qltchspcr, qlstdspcr, qlonlmatr, qlcompitr, qltxtbookr, qlequipr, qllibresr).
```

Figure 6. SPSS syntax used to compute UES focus area scores

```
RECODE DEVELOPMENT (55 THRU 100=1) (MISSING=SYSMIS) (ELSE=0) INTO DEVELOPMENT_SAT.
RECODE ENGAGEMENT (55 THRU 100=1) (MISSING=SYSMIS) (ELSE=0) INTO ENGAGEMENT_SAT.
RECODE TEACHING (55 THRU 100=1) (MISSING=SYSMIS) (ELSE=0) INTO TEACHING_SAT.
RECODE SUPPORT (55 THRU 100=1) (MISSING=SYSMIS) (ELSE=0) INTO SUPPORT_SAT.
RECODE RESOURCES (55 THRU 100=1) (MISSING=SYSMIS) (ELSE=0) INTO RESOURCES_SAT.
```

Figure 7. SPSS syntax used to compute focus area satisfaction variables

At the item level, satisfaction reflects a response in the top two categories of both the four- and five-point response scales. As with the focus area satisfaction variables discussed previously, satisfaction with a particular UEQ item is represented by a binary variable taking the value of one if the student is satisfied and zero otherwise. An excerpt of the SPSS syntax used to generate these item satisfaction variables is presented in Figure 8.

```
RECODE qlovledu (1=0) (2=0) (3=1) (4=1) (ELSE=SYSMIS) INTO qlovledu_sat.
RECODE partidiscus (1=0) (2=0) (3=1) (4=1) (ELSE=SYSMIS) INTO partidiscus_sat.
…
RECODE qllibres (1=0) (2=0) (3=1) (4=1) (ELSE=SYSMIS) INTO qllibres_sat.

RECODE supsettle (1=0) (2=0) (3=0) (4=1) (5=1) (ELSE=SYSMIS) INTO supsettle_sat.
RECODE effenrolm (1=0) (2=0) (3=0) (4=1) (5=1) (ELSE=SYSMIS) INTO effenrolm_sat.
…
RECODE englang (1=0) (2=0) (3=0) (4=1) (5=1) (ELSE=SYSMIS) INTO englang_sat.
```

Figure 8. SPSS syntax used to compute item satisfaction variables

Appendix C: Comparison of the UEQ and CEQ

The Course Experience Questionnaire (CEQ) has been administered as a component of the Australian Graduate Survey (AGS) since 1993. The CEQ measures 11 facets of the higher education experience, which are underpinned by 49 Likert-type items that share a common five-point response scale. The response categories are currently labelled strongly disagree, disagree, neither agree nor disagree, agree and strongly agree.[18]

As part of the 2012 UES, the ACER-led Consortium were asked to investigate the conceptual and empirical relationship between the UES focus areas and CEQ scales. The Consortium concluded that six of the eleven CEQ scales should be fully incorporated into the UES, with the remaining five CEQ scales phased out from national administration and the name "CEQ" itself be discontinued.

During the development of the 2013 UES, it was decided, based on the report of the 2012 UES, that a further trial of the CEQ should be undertaken to provide more comprehensive evidence than that previously presented regarding the role or place of the CEQ in relation to the UES. The six CEQ scales recommended for retention in 2012 were administered on the 2013 UES and presented to students of 14 participating institutions who agreed to participate in the CEQ trial (see Table 3). These six scales and their constituent items are summarised in Table 24. The CGS item denoted with "R" is negatively worded and is reverse coded for analysis. This is necessary to ensure that all items within the scale measure the underlying phenomenon in a uniform direction.

Table 24. CEQ scales and items administered on the 2013 UES

Scale	Item
Good Teaching Scale	The staff put a lot of time into commenting on my work.
	The teaching staff normally gave me helpful feedback on how I was going.
	The teaching staff of this course motivated me to do my best work.
	My lecturers were extremely good at explaining things.
	The teaching staff worked hard to make their subjects interesting.
	The staff made a real effort to understand difficulties I might be having with my work.
Generic Skills Scale	The course helped me develop my ability to work as a team member.
	The course sharpened my analytic skills.
	The course developed my problem-solving skills.
	The course improved my skills in written communication.
	As a result of my course, I feel confident about tackling unfamiliar problems.
	My course helped me to develop the ability to plan my own work.
Overall Satisfaction Item	Overall, I was satisfied with the quality of this course.
Clear Goals and Standards	It was always easy to know the standard of work expected.
	I usually had a clear idea of where I was going and what was expected of me in this course.
	It was often hard to discover what was expected of me in this course. R
	The staff made it clear right from the start what they expected from students.
Graduate Qualities Scale	The course provided me with a broad overview of my field of knowledge.
	The course developed my confidence to investigate new ideas.
	University stimulated my enthusiasm for further learning.
	I learned to apply principles from this course to new situations.
	I consider what I learned valuable for my future.
	My university experience encouraged me to value perspectives other than my own.
Learning Community Scale	I felt part of a group of students and staff committed to learning.
	Students' ideas and suggestions were used during the course.
	I learned to explore ideas confidently with other people.
	I felt I belonged to the university community.
	I was able to explore academic interests with staff and students.

[18] Prior to 2010, only the anchor points were labelled.

During the development of the 2013 UES, it was acknowledged that the wording of the CEQ was aimed at students who had graduated from their course and that the tense of these items may not be appropriate for commencing students in the first year of their studies. As such, it was decided to trial a present tense version of the CEQ. Students identified as commencers received a present tense version of the CEQ, with all other students from the institutions participating in the trial receiving the original, past tense version. The latter wording is the version presented in Table 24.

For consistency with the UES, CEQ items were recoded onto a scale that runs from 0, 25, 50, 75 and 100. Scale scores were computed as the mean of the constituent item scores, and satisfaction variables were then created from these scale scores based on the same threshold of 55 adopted for the UES (see Section 1.3).

Because the UES focus area scores and the CEQ scale scores are represented as binary satisfaction indicators, the tetrachoric correlation is used to compare the two sets of variables.[19] The tetrachoric correlation coefficient estimates the relationship between two dichotomous variables assuming that the underlying trait is continuous and normally distributed. Based on this assumption, the tetrachoric correlation coefficient can be interpreted in the same manner as a correlation coefficient calculated on a continuous scale, with 0 indicating no linear relationship and 1 indicating a perfect relationship. The results of this analysis are presented in Table 25.

Table 25. Tetrachoric correlation matrix of UES focus areas and CEQ scales

	Skills Development	Learner Engagement	Teaching Quality	Student Support	Learning Resources
Good Teaching Scale	0.59	0.37	0.77	0.53	0.50
Generic Skills Scale	0.79	0.47	0.67	0.46	0.49
Overall Satisfaction Item	0.67	0.43	0.80	0.52	0.55
Clear Goals and Standards	0.50	0.31	0.63	0.44	0.45
Graduate Qualities Scale	0.75	0.46	0.74	0.51	0.52
Learning Community Scale	0.65	0.60	0.66	0.52	0.51

It can be seen from Table 25 that the correlations between UES focus areas and CEQ scales range from moderate to strong, with correlation coefficients ranging from 0.31 to 0.80. Given that both of the instruments were developed to measure the higher education experience, a strong relationship is not an unexpected finding. The strongest correlations tended to be observed in relation to the teaching quality and skills development focus areas, with weaker correlations generally observed in relation to the student support, learning resources and learner engagement focus areas. This is not unexpected, given that these last three focus areas appear to be conceptually related to CEQ scales that were not administered as part of the 2013 trial, such as learning resources and student support.

Many of the strongest correlations observed in Table 25 make good intuitive sense, for example, between the UES skills development focus area and the CEQ generic skills scale (0.79) and graduate qualities scale (0.75); and between the UES teaching quality focus area and the CEQ overall satisfaction item (0.80), good teaching scale (0.77) and graduate qualities scale (0.74).

To some extent, the strong correlations between some of the UES focus areas and CEQ scales raises the question of whether there is any utility at all in administering CEQ items as part of the UES. The results in Table 25 suggest that doing so would result in little additional benefit, since the CEQ scales

[19] This analysis was conducted using the *tetrachoric* command in Stata.

appear to be measuring facets of the student experience already addressed by the UES. Considering that both the CEQ and UES were designed to be standalone instruments, the fact that they appear to measure similar things should come as no surprise. Given that UES respondents are already tasked with completing 47 items on their higher education experience, or twice that number if they are completing a double or combined degree, the addition of potentially redundant items may be seen as increasing respondent load without improving measurement of the UES focus areas to an appreciable degree. Any new questionnaire items should be developed with the goal of enhancing the instrument, based on a holistic analysis of its psychometric properties (see Appendix F).

An argument can be made, however, that administering selected CEQ scales on the UES would allow for additional international benchmarking (see Section 4.5). The CEQ and the UK National Student Survey (NSS), administered to final-year undergraduate students at most of the UK's higher education institutions, have several questions with very similar wording. Most notably, both surveys have an overall satisfaction item with near-identical wording,[20] measured on a five-point Likert-type response scale. In light of the importance of international benchmarking to the Australian higher education sector, the inclusion of several CEQ scales on the UES is justified on this basis.

Recommendation 6*:* For purposes of international benchmarking, the CEQ Good Teaching Scale, Generic Skills Scale, Clear Goals and Standards Scale and Overall Satisfaction Item should be administered on the UES to a sample of non-commencing students across all participating institutions.

[20] "Overall, I am satisfied with the quality of the [this] course."

Appendix D: Promotional website tiles

2013 University Experience Survey National Report

March 2014

Figure 9. UES 2013 promotional website tiles

Appendix E: Construction of confidence intervals

The 95 per cent confidence intervals presented in Tables 10 and 11 were calculated using the Finite Population Correction (FPC) to account for the relatively large size of the sample relative to the in-scope population. The FPC is generally used when the sampling fraction exceeds 5 per cent. In order to calculate the standard errors for the survey estimates, no non-response bias was assumed and thus simple random sample survey errors were used. This approach is similar to the one employed to construct confidence intervals for the UES estimates presented on the *MyUniversity* website.

Because percentage agreement scores are reported for the 2013 UES, the formula for the confidence interval of a proportion is used.

$$95\%\ CI\ bound(\hat{p}) = 1.96 \times FPC \times SE(\hat{p}) = 1.96 \times \sqrt{\frac{N-n}{N-1}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where $\hat{p}$ is the estimated proportion of satisfied responses (i.e. the top two response categories), N is the size of the population in the relevant subgroup, n is the number of valid responses in the relevant subgroup, FPC is the Finite Population Correction and $SE(\hat{p})$ is the standard error. The survey frame (see Section 2.3.1) was used to determine the size of the population.

The 95 per cent confidence interval of each estimated proportion is then calculated as the proportion plus or minus its 95 per confidence interval bound.

The use of simple random sample survey errors assumes a simple random sample at the level of estimation. A national estimate, for example, assumes that the survey was a simple random sample at the national level, while subject area estimates assumes that it was a simple random sample at the national subject area level. Because the UES was conducted using stratified sampling at the institution by subject area level (see Section 2.3.2), standard errors calculated at the national level will be upwardly-biased. As such, the confidence intervals presented in Tables 10 and 11 are conservative and should be treated as indicative only.

Weighted stratified estimates would be more efficient and potentially more representative than those presented in this report; however the relatively fine stratification in the UES results in strata sample sizes that are, in many cases, too small to allow the calculation of the standard errors of the weighted estimates (i.e. too many strata with n < 25).

Appendix F: Psychometric analysis of the UEQ

F.1 Introduction

This report summarises a psychometric analysis of the UES items and facets, in which the Rasch measurement model was used to assess the following aspects:

- How well the items in each facet seemed to be working together to measure a common trait.
- Whether the rating scale categories were used by respondents in a consistent manner.
- Whether there was a distinctive hierarchy of items and persons along the measured variable.
- How well the items were matched to the sample of respondents.
- Whether some groups of respondents seemed to be responding to items in very different ways from other respondents.

These aspects will be explored for each facet in turn. Recommendations for refining the existing items and facets will be made.

The work described here complements the analytical work already carried about by GCA, in particular the factor analyses used to confirm the facet membership of each item.[21]

F.2 Analytical approach

The Rasch model provides many outputs that can be used to test how well questionnaire items contribute to an underlying trait (or dimension) and also how consistently respondents answer questions. As explained in Section F.6, the model estimates the probability that a person with a given attitude will choose a particular response to an item. Persons who possess high levels of the underlying trait will be more likely to endorse items. By contrast, persons with low levels of the underlying trait will be more likely to disendorse items. Some items are very easy for respondents to endorse (that is, the item difficulty is low) whereas other items are only endorsed by those respondents with high levels of the underlying trait (such items have a high difficulty).

Overall then, for a set of items that are effectively measuring the underlying dimension, we expect:

- Difficult items that are most likely to be endorsed by those with high levels of the underlying trait;
- Easy items that are most likely to be disendorsed by those with low levels of the trait; and
- A predictable progression in between.

The Rasch model provides a number of statistics summarising how well items and persons fit this expected progression in difficulty and attitude. Items or persons that deviate significantly from our expectation are evidence of items that measure different traits or persons that answer questions in unexpected ways.[22] Both of these are undesirable and degrade the quality of derived measures.

[21] The results of the factor analyses are not presented in the interest of brevity. In relation to each of the focus areas, all items loaded onto a single factor. Each factor analysis was based on a polychoric correlation matrix to account for the ordinal nature of the UES items.

[22] Although it is generally assumed that respondents understand the questions they are asked and then answer them in a coherent way, the Rasch model provides numerous statistics for detecting evidence to the contrary.

The following diagnostic outputs were used in the assessment of quality.

F.2.1 Fit to the Rasch model

The Rasch model calculates several fit statistics for both items and persons. Several that are of particular use are explained in Table 26.

Table 26. Selected Rasch model outputs for item assessment

Statistic	Meaning	Ideal range
Infit mean square	This is the mean of the squared residuals, [23] giving relatively more weight to the performances of persons closer to the item value.	0.6-1.4 for rating scale items with an expected value of 1. Low values indicate items whose responses can be easily predicted from other items. High values indicate unpredictable responses.
Outfit mean square	This is the mean of the squared residuals, across all items	
Point-measure correlation	This is the correlation between the Rasch measures and the responses for an item.	Low values indicate poor fit and negative values suggest miscoding (where a scale is reversed relative to other items).
Item discrimination	This relates to how well an item discriminates between high and low scoring persons.	The expected value is 1. High values indicate better than expected by the model and low values indicate an item that discriminates less than expected.

F.2.2 Ordered category thresholds

In analysing rating scale data, it is important to assess how well the categories are contributing to the creation of interpretable measures. This is determined by checking that the categories fit the model, namely that the difficulty of selecting item categories progresses in a hierarchical manner. Where too few respondents select a category or where there are too many categories, the difficulties of some item categories will be unpredictable. Disordering or instability among item categories can be detected by plotting the probability of responding to any particular category, given the difference between a person's attitude and the item's difficulty.

F.2.3 Item and person targeting

Measurement tools need to be "fit for purpose" in the sense that they must be designed for the persons expected to be measured. For instance, in the field of educational testing, a teacher would not administer a test designed for 12 year-olds to those who are already 15. Although some of the less capable 15 year-olds may struggle with the more difficult questions, we would expect the test to be too easy for the group as a whole. If the Rasch model were used to analyse such data, we would expect to see person scores that are well above the item difficulties. When the items and the persons are not well matched, the instrument has limited use as a diagnostic tool.

Since the Rasch model calculates person and item estimates on the same scale, a side-by-side plot of person and item measures quickly shows how well matched the instrument is to the intended respondents.

[23] The residual values represent the differences between the Rasch model's theoretical expectation of item performance and the performance actually encountered for the item. Following usual statistical convention, residuals are squared to make the difference between actual and predicted values positive.

F.2.4 Unidimensionality

A requirement of the Rasch model is that measures must be unidimensional, so that constructing measures should proceed by one clearly theorized trait at a time. Indicators of misfit are typically used to reveal the extent to which any item or person performance suggests more than one underling latent trait is at work.[24] A further tool is the principal components analysis of Rasch model residuals, which can identify items with substantial variance that remains unexplained by the primary Rasch measure. Given the factor analysis already carried out by GCA to confirm the dimensionality and item membership of the facets, an investigation of Rasch residuals has not been undertaken as part of the work reported here.

The next section summarises these aspects of quality for each of the five UES focus areas.

F.3 Assessing item and facet quality

F.3.1 Learner Engagement

This section summarises results for the Learner Engagement facet, which comprises the following items and response scale types:[25]

Variable	Item text	Scale
feelprepared	Felt prepared for study	Extent
interactdiff	Interacted with different students	Frequency
interactoth	Student interaction outside study	Frequency
opploc	Opportunities to interact with local students	Extent
partidiscus	Online or face-to-face discussions	Frequency
sensebelong	Sense of belonging to university	Extent
workothers	Worked with other students	Frequency

Table 27 summarises the item statistics for the Learner Engagement focus area. These are presented in terms of their difficulty of endorsement. The item *sensebelong* ("At university during 2013, to what extent have you had a sense of belonging to your university?") was the most difficult item for respondents to endorse and *workothers* ("In 2013, how frequently have you worked with other students as part of your study?") was the easiest. Only *feelprepared* ("At university during 2013, to what extent have you felt prepared for your study?") showed any evidence of misfit, although this was only slight (a higher outfit mean-square than other items, accompanied by lower discrimination and lower point-measure correlation).

Table 27. Item statistics for Learner Engagement, in difficulty order

Item	Measure[26]	Infit mean square	Outfit mean square	Point-measure correlation	Item discrimination
sensebelong	0.31	0.97	0.97	0.65	1.03
interactoth	0.25	1.00	0.98	0.69	1.05
opploc	0.11	1.10	1.08	0.67	0.93
feelprepared	-0.02	1.25	1.33	0.49	0.55
interactdiff	-0.07	0.82	0.82	0.68	1.23
partidiscus	-0.15	0.95	0.94	0.65	1.06
workothers	-0.43	0.90	0.89	0.67	1.14

[24] Bond, T.G., & Fox, M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. (2nd ed.) Mahwah, NJ: Erlbaum.
[25] Refer to Appendix A for a list of response scale categories.
[26] The mathematical unit of Rasch measurement is the log-odds unit or "logit" (see Section F.6).

The next aspect of quality to be assessed was the functioning of the categories. The following plots illustrate the probability of responding to any given category, given the difference in estimates of person attitude and any item difficulty. Here, attitude refers to the latent trait "Learner Engagement". In Figure 0, for the *Extent* response scale, a person who is very low on the trait relative to an item (a large negative difference) is most likely to choose "Not at all". Conversely, where someone has a high measure on the trait relative to an item (a large positive difference) they are most likely to choose "Very much". The three middle categories ("Very little", "Some" and "Quite a bit") have much narrower ranges for their probability peaks, however, which is suggestive of too many response categories.

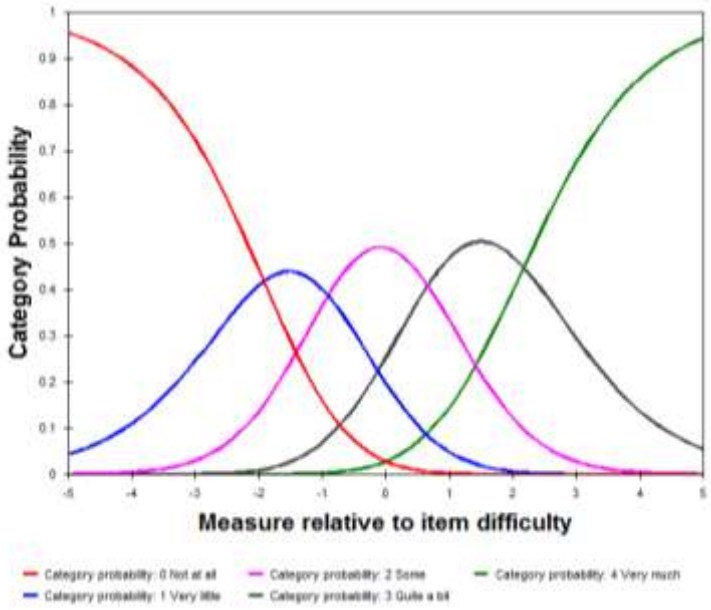


Figure 10. Category probabilities for original Extent scale

Since the other two scales (*Frequency* and *Rating*) each had four categories, several four-category variations of the *Extent* scale were trialled to determine which had the best statistical properties. Table 28 shows results for each of the ways into which the five *Extent* categories can be collapsed into four. The objective is to identify which categorisation of the variable results in a better ordering of the underlying variable, one that is more consistent with the theory that generated the items in the first place. The two statistics used are person and item separation (higher is better), which identify how well the measures produce a clear hierarchy. In this table, it is evident that collapsing the two lowest categories yielded the higher reliability for both persons and items. This categorisation is used subsequently in this report. The category probabilities for the revised scale are shown in Figure 11.

Table 28. Comparison of five categorisations for the Extent scale

Categorisation	Person separation	Item separation
12345	3.58	119.4
11234	3.68	127.9
12234	3.59	119.6
12334	3.2	108.7
12344	3.02	123.7

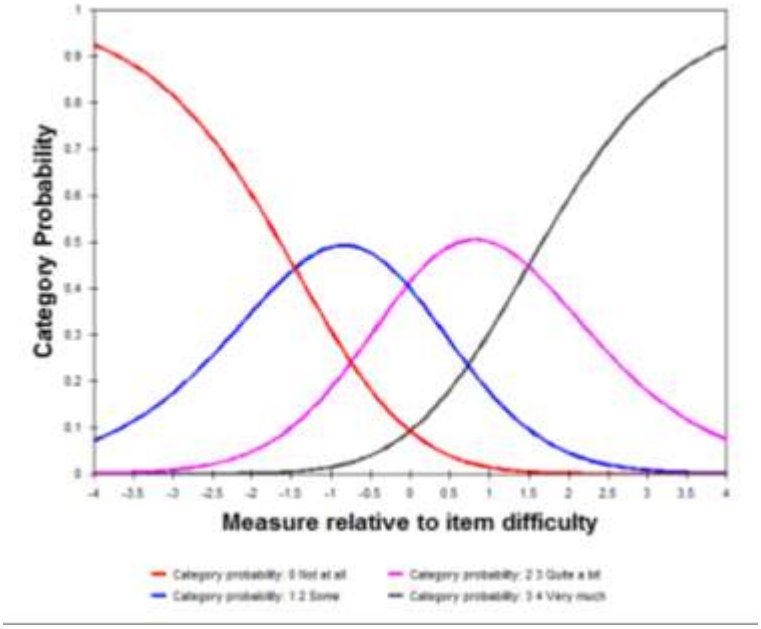


Figure 11. Category probabilities for revised Extent scale

The category probabilities for the *Frequency* scale are shown in Figure 12 below. The very narrow range of peak probability for the "Often" category suggests that reduction of this scale may also be merited but this is not considered further here.

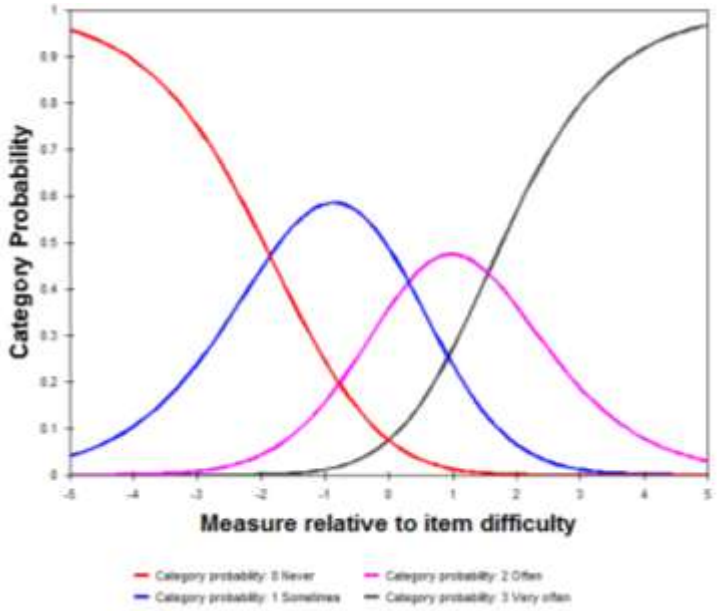


Figure 12. Category probabilities for Frequency response scale

The final aspect of facet quality is how well matched the items are to the persons who answered them. Since the item difficulties and person abilities are on the same scale, a simple histogram of the two sets of parameters will verify the match (see Figure 13). This graph shows the distribution of person measures on the left and the item difficulties on the right.[27] When items and persons are well matched, the two plots will show significant overlap, as they do here. Evident here, however, is that the persons cover a greater range of the underlying trait than do the items. This means that the respondents at the ends of the distribution (those who are very low on the trait and those who are very high) will be measured with limited precision.[28] Measurement would be improved for these respondents if there were additional items, some easier to endorse than the current items and some harder to endorse, that were a closer match to their location on the trait.

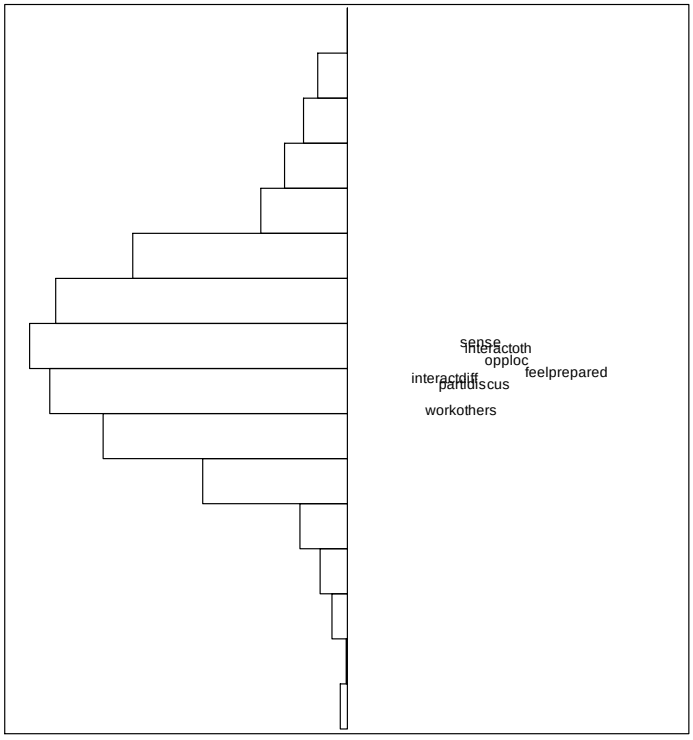


Figure 13. Person-item map for Learner Engagement

F.3.2 Learning Resources

This section presents Rasch model outputs for the Learning Resources focus area, which is made up of the following items:

[27] The x dimension for the items in Figure 13 is related to the outfit mean square (lower to higher from left to right) with an adjustment to minimise overlapping item labels.

[28] In particular, such persons will have large standard errors of measurement relative to those in the middle of the distribution. For example, one respondent approximately in the centre of the distribution had an estimated measure of 0.19 logits with an accompanying standard error of 0.49 logits. Another respondent at the very top of the scale measured 4.69 logits, but this was associated with a standard error of 1.84 logits signifying much less certainty in their location on the underlying trait.

Variable	Item text	Scale
qlcompit	Computing/IT resources - quality	Rating
qlequip	Laboratory or studio equipment - quality	Rating
qllibres	Library resources and facilities - quality	Rating
qlonlmat	Online learning materials - quality	Rating
qlstdspc	Student spaces - quality	Rating
qltchspc	Teaching spaces - quality	Rating
qltxtbook	Textbooks and learning resources - quality	Rating

The various fit statistics are shown in Table 29 and no issues are evident. The category probabilities for the *Rating* response scale are shown in Figure 14 and the person-item map in Figure 15. The issue of inadequate targeting is evident in the last plot, especially at the high end of the trait where a large group of students have achieved the maximum possible measure. If it was desired to estimate these respondents' measures more accurately or to discriminate better between the most satisfied students, more items would be required that are harder for the group to endorse.

Table 29. Item statistics for Learning Resources, in difficulty order

Item	Measure	Infit mean square	Outfit mean square	Point-measure correlation	Item discrimination
qlstdspc	0.45	1.10	1.10	0.72	0.90
qltxtbook	0.35	0.99	1.02	0.69	0.98
qlcompit	0.15	0.92	0.92	0.73	1.09
qlequip	-0.07	0.98	0.98	0.71	1.03
qltchspc	-0.12	0.96	0.96	0.70	1.04
qlonlmat	-0.22	1.02	1.03	0.68	0.97
qllibres	-0.54	0.99	0.97	0.69	1.03

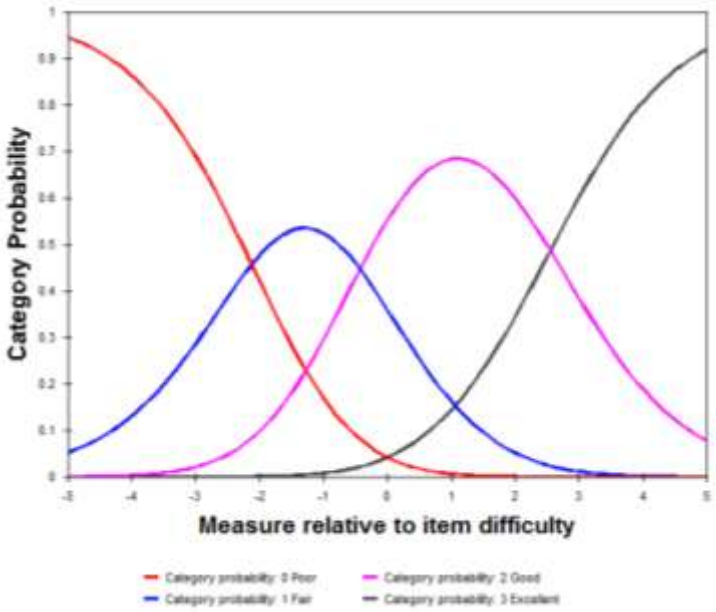


Figure 14. Category probabilities for the Rating response scale

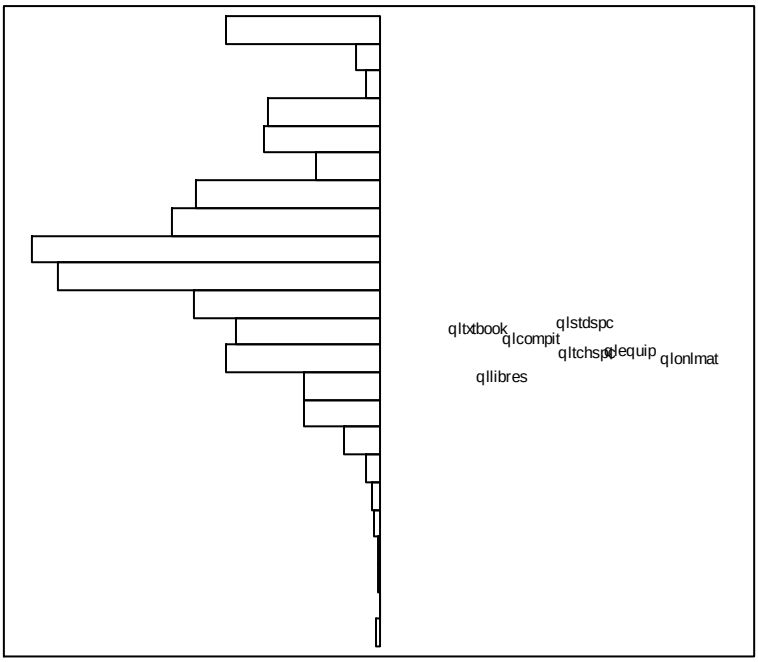


Figure 15. Person-item map for Learning Resources

F.3.3 Skills Development

This section summarises results for the Skills Development focus area, which comprises the following items, all on the *Extent* response scale:

Variable	Item text	Scale
expconfind	Independent learning developed by course	Extent
expknowl	Knowledge of study areas developed by course	Extent
expprbslv	Complex problem solving developed by course	Extent
expspeak	Spoken communication developed by course	Extent
expthink	Critical thinking skills developed by course	Extent
expwriting	Written communication developed by course	Extent
expwrkoth	Teamwork developed by course	Extent
expwrkskill	Work readiness developed by course	Extent

The various fit statistics are shown in Table 30 and no issues are evident. The category probabilities for the revised *Extent* response scale are shown in Figure 16, and the person-item map in Figure 17. Similar to the Learning Resources facet, additional items may be required to obtain better estimates for students at either end of the distribution, especially at the top.

Table 30. Item statistics for Skills Development, in difficulty order

Item	Measure	Infit mean square	Outfit mean square	Point-measure correlation	Item discrimination
expspeak	0.71	1.08	1.08	0.75	0.91
expprbslv	0.33	0.81	0.82	0.78	1.21
expwrkoth	0.22	1.18	1.20	0.71	0.79
expwrkskill	0.21	1.05	1.06	0.74	0.94
expwriting	0.15	1.05	1.07	0.73	0.94
expthink	-0.31	0.77	0.77	0.78	1.26
expconfind	-0.42	0.98	0.98	0.73	1.03
expknowl	-0.89	1.02	1.03	0.70	0.97

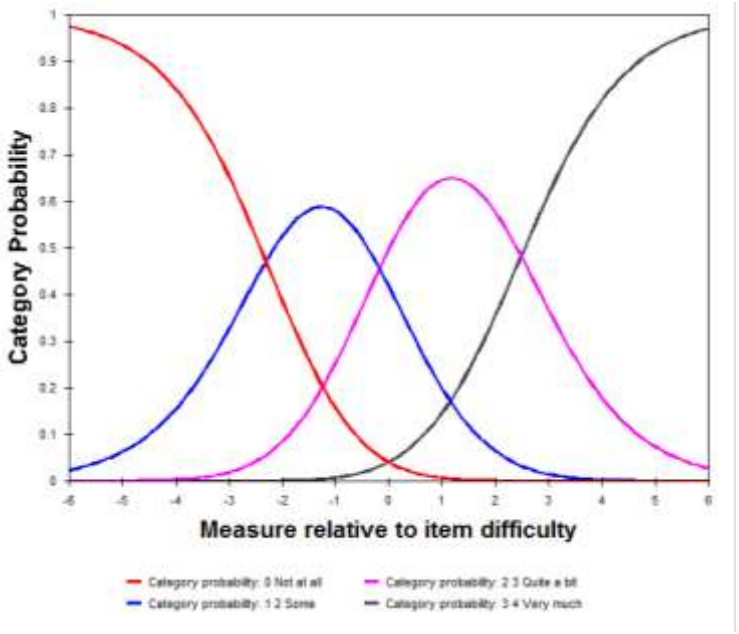


Figure 16. Category probabilities for the Extent response scale, Skills Development

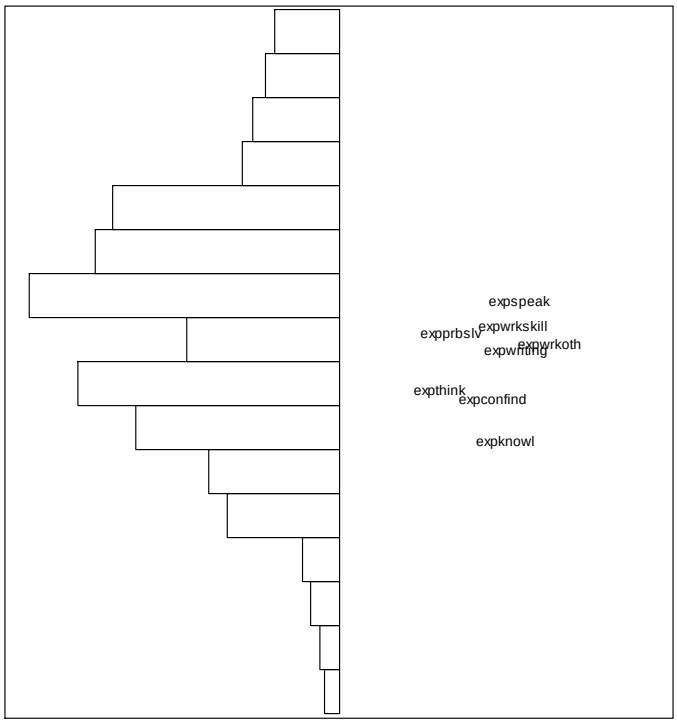


Figure 17. Person-item map for Skills Development

F.3.4 Student Support

This section summarises results for the Student Support focus area, which comprises the following items, all on the *Extent* response scale:

Variable	Item text	Scale
acdavail	Academic or learning advisors available	Extent
acdhelp	Academic or learning advisors helpful	Extent
admavail	Admin staff/systems available	Extent
admhelp	Admin staff/systems helpful	Extent
caravail	Careers advisors available	Extent
carhelp	Careers advisors helpful	Extent
effenrolm	Efficient enrolment and admissions processes	Extent
englang	English language support received	Extent
indorien	Induction / orientation activities relevant / helpful	Extent
offsup	Have been offered relevant support	Extent
supavail	Other advisors available	Extent
suphelp	Other advisors helpful	Extent
supsettle	Supported to settle into university	Extent
uniservices	University services used to support study	Extent

The fit statistics are shown in Table 31. One item *uniservices* ("At university during 2013, to what extent have you used university services to support your study?") fits the model less well than the others and may be measuring something different. Whereas other items in this facet are seeking ratings on helpfulness and availability, this item is about use of services. The category probabilities for the revised *Extent* response scale are shown in Figure 18, and the middle two categories have a particularly narrow window of being the most likely response – four categories may be too many for

this scale. The person-item map is shown in Figure 19 and the persons and items are well targeted, albeit with a likely loss of precision at the extremes, as seen previously.

Table 31. Item statistics for Student Support, in difficulty order

Item	Measure	Infit mean square	Outfit mean square	Point-measure correlation	Item discrimination
uniservices	1.53	1.69	1.79	0.45	0.55
englang	1.28	1.54	1.43	0.55	0.73
offsup	0.81	1.25	1.17	0.62	0.88
carhelp	0.12	0.93	0.93	0.66	1.09
caravail	0.11	0.90	0.92	0.66	1.09
supavail	-0.11	0.89	0.89	0.67	1.14
suphelp	-0.15	0.93	0.92	0.67	1.12
supsettle	-0.19	0.93	0.95	0.65	1.08
indorien	-0.23	1.05	1.09	0.61	0.89
admhelp	-0.49	0.79	0.79	0.68	1.28
acdavail	-0.54	0.72	0.74	0.68	1.32
admavail	-0.59	0.76	0.77	0.67	1.28
acdhelp	-0.64	0.75	0.76	0.68	1.30
effenrolm	-0.91	1.12	1.18	0.58	0.82

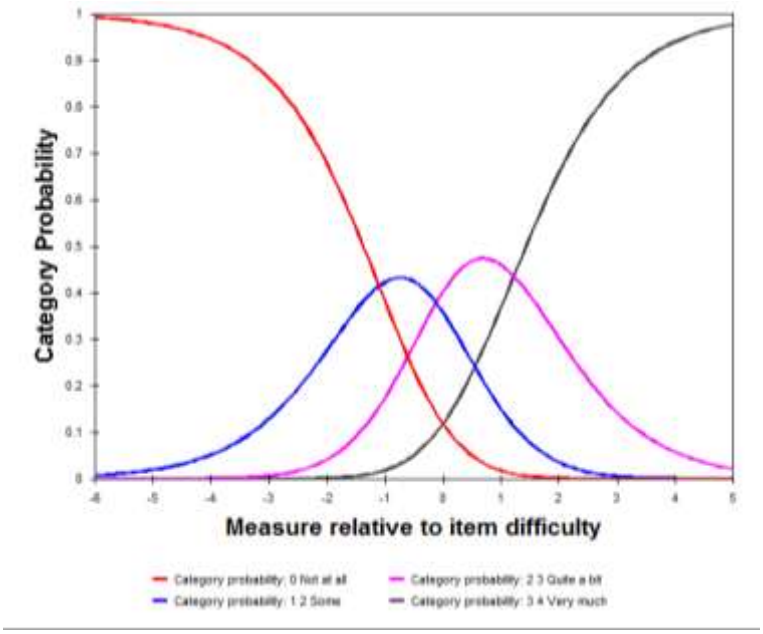


Figure 18. Category probabilities for the Extent response scale, Student Support

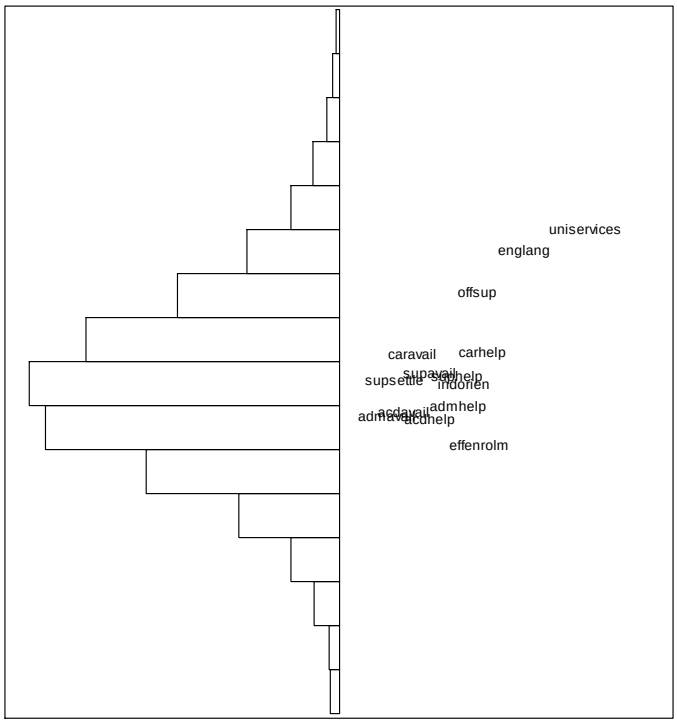


Figure 19. Person-item map for Student Support

F.3.5 Teaching Quality

This section summarises results for the Teaching Quality focus area, which comprises the following items, based on a combination of the *Extent* and *Rating* response scales:

Variable	Item text	Scale
qlovledu	Quality of overall educational experience	Rating
qlteach	Quality of teaching	Rating
stdrelev	Course relevant to education overall	Extent
stdstruc	Course well structured	Extent
tchactiveng	Teaching staff actively engaged students	Extent
tchasschlng	Teaching staff set challenging assessments	Extent
tchclexpec	Teaching staff explained coursework and assessment	Extent
tchconlrn	Teaching staff concerned about student learning	Extent
tchfeedbck	Teaching staff provided constructive feedback	Extent
tchhelpapp	Teaching staff were helpful and approachable	Extent
tchstimint	Teaching staff provided intellectual stimulation	Extent

The item statistics for Teaching Quality are summarised in Table 32, where *tchfeedbck* ("During 2013, to what extent have your lecturers, tutors and demonstrators commented on your work in ways that help you learn?") was the most difficult item to endorse and *qleach* ("Thinking of this year, overall at your university, how would you rate the quality of the teaching you have experienced?") was the easiest. There was no evidence of misfitting items. Category responses are shown in Figures 20 and 21 for the *Extent* and *Rating* scales, respectively. The person-item map for this facet appears in Figure 22, where we see the familiar clustering of items around the middle of the person distribution.

Table 32. Item statistics for Teaching Quality, in difficulty order

Item	Measure	Infit mean square	Outfit mean square	Point-measure correlation	Item discrimination
tchfeedbck	1.07	1.21	1.20	0.73	0.77
tchconlrn	0.54	1.03	1.02	0.76	0.98
tchactiveng	0.35	0.84	0.83	0.79	1.19
tchclexpec	0.27	1.09	1.10	0.72	0.89
stdstruc	0.23	0.85	0.88	0.76	1.15
tchstimint	-0.02	0.92	0.91	0.76	1.09
stdrelev	-0.23	1.07	1.10	0.71	0.90
tchhelpapp	-0.32	0.98	0.96	0.75	1.03
tchasschlng	-0.47	1.10	1.10	0.69	0.89
qlovledu	-0.66	0.96	1.00	0.72	1.03
qlteach	-0.75	0.88	0.89	0.75	1.11

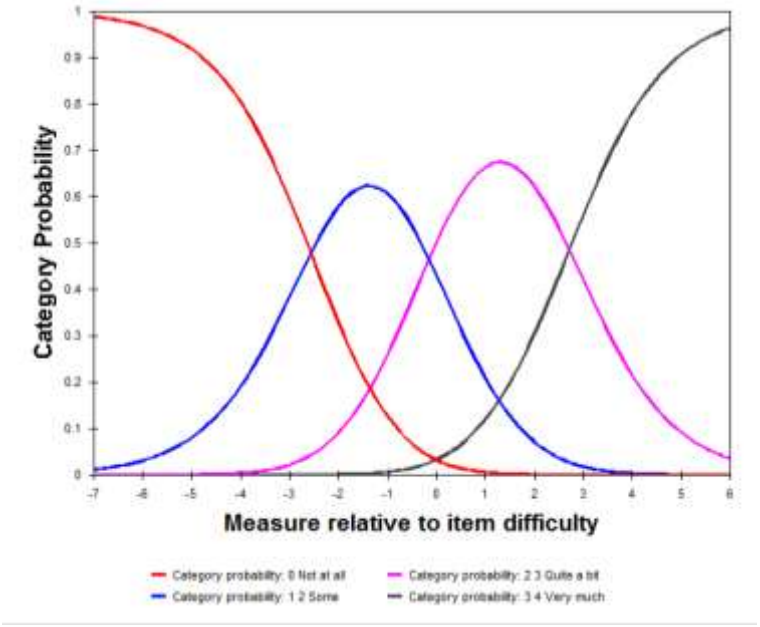


Figure 20. Category probabilities for the Extent response scale, Teaching Quality

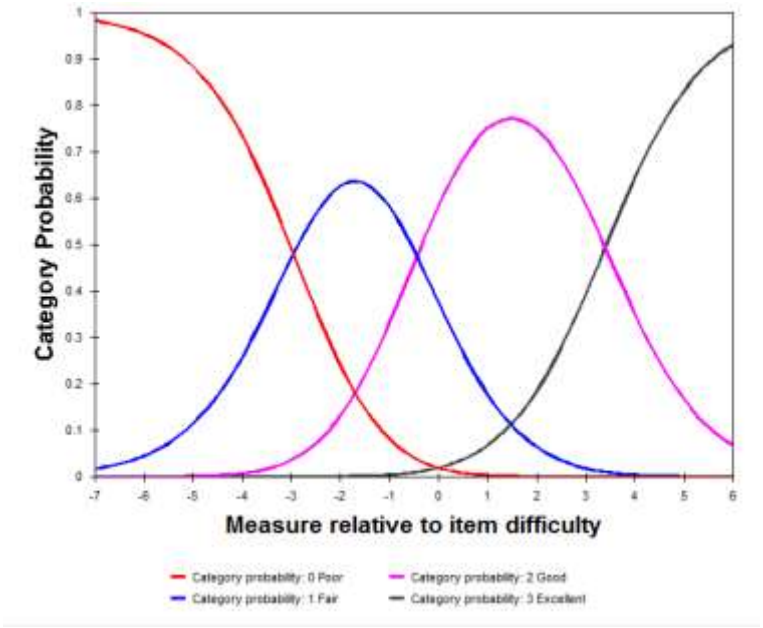


Figure 21. Category probabilities for the Rating response scale, Teaching Quality

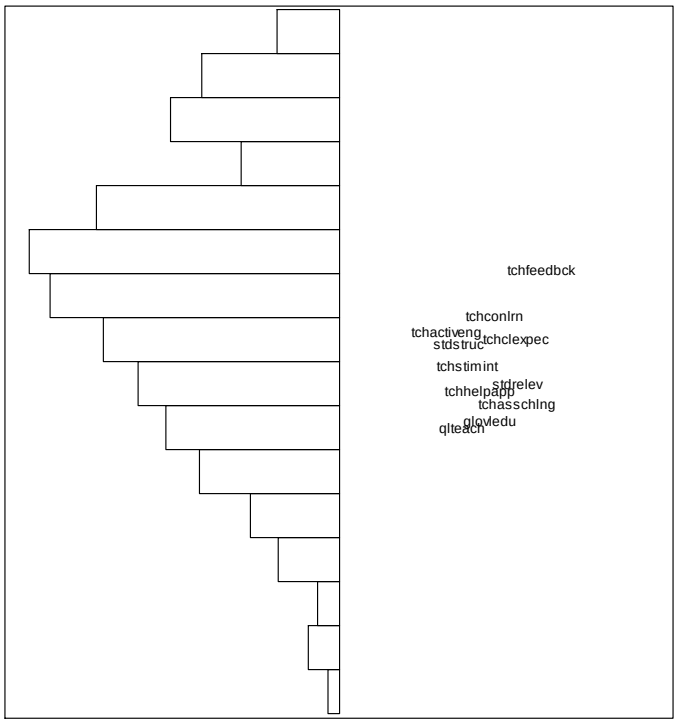


Figure 22. Person-item map for Teaching Quality

F.4 Differential item functioning

Differential Item Functioning, DIF, indicates that one group of respondents scored better than another group of respondents on an item (after adjusting for the overall scores of the respondents). In general, we do not expect an item to be unexpectedly easy (or difficult) for any particular group to endorse. The Rasch model statistics cannot determine the reason for differential item functioning, but does identify items for closer scrutiny to ensure that including them in the instrument does not "disadvantage" any particular group.

The UES collects a great many demographic characteristics of respondents, including gender, age, course type, attendance mode, citizenship status, and so on. Each of these can be inspected for evidence of DIF, and this section presents selected results of these comparisons. The rule of thumb for assessing DIF is that difficulty differences greater than 0.5 logits may be evidence of substantive DIF effects.[29]

For gender, the largest difference between females and males was for *expwriting* ("To what extent has your experience at university developed your written communication skills?"). This difference was only 0.24 logits (more difficult for males than females) so had no substantive impact on measures. This item showed a difficulty difference of 0.49 logits in relation to stage of study, with commencing students finding the item harder to endorse than later year students.

For age, the largest difference of 0.82 logits was for those aged 25+ on *interactoth* ("In 2013, how frequently have you interacted with students outside study requirements?"). Older respondents were much less likely than expected to endorse this item than younger respondents. The same item yielded a difficulty difference of 0.96 logits for those students who completed their study by external or mixed-mode delivery.

With respect to Indigenous status, the largest difference of 0.57 logits was observed for *offsup* ("At university during 2013, to what extent have you been offered support relevant to your circumstances?*")*. The Indigenous group found this item easier to endorse than non-Indigenous students. This item also showed the greatest difference based on whether students reported a disability, with those who did finding the item 0.85 logits easier to endorse than the other students.

On the basis of these comparisons, there is limited evidence of DIF effects among groups of respondents, and a more comprehensive analysis across other respondent characteristics may yield more substantive differences.

F.5 Summary

This report has summarised the results of a psychometric analysis of the UES items, undertaken through the use of the Rasch measurement model. In the process of constructing linear measures from rating scale data, the model produces a great many outputs that can be used to assess the quality of the scales and items. The main findings of the analysis were as follows:

[29] Linacre, J.M. (2013). *Winsteps® Rasch measurement computer program user's guide*. Beaverton, Oregon: Winsteps. com

- The items in each of the facets seemed to act in unison to measure an underlying variable, with the possible exception of *uniservices* ("At university during 2013, to what extent have you used university services to support your study?"). This item should be revisited to confirm the theoretical basis for its inclusion in Student Support, and its omission from the analysis would improve measures for this focus area. The Learner Engagement item *feelprepared* ("At university during 2013, to what extent have you felt prepared for your study?") showed slight evidence of misfit although to a lesser degree.

- The items were generally well-targeted to the population of respondents. The precision of measures could be improved for those on the tails of the underlying variables through the development of additional items specifically targeted at those who scored high or low on the variables.

- The rating scale categories generally functioned acceptably, particular at the extremes of the scale. Collapsing the two lowest *Extent* categories proved beneficial to measurement, and it is recommended to make this a four-category scale in future.

- There was some evidence of differential item functioning for a handful of items, and the relevance of these items to different groups of respondents should be confirmed.

Rasch measures have the interval-scale properties necessary for statistical purposes, so may be used in models and summaries of UES results.[30] By fixing the item difficulties to their 2013 values a "UES scale" may then be developed that would enable direct comparisons across time and institutions, a necessity for valid benchmarking.

Recommendation 7: The item, "At university during year x, to what extent have you used university services to support your study?" should be omitted on the basis of psychometric evidence of lack of fit within the Student Support focus area.

F.6 About the Rasch model

Responses to the questionnaires were analysed using the Rasch measurement model. Rasch analysis is a method for obtaining objective, fundamental, linear measures from stochastic observations of ordered category responses.[31] It calculates measures that are directly comparable across different administrations of a questionnaire and seamlessly accommodates missing data. As already described, in the process of deriving measures the Rasch model provides a large range of diagnostics about the quality of the measures and of the items used in their construction. These diagnostics enable the direct assessment of how well the measure "holds together" and whether or not the individual items contribute usefully to the construction of scores.

In mathematical notation, the Rasch model may be represented by the following equation:

$$\log_e\left(\frac{\pi_{nik}}{1-\pi_{nik}}\right) = \beta_n - \delta_n - \tau_k$$

[30] This is in contrast to "raw score" methods, such as adding or averaging scores, which treat ordinal categories as if they represent interval-scale measures.
[31] Linacre, *loc. cit.*

where π_{nik} is the probability of person n on item i choosing category k; β_n is the person attitude, δ_i is the item difficulty, and τ_k is the difficulty of threshold k. This equation is solved iteratively to yield estimates and standard errors for each of these parameters (one for each person, item and item threshold).

An attractive and intuitive feature of the Rasch model is that a person's likelihood of endorsing a particular item depends only on the person's attitude and the item difficulty.

Measures were calculated for each person who completed one or more of the items in each facet. The usual scale for Rasch measures is logits (log-odds), which has a theoretical range of $(-\infty, +\infty)$.

Appendix G: Response category percentages

Table 33. Skills Development item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Developed critical and analytical thinking	Not at all	1	1	1
	Very little	5	4	5
	Some	28	22	26
	Quite a bit	45	43	44
	Very much	21	30	25
Developed ability to solve complex problems	Not at all	2	1	2
	Very little	9	6	8
	Some	36	28	33
	Quite a bit	39	42	40
	Very much	14	22	17
Developed ability to work effectively with others	Not at all	3	2	3
	Very little	9	8	9
	Some	31	26	29
	Quite a bit	38	39	39
	Very much	19	25	21
Developed confidence to learn independently	Not at all	1	1	1
	Very little	6	4	5
	Some	26	20	23
	Quite a bit	43	42	42
	Very much	25	33	28
Developed written communication skills	Not at all	2	2	2
	Very little	9	7	8
	Some	32	24	29
	Quite a bit	39	41	40
	Very much	17	28	22
Developed spoken communication skills	Not at all	4	3	4
	Very little	14	10	13
	Some	35	29	32
	Quite a bit	33	37	34
	Very much	14	22	17
Developed knowledge of field studying	Not at all	1	1	1
	Very little	3	3	3
	Some	20	18	19
	Quite a bit	44	42	43
	Very much	32	36	34
Developed work-related knowledge and skills	Not at all	2	2	2
	Very little	8	9	9
	Some	30	27	29
	Quite a bit	40	38	40
	Very much	19	24	21

Table 34. Learner Engagement item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Felt prepared for your study	Not at all	2	2	2
	Very little	9	7	8
	Some	33	30	32
	Quite a bit	39	40	40
	Very much	17	20	18
Had a sense of belonging to your university	Not at all	3	4	3
	Very little	12	15	13
	Some	34	33	33
	Quite a bit	36	32	34
	Very much	16	15	16
Participated in discussions online or face-to-face	Never	9	8	9
	Sometimes	37	34	36
	Often	35	35	35
	Very often	20	23	21

Worked with other students as part of your study	Never	7	7	7
	Sometimes	33	29	31
	Often	37	35	36
	Very often	24	30	26
Interacted with students outside study requirements	Never	16	15	16
	Sometimes	38	38	38
	Often	27	28	28
	Very often	18	20	19
Interacted with students who are very different from you	Never	7	8	8
	Sometimes	39	40	39
	Often	36	34	35
	Very often	18	18	18
Been given opportunities to interact with local students	Not at all	4	5	4
	Very little	12	14	13
	Some	29	28	28
	Quite a bit	32	29	31
	Very much	24	25	24

Table 35. Teaching Quality item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Study well structured and focused	Not at all	1	2	1
	Very little	4	7	5
	Some	27	31	29
	Quite a bit	49	45	47
	Very much	19	16	18
Study relevant to education as a whole	Not at all	1	1	1
	Very little	4	5	4
	Some	25	26	26
	Quite a bit	45	43	44
	Very much	26	25	25
Teachers engaged you actively in learning	Not at all	1	2	1
	Very little	7	8	7
	Some	31	31	31
	Quite a bit	43	41	42
	Very much	19	19	19
Teachers demonstrated concern for student learning	Not at all	2	3	2
	Very little	9	10	9
	Some	31	31	31
	Quite a bit	39	37	38
	Very much	19	19	19
Teachers provided clear explanations on coursework and assessment	Not at all	1	2	1
	Very little	7	8	7
	Some	29	30	29
	Quite a bit	43	42	42
	Very much	21	19	20
Teachers stimulated you intellectually	Not at all	1	2	1
	Very little	5	7	6
	Some	26	27	26
	Quite a bit	44	42	43
	Very much	24	23	24
Teachers commented on your work in ways that help you learn	Not at all	3	4	3
	Very little	14	14	14
	Some	34	34	34
	Quite a bit	34	34	34
	Very much	15	15	15
Teachers seemed helpful and approachable	Not at all	1	2	1
	Very little	5	6	5
	Some	24	25	25
	Quite a bit	41	40	41
	Very much	29	28	29

Teachers set assessment tasks that challenge you to learn	Not at all	1	1	1
	Very little	3	5	4
	Some	20	23	21
	Quite a bit	49	46	48
	Very much	28	25	27
Quality of teaching	Poor	2	4	3
	Fair	17	20	18
	Good	54	52	53
	Excellent	27	24	26
Quality of entire educational experience	Poor	2	4	3
	Fair	16	20	18
	Good	56	54	55
	Excellent	25	21	24

Table 36. Student Support item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Experienced efficient enrolment and admissions processes	Not at all	3	4	3
	Very little	7	9	8
	Some	23	24	23
	Quite a bit	35	33	35
	Very much	32	30	31
Induction/ orientation activities relevant and helpful	Not at all	4	8	6
	Very little	13	18	15
	Some	32	32	32
	Quite a bit	31	27	29
	Very much	20	15	19
Received support from university to settle into study	Not at all	4	8	6
	Very little	12	18	14
	Some	32	34	33
	Quite a bit	32	26	29
	Very much	20	14	18
Administrative staff or systems: available	Not at all	2	3	2
	Very little	7	10	8
	Some	28	31	29
	Quite a bit	40	37	39
	Very much	23	20	22
Administrative staff or systems: helpful	Not at all	3	5	4
	Very little	8	12	10
	Some	28	30	29
	Quite a bit	38	34	36
	Very much	23	19	22
Careers advisors: available	Not at all	4	6	5
	Very little	14	17	16
	Some	34	35	34
	Quite a bit	31	28	29
	Very much	16	14	16
Careers advisors: helpful	Not at all	6	9	7
	Very little	13	17	15
	Some	34	33	34
	Quite a bit	30	26	28
	Very much	17	15	17
Academic or learning advisors: available	Not at all	2	2	2
	Very little	7	10	8
	Some	31	31	31
	Quite a bit	39	37	38
	Very much	21	20	21
Academic or learning advisors: helpful	Not at all	2	4	3
	Very little	6	9	7
	Some	28	29	28
	Quite a bit	39	36	38
	Very much	25	23	24

Support services: available	Not at all	4	6	5
	Very little	11	14	12
	Some	31	32	31
	Quite a bit	32	29	31
	Very much	21	20	21
Support services: helpful	Not at all	5	7	6
	Very little	11	13	12
	Some	30	29	30
	Quite a bit	31	29	30
	Very much	23	22	23
Used university services to support study	Not at all	45	47	46
	Very little	20	20	20
	Some	19	18	19
	Quite a bit	10	9	9
	Very much	7	6	6
Offered support relevant to circumstances	Not at all	45	47	46
	Very little	20	20	20
	Some	19	18	19
	Quite a bit	10	9	9
	Very much	7	6	6
Received appropriate English language skill support	Not at all	35	39	36
	Very little	14	15	14
	Some	24	22	23
	Quite a bit	17	14	16
	Very much	11	10	11

Table 37. Learning Resources item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Quality of teaching spaces	Poor	2	4	3
	Fair	12	16	14
	Good	47	48	47
	Excellent	39	32	36
Quality of student spaces and common areas	Poor	5	9	6
	Fair	17	21	19
	Good	44	43	43
	Excellent	34	27	31
Quality of online learning materials	Poor	2	4	3
	Fair	12	15	13
	Good	45	47	46
	Excellent	41	34	39
Quality of computing/IT resources	Poor	3	5	4
	Fair	15	18	16
	Good	47	47	47
	Excellent	35	30	33
Quality of assigned books, notes and resources	Poor	3	4	3
	Fair	17	20	18
	Good	50	50	50
	Excellent	31	25	29
Quality of laboratory or studio equipment	Poor	3	5	4
	Fair	13	17	14
	Good	45	46	45
	Excellent	40	32	37
Quality of library resources and facilities	Poor	2	3	2
	Fair	10	13	11
	Good	42	44	43
	Excellent	47	41	44

Appendix H: Subject area definitions

Table 38. UES subject areas and corresponding ASCED fields of education

Subject area	ASCED field of education[a]
Natural & Physical Sciences	0103 (Physics and Astronomy), 0105 (Chemical Sciences), 0107 (Earth Sciences), 010000 (Natural & Physical Sciences), 019900 (Other Natural & Physical Sciences), 019999 (Natural & Physical Sciences n.e.c.)
Mathematics	0101 (Mathematical Sciences)
Biological Sciences	0109 (Biological Sciences)
Medical Science & Technology	019901 (Medical Science), 019903 (Forensic Science), 019905 (Food Science and Biotechnology), 019907 (Pharmacology), 019909 (Laboratory Technology)
Computing & Information Systems	0201 (Computer Science), 0203 (Information Systems), 0299 (Other IT), 020000 (Information Technology)
Engineering – Other	0301 (Manufacturing Engineering), 0305 (Automotive Engineering), 0311 (Geomatic Engineering - includes Surveying), 0399 (Other Engineering and Related Technologies), 0317 (Maritime Engineering and Technology), 030000 (Engineering and Related Technologies).
Engineering – Process & Resources	030300 (Process & Resources Engineering), 030301 (Chemical Engineering), 030303 (Mining Engineering), 030305 (Materials Engineering), 030307 (Food Processing Technology), 030399 (Process & Resources Engineering n.e.c.)
Engineering – Mechanical	0307 (Mechanical & Industrial Engineering & Technology)
Engineering – Civil	0309 (Civil Engineering)
Engineering – Electrical & Electronic	0313 (Electrical & Electronic Engineering & Technology)
Engineering – Aerospace	0315 (Aerospace Engineering & Technology)
Architecture & Urban Environments	0401 (Architecture & Urban Environment), 040000 (Architecture and Building)
Building & Construction	0403 (Building)
Agriculture & Forestry	0501 (Agriculture), 0503 (Horticulture and Viticulture), 0505 (Forestry Studies), 0507 (Fisheries Studies), 0599 (Other), 050000 (Agriculture, Environmental and Related Studies)
Environmental Studies	0509 (Environmental Studies)
Health Services & Support	0609 (Optical Science), 0615 (Radiography), 061700 (Rehabilitation Therapies), 061705 (Chiropractic & Osteopathy), 061707 (Speech Pathology), 061709 (Audiology), 061711 (Massage Therapy), 061713 (Podiatry), 061799 (Rehabilitation Therapies n.e.c.), 0619 (Complementary Therapies), 0699 (Other Health), 060000 (Health)
Public Health	0613 (Public Health)
Medicine	0601 (Medical Studies)
Nursing	0603 (Nursing)
Pharmacy	0605 (Pharmacy)
Dentistry	0607 (Dental Studies)
Veterinary Science	0611 (Veterinary Studies)
Physiotherapy	061701 (Physiotherapy)
Occupational Therapy	061703 (Occupational Therapy)
Teacher Education – Other	070107-070199 (Teacher-Librarianship through to Teacher Education n.e.c.), 0703 (Curriculum and Education Studies), 0799 (Other), 070100, 070000
Teacher Education – Early Childhood	070101(Teacher Education: Early Childhood)
Teacher Education – Primary & Secondary	070103 (Teacher Education: Primary), 070105 (Teacher Education: Secondary)
Accounting	0801(Accounting)
Business Management	0803 (Business & Management)
Sales & Marketing	0805 (Sales & Marketing)
Management & Commerce – Other	0809 (Office Studies), 0899 (Other Management & Commerce), 080000
Banking & Finance	0811 (Banking & Finance)
Political Science	0901 (Political Science and Policy Studies)
Humanities inc History & Geography	090300 (Studies in Human Society), 090301 (Sociology), 090303 (Anthropology), 090305 (History), 090307 (Archaeology), 090309 (Geography), 090311 (Indigenous Studies), 090313 (Gender Specific Studies), 090399 (Studies in Human Society n.e.c.), 0999 (Other Society and Culture), 0913 (Librarianship, Information Management and Curatorial Studies), 0917 (Philosophy and Religious Studies), 090000 (Society and Culture)
Language & Literature	0915 (Language and Literature)

Social Work	0905 (Includes Social Work and Counselling)
Psychology	0907 (Includes Psychology and Behavioural Science)
Law	0909 (Law)
Justice Studies & Policing	0911(Justice and Law Enforcement)
Economics	0919 (Economics and Econometrics)
Sport & Recreation	092100 (Sport and Recreation), 092101 (Sport and Recreation Activities), 092103 (Sports Coaching), 092199 (Sport and Recreation n.e.c.)
Art & Design	1003 (Visual Arts & Crafts), 1005 (Graphic & Design Studies), 1099 (Other Creative Arts), 100000 (Creative Arts)
Music & Performing Arts	1001 (Includes Music, Dance & Theatre Studies, Dance, Performing Arts n.e.c.)
Communication, Media & Journalism	1007 (Includes Journalism)
Tourism, Hospitality & Personal Services	1101 (Food & Hospitality) and 1103 (Personal Services), 0807 (Tourism), 110000 (Food, Hospitality and Personal Services), 1201 (General Education Programmes), 1203 (Social Skills Programmes), 1205 (Employment Skills Programmes), 1299 (Other Mixed Field Programmes), 120000 (Mixed Field Programmes).

[a] ASCED refers to the Australian Standard Classification of Education, published by the Australian Bureau of Statistics. The code "n.e.c." denotes fields of education not elsewhere classified.

Appendix I. Independent review of the 2013 UES

REPORT TO GRADUATE CAREERS AUSTRALIA ON THE QUALITY REVIEW OF THE UNIVERSITY EXPERIENCE SURVEY, 2013

Dennis Trewin AO, Statistical Consultant

1. Terms of Reference

I was asked to review the 2013 University Experience Survey (UES) from the point of view of its statistical validity. This is the main purpose of this report.

This follows similar reports I did for the 2011 pilot study and 2012 UES.

2. My Qualifications

My main qualification for this review was that I was Australian Statistician from 2000 until 2007. This was a culmination of a long career in official statistics. Much of my early career was in survey methods. I was Director of Statistical Methods at the Australian Bureau of Statistics (ABS) in the late 1970s and have retained that interest since then.

I have formally been accredited as a statistician by the Statistical Society of Australia. I have undertaken a number of statistical reviews since leaving the ABS. For example, I am currently undertaking a quality audit for Statistics Sweden focussing on their twelve most important statistical outputs.

I have been active in international consulting in a number of countries. Over the last 3 years, I have been reviewing the statistical outputs of the Swedish National Statistics Office. One of the problems they are trying to address is the increasing non-response rates in their household surveys.

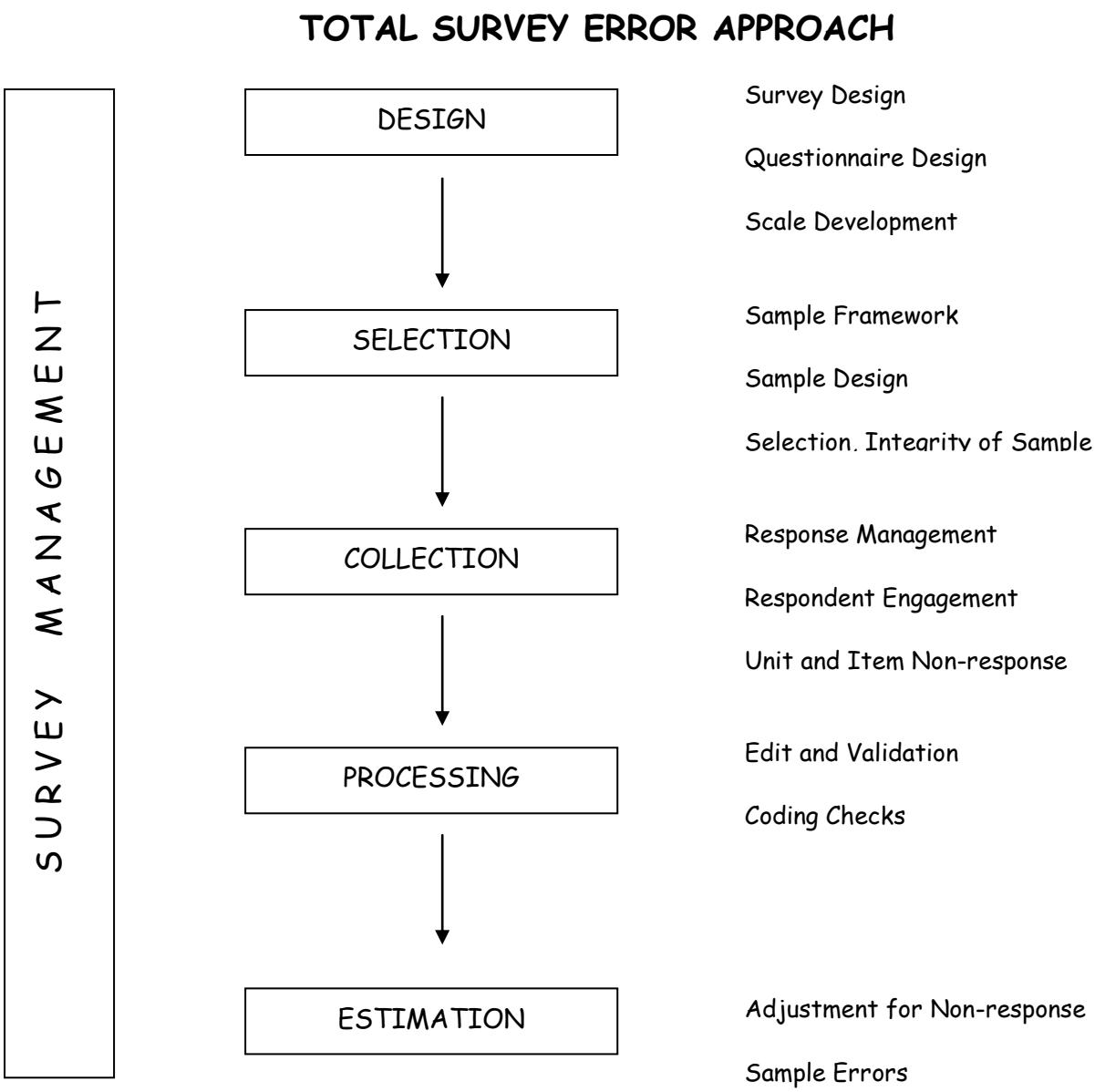
Other relevant external appointments are Past President of the International Statistical Institute, Past President of the International Association of Survey Statisticians, Chairman of the Policy and Advocacy Committee of the Academy of the Social Sciences of Australia, and Associate Commissioner of the Productivity Commission for the Inquiry into the Not-for-Profit Sector.

I have active associations with three Australian Universities. I don't think they represent a conflict of interest. First, I am a Council member at the University of Canberra. I am also Chairman of their Audit and Risk Management Committee. I work as a Statistical Adviser at James Cook University primarily working on a report on the State of the Tropical Regions. At Swinburne University I chair the Advisory Board for the Institute for Social research.

3. My Approach

I have studied the various documents that were provided to me by Graduate Careers Australia (GCA) and The Social Research Centre (SRC). I also visited both Offices on 8 January 2014 where I had an opportunity to ask a range of questions about the survey processes and how effectively they worked in practice. I was also able to obtain additional information relevant to this report including some information of a quantitative nature.

The following framework was used for the preparation of this report. It is based on the business process model used for statistical collections, starting from design through to estimation and analysis based on the actual data collection. The framework has been used to identify the main potential sources of error in the UES. It is the same framework that I used for my report on the 2012 UES.



4. Summary of Findings

I have provided an overview of my assessment here. More details, using the framework above, are provided in the next part of the Report.

A very professional approach has been taken to the design of the UES right through the various design stages. Even though the survey providers are different, it has built on the previous UES. Furthermore, the survey has been undertaken consistently with the design. In particular, significant effort was devoted to maximising response (for a population where response rates are traditionally low) and ensuring the achieved sample was representative. An overall response rate of 29.3% was achieved, whilst less than the target of 35%, this is good for a survey of this type. Opportunities for further improvement should be explored in future surveys.

Important steps taken to improve response and representation were (1) to engage as closely as possible with each University through an especially appointed survey manager, (2) a promotion and marketing campaign, and (3) the use of incentives. These initiatives appear to be successful but no doubt could be improved in light of experience. Steps were also taken to improve the representativeness of the achieved sample especially in the targeting of the reminder action. As a consequence, the sample was reasonably representative in respect of most of the key population characteristics. These include field of study, year of study, mode and type of attendance, Indigenous/Non-Indigenous, Disability and International/Domestic.

The sample was not representative on gender (43% male in the population, 33% in the sample) and the response rates for Universities varied from 17.9% to 50.5% due to, in large part, the extent of student engagement at the University. Furthermore, samples were used in the larger universities so some in-scope students were deliberately excluded on a random basis.

In addition to the steps taken to ensure the representativeness of the sample, weighting procedures can mitigate against non-response bias. They should also be used to adjust for the fact that students from larger universities have a smaller probability of selection than for smaller universities. The question has been asked as to whether to weight the responses or not. I would strongly recommend that weights be used. It is consistent with good practice and it would be most unusual if weights were not used in a survey of this type with differential response rates and different sample fractions across strata.

Why does it matter? If you don't weight, the estimates will be biased towards the characteristics for those that are over-represented. In the UES 2013, the unweighted estimates would be biased towards the responses of females, Universities with low response rates and smaller Universities where no sampling was taking place. Based on the information I was shown:

- Females tend to have slightly higher satisfaction levels than males,
- Not surprisingly, the Universities with lower response rates have lower satisfaction levels as the students are less engaged, and
- There did not appear to be strong association between satisfaction levels and the size of Universities.

In summary, unweighted answers would have a small upward bias. Although the bias may be small at the most aggregated levels, it is preferable not to have this bias and improve the 'face validity' of

the survey to potential critics. The biases at lower levels of aggregation are not clear. Also, although the biases may be small when looking at levels they will become more significant when looking at differences, either over time or between population groups (e.g. fields of study) within the same survey.

The strata were designed to support estimation for the most important domains (Universities, field of study, year of study) and weighting should take place at that level. Separate weights should be used for each stratum based on the effective sample sizes. If this happens, the stratification and weighting procedures mitigate against potential non-response bias from differential response rates at these levels. Furthermore, it is recommended that post-stratification undertaken on the gender variable as the response rate for females is much higher than that for males.

This approach is consistent with those used by the Australian Bureau of Statistics in their population surveys. The strata they use are geographically based but the sample fractions vary from stratum to stratum. Achieved response rates vary by age group and gender so they also post stratify on age group and gender using external benchmarks to overcome potential biases from young adults (especially males) being under-represented for example.

The focus on a representative sample and the use of weighting to population benchmarks to stratum/post-stratum population benchmarks is consistent with international trends for dealing with non-response as higher response rates are becoming harder and harder to obtain especially when face to face interviewing is not used.

One major risk in UES 2013 was that the questionnaire was not tested and evaluated. As SRC recognise, this is not good practice but the timetable they were given did not allow for testing of the questionnaire. Nevertheless, it appears to have worked reasonably well. One test of the adequacy of a questionnaire is the item non-response. If the questions are confusing, respondents will tend not to answer them. The average item non-response was 1.0% which is low compared with most surveys. It is also low compared with the 7.7% achieved in UES 2012. Nevertheless, the timetable for future surveys should allow adequate time for testing.

An important initiative in this survey was the availability of the HEIMS data base as a sample framework. This provides greater assurance that the whole student population is being covered, simplifies the sample selection process, provides the survey taker control over who is included in the survey and prevents the gaming problems that have apparently existed in some past surveys (but not UES). The accuracy of this survey framework would have contributed to the improved response rates. Furthermore, it has enabled centralised data collection by SRC for most Universities. As I commented in my last Report, the response rates in UES 2012 were much higher for those Universities where the data collection was centralised, but this was not replicated in UES 2013 although centralised data collection was used in many more Universities.

A particular issue for the survey has been those students doing double degrees. This is about 20% of students. It has been decided to treat each course undertaken by a student as a unit of analysis. That is, a student undertaking a double degree will appear twice in the data base. Ideally, the weights referred to above should allow for double degrees. However, given the relatively low sensitivity of the estimates to the weights, it should suffice if the stratum weights are based on student numbers.

That is, the weight could be based on the ratio of the number of students in the population in the stratum to the number of students in the achieved sample for the stratum.

Finally, there is some duplication between the UES and other surveys (including those run by the Universities themselves). This imposes additional costs on Universities and burdens on students and must have some influence on response rates. There should be some effort at rationalisation working with the Universities themselves possibly through Universities Australia. Also, the data collected in UES 2013 should allow analysis of the extent of duplication between the instruments used in UES and CEQ.

REPORT AGAINST ERROR FRAMEWORK

1. **Survey Design**

In my previous reports I have pointed out that the survey design is very dependent on the survey objectives and that the accuracy of survey estimates is largely dependent on the size of the sample rather than the fraction of the population that is being sampled. Consequently, the sample size should be much the same for the smallest and largest universities if you want estimates of the same accuracy. These design principles were used and the survey design for UES 2013 is appropriate for the objectives of the survey.

One aspect that was treated differently in UES 2013 was students doing double degrees. This is about 20% of students. It has been decided to treat each course undertaken by a student as a unit of analysis. That is a student undertaking a double degree will appear twice in the data base.

2. **Questionnaire Design**

There was no evaluation and testing of the survey instruments for the UES 2013. I was advised that this was because the timetable imposed on the survey providers did not allow sufficient time for testing. This is not good practice and should be avoided in future surveys or there is the risk the questionnaire will not be consistent with the data items being sought. Nevertheless, the questionnaire appeared to work quite well. One test of the adequacy of a questionnaire is the item non-response. If the questions are confusing, respondents will tend not to answer them. The average item non-response was 1.0% which is low compared with most surveys. It is also low compared with the 7.7% achieved in UES 2012. Nevertheless, the timetable for future surveys should allow adequate time for testing.

Unlike UES 2012, item non-response did not vary much by item except for those students doing a double degree where fatigue might have been a factor. This is a positive sign.

One of the interesting features of the questionnaire design was the use of different orderings of the main question blocks to mitigate against possible ordering effects possibly caused by fatigue towards the end of the questionnaire. It is a sensible precaution. However, there has been no analysis yet of the ordering effects. This feature adds a complication so there should be some analysis of whether ordering effects exist or not. If not, this 'complication' could be removed from future surveys.

3. **Scale Development**

There was no work on scale development for UES 2013. For the UEQ part of the UES, the scales developed by ACER were used. The work done on these scales was professional and so this is a reasonable approach. It also supports consistency between the UES 2012 and UES 2013 surveys and would help analysis of changes between the two years.

For the CEQ part of the UES, I understand the Likert scales were developed many years back using a restricted sample. It may no longer be valid and re-compilation of the scales using current data would be warranted especially to see if there are significant differences with the UES scales. These comparisons are important for consideration of the content of future UES surveys and should be done sooner rather than later so that the key decisions for UES 2014 can be made relatively early.

4. Framework of Students

An important initiative in this survey was the availability of the HEIMS data base as a sample framework. This provides greater assurance that the whole student population is being covered, simplifies the sample selection process, provides the survey taker with greater control over who is included in the survey and prevents the gaming problems that have apparently existed in some past surveys administered by Universities in a decentralised way (but not UES). The availability of an accurate survey framework would have contributed to the improved response rates as non-contacts (e.g. persons no longer studying) would have reduced. Furthermore, it has enabled centralised data collection by SRC at most Universities. As I commented in my last Report, the response rates in UES 2012 were much higher for those Universities where the data collection was centralised.

Algorithms were developed to identify final year students as this was an area of confusion in UES 2012. The algorithm seemed to have worked reasonably well but there should be some evaluation to see whether there is scope for improvement in future surveys.

5. Sample Design

The sample design was straightforward and appropriate for the objectives of the survey.

Stratification was based on a combination of University x Type of course x Year of Study. This was appropriate given that each of these variables is an area of disaggregation of particular interest and the stratification will assist the accuracy of the estimates including allowances for differences in the response rates and sampling fractions.

Where the number of students in a stratum was less than 1333, all students were included in the survey. This was true for most strata. If there were more than 1333 students in a stratum, a sample of 1333 students was chosen randomly. The logic for this cut-off is not entirely clear but it was stated in UES2012 to ensure a sample size of 200 but this seemed to be assuming a relatively low response rate. Presumably, it was based on sample error considerations. This should be revisited prior to the 2014 survey to assess whether such a large sample is required.

Sample errors were estimated assuming the achieved sample was a random sample. This was a reasonable assumption in the circumstances. I have seen the estimated sample errors for selected variables for population groups and they were acceptably low. The sample size is quite large so it is not surprising. At the time of my visit they had not been estimated for individual Universities. I would think there would be interest in Universities knowing their results so I would recommend that this be done. They would also be interested in disaggregation by field of study but, in some cases, the student numbers will be small so some aggregation may be required.

6. Sample Selection and Administration

The sample selection was undertaken by SRC for all Universities based on the HEIMS framework. The student contact details were provided by the Universities. A centralised approach was largely taken for data collection except for a relatively small number of Universities who, for privacy reasons, undertook the data collection themselves. The level of co-operation in these Universities was slightly higher than other Universities although it was too small a sample to read much into this. This was a

different finding to UES 2012 although the number of 'decentralised' Universities was larger in UES 2012.

Ideally, there would be a consistent approach but it is not a high priority as there does not appear to be any adverse impact on student co-operation from decentralised data collection for some Universities.

7. Response Management

The overall response rate is 29.3% less than the target of 35% but significantly higher than UES 2012. A lower response rate will increase sampling errors and increase the risk of non-response bias. I have deliberately used the word risk. For example, if the achieved sample is still representative, there will be no non-response bias. Furthermore, even if some parts of the population are under-represented, there will only be non-response bias if their characteristics tend to be different to the rest of the population. Representativity is a very important objective for surveys that have inherently low response rates like those of student populations. In fact, it can be shown that following up non-respondents that are more typical of current respondents than the majority of non-respondents will reduce the representativeness of the sample. SCR have managed the reminder action (using a targeted approach) in a way that improved the representativity of the sample and should be complimented for that. As a consequence, the achieved sample is reasonably representative of the population. There is some variation in response rates across Universities. However, the main concern is the differential response rates for females and males. The same outcome occurred in UES 2012.

It is becoming more common to compile statistical measures that assess the representativity of the sample on a dynamic basis. One such statistic is known as the R-factor and another is known as the distance function which essentially measures the 'distance' between respondents and non-respondents on certain attributes that are known for both. These types of measure might be considered for future surveys (References: Schouten B, Cobben F and Bethleham J, (2009), "Indicators for the representativeness of survey response", Survey Methodology, 35, 101-113; Sarndal C-E (2007) "The Calibration approach in survey theory and practice", Survey Methodology, 33, 99-119).

Reminders were definitely important in increasing the response rate and sample size. It would be worth analysing whether they have increased the representativeness of the sample. Certainly the higher sample size as a result of reminder action will reduce the size of sampling errors so that is one positive outcome.

8. Respondent Engagement

A significant and impressive effort was put into the promotion and marketing campaign and the engagement of Universities. The key elements were:

- The nomination of a senior survey contact and a survey manager; communication to Universities went through these persons.
- The use of incentives to increase student interest and co-operation.
- Regular feedback to Universities on progress during the data collection phase.
- Targeted non-response follow-up activities.

How did these initiatives work? The improved response rate, and reasonable representativity in the sample, suggests they worked well. I think the appointment of survey managers was an essential step. I have received comments from the sector that the engagement on UES 2013 was far superior to UES 2012. The appointment of survey managers would have been important for this achievement.

As a general comment, the extensive surveying of the student population makes their co-operation more difficult. There appears to be scope for some rationalisation of surveys and this should be examined. If students are surveyed excessively, their co-operation is likely to diminish. Student surveys are initiated by both the Government and the University sectors (including surveys run by Universities of their own students). Rationalisation cannot occur without the agreement of both so efforts at rationalisation need to involve both sectors.

9. Unit and Item Non-response

Significant effort was devoted to maximising response (for a population where response rates are traditionally low and ensuring the achieved sample was representative. An overall response rate of 29.3% was achieved, whilst less than the target of 35%, this is good for a survey of this type. Opportunities for further improvement should be explored in future surveys.

As discussed above, steps were taken to improve response. These appear to have been successful but no doubt could be improved in light of experience. Steps were also taken to improve the representativeness of the achieved sample especially in the targeting of the reminder action. As a consequence, the sample was reasonably representative in respect of most of the key population characteristics. These include field of study, year of study, mode and type of attendance, Indigenous/Non-Indigenous, Disability and International/Domestic.

The sample was not representative on gender (43% male in the population, 33% in the sample) and the response rates for Universities varied from 17.9% to 50.5% representing, in large part, the extent of student engagement at that level.

In addition to the steps taken to ensure the representativeness of the sample, weighting procedures can mitigate against non-response bias. These are discussed below.

The average item non-response was 1.0% which is low compared with most surveys. It is also low compared with the 7.7% achieved in UES 2012. This does not appear to be a significant issue for the reliability of the survey.

10. Edit and validation

I have not looked closely at the procedures except for the information provided in the Methodological Report. Based on the documentation, the procedures used were consistent with good practice.

11. Coding of Open Ended Responses

Most of the questions are self-coding. The only coding required was to code the courses undertaken by students to the fields of study used by UES. This work was undertaken by experienced processing staff who were familiar with this type of coding. Although I couldn't check, SRC were confident this work was done accurately.

There was coding of areas of concern by GCA. I did not obtain sufficient information to comment on this. However, this is a supplementary output rather than one of the main outputs.

12. Estimation, including adjustment of non-response

The sample was not representative on gender (43% male in the population, 33% in the sample) and the response rates for Universities varied from 17.9% to 50.5% representing, in large part, the extent of student engagement. Furthermore, samples were used in the larger universities so some in-scope students were deliberately excluded on a random basis. Adjustments can be made as part of the estimation process.

There has been some discussion of whether weighting should be used or not. Stratum level weighting should be used and this will adjust for differential non-response across the strata (university, field of study, first/final year students). It will also allow for differential sampling fractions across strata. It will not adjust for any differences between respondents and non-respondents within strata. Post-stratification may assist in this respect. This involves weighting respondents within a stratum differently according to their characteristics with respect to potential post-stratification variables.

To warrant the use of post-stratification to reduce possible non-response bias, there has to be both a differential non-response rate for the categories (within a post-stratification variable such as gender) AND the survey characteristics for these categories have to be different. For example, the response rate for females was much higher than that for males. If the characteristics of females were different to those for males, the use of post-stratification would reduce non-response bias.

There is another reason for applying post-stratification. If it is important for the estimates of particular categories (e.g. females) across strata to add to population totals, post-stratification can be used to affect this. For example, post-stratification is used in the ABS Labour Force Survey to force estimates to add to independent benchmarks of population disaggregated by State/Territory, age group and gender. A similar approach could be used in UES for estimates disaggregated by gender. Without the use of post-stratification, females would be over-estimated and males would be underestimated.

Having looked at the potential post–stratification variables that might be considered for UES, only gender is worth considering. As mentioned above there were differential response rates across the two gender categories. However, for the other potential post-stratification variables, there was little difference in response rates across the categories so there would be no gains from using these variables for post-stratification purposes.

Although I cannot be certain, I think the steps taken with the survey design to improve the representativeness of the sample, and the use of stratification and post-stratification with separate weighting of strata/post-strata, should mean that non-response bias is low. Why do I say this? For a CEQ study in 2005, analysis was undertaken by ACER which showed the most important determinants of student 'satisfaction' scales were the variables used for stratification in UES and gender. The most important variable for the CEQ was field of study and this might be the case for UES as well. It was used as a stratification variable.

With respect to the question of whether to weight the responses or not, I would strongly recommend that weights be used. It is consistent with good practice and it would be most unusual if weights were not used in a survey of this type with differential response rates.

What would be the impact if you didn't weight? If you don't weight, the estimates will be biased towards the estimates for those that are over-represented. In the UES 2013, the unweighted estimates would be biased towards the responses of females, Universities with low response rates and smaller Universities where there was sampling taking place. Based on the information I was shown:

- Females tend to have slightly higher satisfaction levels than males,
- Not surprisingly, the Universities with lower response rates have lower satisfaction levels as the students are less engaged, and
- There did not appear to be strong association between satisfaction levels and the size of Universities.

In summary, unweighted answers would have a small upward bias. Although the bias may be small at most aggregated levels, it is preferable not to have this bias and improve the 'face validity' of the survey to potential critics. Also, although the biases may be small when looking at levels they will become more significant when looking at differences, either over time or between population groups (e.g. fields of study) within the same survey.

The focus on a representative sample and the use of weighting to population benchmarks to stratum/post-stratum population benchmarks is consistent with international trends for dealing with non-response as higher response rates are becoming harder and harder to obtain especially when face to face interviewing is not used.

13. Sampling errors

Details about sampling errors should be estimated and presented to assist with analysis of report. This should be based on actual data. These do not have to be calculated for every estimate – only sufficient estimates to provide readers with a feeling for the size of the sampling errors. However, they should be estimated for each level of estimation – e.g. total population, field of study, University.

Strictly speaking the sample is not a random sample which is an important assumption for most estimates of sampling errors. The students in the sample have self-selected to the extent they have agreed to respond to the UES. However, for the purposes of estimating sample errors, I believe it is OK to assume the sample is random. In fact, there is no other realistic assumption.

14. Modelling

Another question that might be asked is the impact of non-response on the modelling that is undertaken to estimate the scales. The answer is that the impact should be negligible. If the model is valid, it should apply to both respondents and non-respondents. Therefore, a model based on respondents only should still be representative of the whole population.

Another question that might be asked is whether to use weighted or unweighted data when modelling. The answer is that it does not matter greatly. If the underlying model is correct, models estimated on either weighted or unweighted data should both be unbiased. However, if weighted data is available I suggest you use it as studies have shown this will provide more accurate estimates of the parameters of the model but there is not a lot in it.

15. Publication

A criterion has been established for determining what data should be published. I understand only cells with 25 or more will be published. I support having the criterion established in advance.

I understand the logic behind this criterion is to protect confidentiality. This is probably too conservative if that is the sole reason (the ABS has a cut-off of 3 persons). However, estimates based on this small a sample will have relatively high sampling errors so it may be a reasonable cut-off when you consider both these factors.

There is an alternative approach that might be considered. I will illustrate by some analysis I did for UES 2012. According to the criteria being proposed in UES 2012 (which are different to those being proposed for UES 2013), (1) data for all Universities would have been published, (2) data for Universities by year of study except for one combination would be published, (3) data for about 50% of fields of study by Universities would be published, (4) and 20% of Universities by field of study and year of study would be published.

I suggested that all data for Tables (1), (2) and (3) at least be published together with information on sample errors so users can assess the reliability of the data for their purposes. Furthermore, data in these tables with high sampling errors (e.g. standard errors of 20% or higher) could be marked with an asterisk to highlight the high sampling errors. This is the ABS practice rather than suppressing cells with high sampling errors. That is, a 'user beware' approach be adopted rather than suppressing those cells with high sampling errors. I am not sure if this approach was used in UES 2012 or not but I recommend a similar approach in UES 2013.

Of course, any cells that are confidential should be suppressed.

The publication should also provide readers some information to enable them to assess the accuracy of the survey for their purposes. This would include sampling errors. Furthermore, there should be a description of the more significant non-sampling errors and a discussion of the risks they pose to use of the estimates. Quantitative data should be provided wherever possible.

Information on the design of the survey, survey variables and other meta data should be published especially to assist the more informed reader. I have seen a draft of the Data Dictionary. It is a good quality document that will considerably assist users of the UES results.

I understand there is interest in a microdata release especially among the Universities. It is now a common output for surveys conducted by the ABS and AIHW for example. If this is being considered seriously, it would be worth getting an independent expert to provide advice on confidentiality practices.

The Universities are a key stakeholder. The success of the survey depends on their co-operation. This is more likely if they are significant user of the UES itself and also possibly use it as a vehicle to ask for information of particular interest to them. The Department should work on this possibly in collaboration with Universities Australia.

Conclusions

The survey design was appropriate for satisfying the objectives of the survey. The main risks to quality are sampling errors and non-response bias both of which have been discussed in detail above.

The sample size is sufficiently large to enable most of the large aggregates to be published. Also, many of the detailed statistics can be published although some will be subject to high sampling error.

The relatively low response rate of 29.3% leaves open the potential for non-response bias to be an important influence. However, reasonable steps have been taken at the data collection stage to mitigate the impact of non-response bias. If weighting, as described in this document, is used in estimation the residual impact should not be large for the major aggregates. It may be more significant in relative terms for smaller aggregates so care should be taken in interpreting these estimates.

JANUARY 2014