2014 University Experience Survey National Report

March 2015

ISBN 978-1-74361-508-9 (PDF)

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 2014 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 Department of Education, which funded this work and provided support throughout the project. Particular thanks go to Phil Aungles, Sam Pietsch, Dr Andrew Taylor, Gabrielle Hodgson and Ben McBrien.

Dennis Trewin (Dennis Trewin Statistical Consulting) provided an independent technical review of the 2013 UES project. This review is presented in Appendix J.

The UES Project Director was Dr Noel Edge (GCA). The Project Team consisted of Graeme Bryant (GCA), David Carroll (GCA), Bruce Guthrie (GCA), Daniela Iarossi (SRC), Rebecca Lam (SRC), Darren Matthews (GCA), Dianna McDonald (SRC), Eric Skuja (SRC), Daniel Smith (SRC), Sonia Whiteley (SRC) and Alistair Wilcox (SRC). Andrew Ward (Senior Statistician, SRC) conducted and reported on the psychometric analyses of the University Experience Questionnaire presented in Appendices B and D.

For more information on the conduct and results of the 2014 UES, please contact the UES Consortium at ues@graduatecareers.edu.au.

Executive summary

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 commencing and later-year undergraduate students. The UES was developed in 2011 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 and 2014.

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).

The core features of the 2013 UES were retained for the 2014 survey, including a centralised approach to sampling using data from the Higher Education Information Management System (HEIMS) and online-only data collection. There were, however, several differences between the 2013 and 2014 administrations. First, while the approach to sampling in 2013 involved drawing a fixed number of students from large population strata, required sample sizes in 2014 were calculated taking into account the number of records available and response rates from the 2013 UES. Second, one questionnaire item from the student support focus area was removed on the basis of psychometric evidence showing lack of fit. Third, module rotation was removed on the basis of analysis conducted in 2013 that showed no substantial variation between questionnaire sequences (Section 1.2).

The 2014 UES project included a pilot of the UES questionnaire and methodology on students of 15 volunteer non-university higher education institutions (NUHEIs) under the name Student Experience Survey (SES). The pilot administration was a success, achieving a response rate of 47.9 per cent (Section 1.2.1). Because the SES was a small-scale pilot study, only high-level aggregate results for NUHEI students are presented in this report. NUHEI students were generally more likely to be satisfied with their educational experience than university students. For example, 85 per cent of NUHEI students indicated satisfaction with their entire educational experience, compared with 81 per cent of university students; however it is important to note that these results do not account for any differences in student characteristics between NUHEI and universities (Section 4.7). An analysis of the psychometric properties of the University Experience Questionnaire (UEQ) as administered to NUHEI students is presented in Appendix D.

As in 2013, the results of the 2014 UES are reported as the percentage of surveyed students who expressed satisfaction with their higher education experience. One outcome of this is that the results presented in the 2013 and 2014 UES reports are not directly comparable to those presented in the 2012 report (Section 1.3).

As was the case in 2012 and 2013, all 40 Australian universities participated in the 2014 UES. The fieldwork period ran from August to October 2014 (Section 2.1). The “in-scope” population consisted of 330,772 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 between five and ten 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. An engagement campaign was conducted in the lead-up to and throughout the fieldwork period to build awareness of the UES and encourage participation (Section 2.6).

The response rate for the 2014 UES was 30.1 per cent, up from 29.3 per cent in 2013. Institutional response rates ranged from 50.4 per cent to 20.9 per cent. Responses were received from 99,112 students, which equated to 108,322 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 a substantial impact on the results at the national level, so it was decided to analyse the data without applying weights (Section 3.3). Stratum-level sampling and response maximisation adopted for the 2014 UES resulted in a general increase in the number of strata that achieved the desired level of statistical precision (Section 3.4).

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 direct 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	79	60	84	76	88	83	83
Later year	85	63	79	68	81	77	77
Total	**81**	**61**	**82**	**73**	**86**	**81**	**81**

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 15 percentage points separating the subject areas with the highest and lowest results. The widest range was observed in relation to learner engagement, with 30 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 2013 and 2014 UES collections are compared (see table below), the largest difference in terms of focus area results was seen in relation to student support, with 14 percentage points separating 2013 and 2014. A year-on-year difference of this magnitude is most likely the result of the aforementioned changes to the questionnaire and survey method in 2014 (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**
2013	79	57	79	59	83	79	79
2014	81	61	82	73	86	81	81

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, 81 per cent of students indicated satisfaction with both the quality of teaching and their entire educational experience. Many of the lowest results were associated with the student support and learner engagement focus areas (Section 4.3).

When institutional percentage satisfied results on the quality of teaching and the entire educational experience 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 later-year students. A similar result is obtained when later-year respondents to the UES are compared with final-year students who responded to the National Student Survey (NSS) in the UK. It is important to note, however, that these results do not account for potential differences in the composition of the respective student populations, nor methodological differences between the 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 2014. Overall, 17 per cent indicated that they had considered leaving. Commencing students, older students, Aboriginal and Torres Strait Islander students, students with a disability and students 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. The most common reasons given for considering early departure are situational in nature, including health or stress, difficulties relating to finances and workload, and study/life balance (Section 4.6).

Contents

List of tables

List of figures

1 Introduction and overview

1.1 Background to 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.

Specifically, the UES was designed to measure five facets of the higher education student experience: Skills Development, Learner Engagement, Teaching Quality, Student Support and Learning Resources. These are operationalised by means of summated rating scales, underpinned by 46 individual questionnaire items. These items are 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 UEQ items is presented in Appendix A.

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, which involved conducting a pilot of the UES on a sample from 24 universities. From this they made 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 a mixed-mode online-telephone survey.

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. The survey instrument was largely unchanged from 2012, with the major difference between the 2012 and 2013 collections being 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.

The approach to sampling and data collection for 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. Moreover, while the 2012 UES had been administered as a mixed-mode survey, the 2013 UES was administered entirely online. Compared to a mixed-mode online-telephone survey, an online-only data collection methodology is more cost effective, convenient for students, requires a shorter fieldwork period and mitigates potential problematic survey mode effects.

[1] Throughout this report, "course" is synonymous with "degree program".

1.2 The 2014 UES

The core features of the 2013 UES were retained for the 2014 survey, including a centralised approach to sampling based on HEIMS data (see Section 2.3) and online-only data collection. Table 1 presents an overview of the 2013 and 2014 UES collections. The in-scope population definition was unchanged from previous implementations of the survey and consisted of commencing and later-year onshore undergraduate students (see Section 2.2). The in-scope population size was smaller in 2014 than 2013 due to relatively higher proportions in the overall population of students in postgraduate and non-award courses, offshore students and students in the middle years of their courses. Also, a slightly higher proportion of students were excluded by institutions during the sample verification process in 2014 compared with the previous year.

Table 1. UES project overview, 2013 and 2014

Project element	2013	2014
Number of participating institutions	40	40
Number of "in-scope" students	341,343	330,772
Data collection period	August-November	August-October
Data collection mode	Online	Online
Overall response rate (%)	29.3	30.1
Number of completed surveys (student level)	100,225	99,112
Number of valid surveys (course level)	108,940	108,322
Analytic unit	Course	Course

The approach to sampling changed markedly in 2014. In the 2012 and 2013 UES collections, a fixed number of students were sampled from large population strata (> 1,333 students), which often led to oversampling. In 2014, required sample sizes were calculated at the stratum level, taking into account the number of records available and response rates from the 2013 UES. As a result, although marginally fewer surveys were completed in 2014 than 2013, the number of strata meeting desired precision targets increased considerably (see Section 3.4).

There were only two major questionnaire changes between 2013 and 2014. First was the removal of the item "At university during year x, to what extent have you used university services to support your study?" on the basis of psychometric evidence showing lack of fit within the student support focus area.[2] To enable valid comparisons between the 2013 and 2014 UES collections, the 2013 results have been recalculated without this questionnaire item. Consequently, the 2013 results for the student support focus area presented in this report will differ from those published in the 2013 report. An analysis of the psychometric properties of the revised student support focus area is presented in Appendix B. The key findings from this analysis are, firstly, that there remains one poorly fitting questionnaire item within the student support focus area (relating to English language support), and secondly, that a number of items in this focus area display a lack of applicability to students.

Second was the removal of module rotation in 2014. Unlike 2013, in which five module rotations were presented to students, only one questionnaire sequence was employed for the 2014 UES (see Section 2.5).

[2] Refer to Graduate Careers Australia and the Social Research Centre. (2014). *2014 University Experience Survey National Report*. Retrieved 4 Feb., 2015, from https://docs.education.gov.au/system/files/doc/other/ues13_report.pdf

As part of the 2013 UES, six scales from the Course Experience Questionnaire component of the Australian Graduate Survey (AGS) were administered on a trial basis to students from 14 institutions. This trial resulted in a recommendation that the Good Teaching Scale (GTS), Generic Skills Scale (GSS), Clear Goals and Standards Scale (CGS) and Overall Satisfaction Item (OSI) be administered to a sample of later-year students across all participating institutions to facilitate international benchmarking. It was further recommended that the CEQ scales should only be presented to a small sample of students of a sufficient size to yield national-level estimates that are precise within ±7.5 percentage points of the true population value at a 90 per cent confidence level. As with the UEQ, sampled students in double degrees were provided with the opportunity to complete the CEQ for each course element individually. A list of CEQ items administered in the 2014 UES is presented in Appendix C.

As was the case in previous implementations, 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,[3] resulting in a clear demarcation between the two.

1.2.1 The Student Experience Survey

The 2014 UES project included a pilot of the UES questionnaire and methodology on students of 15 volunteer non-university higher education institutions (NUHEIs). These are listed in Table 2. It was recognised early in the process that the name "University Experience Survey" could discourage participation by students of NUHEIs, so the survey was retitled the Student Experience Survey (SES) for this population. Where required, questionnaire item wordings were similarly modified. These differences aside, the administration of the SES was essentially identical to the UES. A total of 1,444 completed surveys were returned from a sample of 3,039 in-scope students, representing a strong response rate of 47.9 per cent (cf. 30.1 per cent for the UES).

Table 2. Institutions that participated in the SES pilot

Alphacrucis College	Polytechnic West
Avondale College of Higher Education	Raffles College of Design and Commerce
Blue Mountains International Hotel Management School	Sydney College of Divinity
Christian Heritage College	Tabor Adelaide
Holmesglen Institute of TAFE	Tabor College NSW
Marcus Oldham	Tabor College Tasmania
Melbourne Institute of Technology	Tabor College Victoria
Melbourne Polytechnic (formerly NMIT)	

Because the 2014 SES was a small-scale pilot study involving a convenience sample of 15 of the approximately 130 NUHEIs currently operating in Australia, detailed results for NUHEI students are not presented in the body of this report as they may not necessarily be representative of the non-university higher education sector as a whole. High-level aggregate SES results are presented in Section 4.7.

An analysis of the psychometric properties of the UEQ as administered to NUHEI students is presented in Appendix D. The main purposes of this analysis are to investigate how well the UEQ

[3] "UEQ" is used where necessary in this report to distinguish between the UES questionnaire proper and additional non-standard elements, including the CEQ and institution-specific items.

items fit within their respective focus areas when administered to NUHEI students, and whether the two cohorts respond to the questionnaire differently, with the aim of determining whether the current UEQ is an appropriate instrument for measuring NUHEI students' engagement and satisfaction with their higher education experience. The results of the analysis indicate that most items demonstrate good fit within their respective focus areas, and most items are answered similarly by university and NUHEI respondents. A number of items, especially in the student support focus area, did not apply or were not answered by large proportions of NUHEI and university students alike, echoing the findings of the previously discussed analysis of the student support focus area (see Appendix B).

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 the results presented in the 2013 and 2014 UES reports 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 E.

1.4 Important caveats

While the UES has now been trialled and administered three times 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 2014 UES. Similar analysis undertaken for the 2013 UES report yielded essentially identical results. As such, all results presented in this report are based on unweighted data unless otherwise noted.

It is important to consider that bias on the basis of unobservable characteristics may still be influencing the results—in any case this type of bias may not be corrected by the application of post-stratification weights. It is difficult to say whether any such bias exists without gathering data from non-respondents, which was not undertaken as part of the 2014 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 explicitly measure the relative importance that students place on different aspects of their higher education 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 2014 UES, including the conduct and administration of the survey, and key results based on the national UES data file, which consists of 108,322 responses from 99,112 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 90 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. 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 2014 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 13 operational cohorts; the largest comprising five institutions and the smallest comprising single institutions. Table 3 lists the institutions in each cohort, along with corresponding fieldwork commencement and completion dates.

Table 3. Operational cohorts for the 2014 UES

Operational cohort	Institution	Commencement date	Completion date
Cohort 1a	Deakin University	4 August	8 September
	Griffith University		9 September
	University of Melbourne		10 September
	James Cook University		11 September
Cohort 1b	University of Southern Queensland	5 August	11 September
	University of South Australia		17 September
	University of Western Sydney		17 September
	University of Divinity		17 September
Cohort 1c	The University of Adelaide	7 August	11 September
Cohort 2a	Charles Sturt University	11 August	19 September
	The Australian National University		19 September
	Edith Cowan University		26 September
	Macquarie University		30 September
	Murdoch University		1 October
Cohort 2b	Federation University	12 August	29 September
	University of the Sunshine Coast		30 September
	University of Wollongong		1 October
	University of Canberra		3 October
Cohort 2c	University of Sydney	13 August	26 September
	University of Queensland		2 October
	University of Notre Dame Australia	14 August	6 October
Cohort 3a	Central Queensland University	18 August	26 September
	Southern Cross University		10 October
	Charles Darwin University		15 October
	Australian Catholic University		10 October
Cohort 3b	Monash University	19 August	19 September
	Queensland University of Technology		30 September
	University of Tasmania		6 October
	University of Newcastle		14 October
	Flinders University		15 October
Cohort 3c	Victoria University	20 August	6 October
	La Trobe University		10 October
Cohort 3d	The University of New South Wales	21 August	30 September
	The University of New England		2 October
	The University of Western Australia		3 October
Cohort 4a	University of Technology, Sydney	1 September	7 October
	RMIT University		22 September
	Curtin University of Technology		3 October
Cohort 4b	Bond University	2 September	9 October
Cohort 4c	Swinburne University of Technology	1 October	27 October

2.2 Survey population

The in-scope survey population for the 2014 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 330,772 students. The definitions used for commencing and later-year students in the 2014 UES are essentially unchanged from 2013. These are described 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 verify, 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. In 2013, the complexity of identifying later-year students was acknowledged 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 (E339) to the total EFTSL for the course (E350) proved the better option for 9 of the 40 institutions in 2013.[4] In 2014 this was re-evaluated and two institutions were changed to Option 2, for a total of seven institutions using this solution.

Option 2, which adjusts for attendance mode (E330) and course duration (E350) was the standard solution used for the remaining 33 institutions. This solution may under-sample if large numbers of students engage in accelerated progression during Semester 2 or Summer Term. As such, specific adjustments were required to accommodate the idiosyncrasies of a small number of 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 verify, where possible, that the selected students were still enrolled.

2.3 Sampling design

2.3.1 Sample frame

As with 2013, the sample frame for the 2014 UES was based on a "top-down" approach using population data from HEIMS to create the sample frames for individual universities. Compared with

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

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 and 2014 UES implementations reduces the likelihood of accidental bias being introduced to the sample selection process and ensures 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.

2.3.2 Approach to sampling

As in previous implementations, a stratified sampling approach was employed for the 2014 UES collection, with strata defined on the basis of institution and subject area;[5] however the approach to determining sample size differed markedly compared with the earlier surveys. In the 2013 UES, the approach to sampling was broadly consistent with that of the 2012 survey. The number of students for each stratum was calculated using the approach described in the 2012 UES National Report.[6] 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 students was drawn in the hope that this would yield at least 200 responses. According to the report, this value was derived from a desire for error bands of ±5 per cent at a 95 per cent level of confidence.[7]

An analysis of this approach suggested that it had a number of shortcomings. In general, large strata were substantially oversampled and often achieved completed surveys well in excess of the target of 200, with the result that students from large strata were substantially over-represented. This had the flow-on effect of increasing the gender imbalance in the sample of secured responses, as many of the large strata consisted of course offerings where males are traditionally underrepresented, such as nursing and education. Lastly, the sampling approach did not take into consideration the differential response rates across strata.

In 2014, required sample sizes were calculated at the stratum level taking into account the number of records available and the goal of reporting stratum-level results at a level of precision of ±7.5 percentage points at a 90 per cent level of confidence.[8] In order to establish the required sample sizes, a target number of completed surveys was calculated for each stratum in order to achieve the desired level of precision. The number of students to be sampled from each stratum to achieve this target was estimated using the response rate for that stratum from the 2013 UES, or the overall response rate for the institution if no stratum-level response rate was available (i.e. no in-scope students fell into the stratum in the 2013 collection).

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 I.
[6] Radloff, A., Coates, H., Taylor, R., James, R. & Krause, K. (2012). *2012 University Experience Survey National Report*. Retrieved 15 Dec., 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.
[8] The original precision target was ±5 percentage points at a 90 per cent level of confidence; however it became apparent that, when the required sample sizes were compared with the response rates achieved in 2013, it would not be possible to achieve the required number of responses for a substantial proportion of the strata.

2.4 Additional questionnaire elements and populations

2.4.1 Institution-specific items

As with previous implementations, institutions were offered the option of including non-standard, institution-specific items as part of the 2014 UES. Fifteen institutions chose to do so, up from three institutions who added institution-specific items to their 2013 UES. Frequent inclusions were the Workplace Relevance Scale, originally developed for (but not incorporated into) the CEQ, and an item to monitor students at risk of discontinuing their studies.

2.4.2 Additional populations

Institutions were given the opportunity to add additional populations to the UES. Ten institutions surveyed 13 additional populations, including postgraduate, middle-year, offshore and enabling students. Responses from students in these populations are not included in the national data file and therefore do not appear in any of the results presented in this report.

2.5 Online survey

As was the case in 2013, the 2014 UES was administered entirely online. The 2012 UES was administered as a mixed-mode online-telephone survey. The move to a single-mode survey in 2013 was motivated by a desire for a cost-effective methodology that would still achieve the necessary response whilst mitigating 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 to complete the survey.

The UES items were organised into a modular structure for ease of administration. Unlike the 2013 survey, in which different module rotations were presented to students to counteract potential order effects, only one questionnaire sequence was used for the 2014 UES on the basis of analysis conducted in 2013 that showed no substantial variation between rotations.[9]

2.6 Student engagement strategy

GCA-SRC designed and disseminated a range of promotional methods and materials to build UES awareness in the higher education sector and encourage participation amongst the student population. There were two main phases of student engagement. 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.

[9] Version A of the rotations was selected as it had the most logical flow of modules and showed negligible non-response in 2013.

The following media formed part of the national engagement strategy:

- The standalone UES website (www.ues.edu.au), which provided information about the survey, including FAQs and results from the 2013 UES.
- Website tiles, which were designed to be placed on institutional websites, learning management systems, news feeds, student association websites, etc. Example website tiles are presented in Appendix F.
- 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 invitation and between five and ten email reminders, depending on length of time in field and response rate. The majority of institutions opted to allow GCA-SRC to approach students directly via email, while two 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 were removed from each email reminder sample file prior to the email reminders being sent.

Seventeen institutions provided mobile telephone numbers to GCA-SRC to facilitate SMS follow-up and one institution sent SMS internally. SMS follow-up was primarily used as a means of alerting students to the impending closure of the prize draw or an alert to the survey period closing.

A hardcopy letter was sent to non-responding students (for whom a postal address was available) after the initial email invitation. The letter was timed to arrive prior to prize draw cut-off, typically within the first two weeks of the fieldwork period. A small number of institutions were selected to receive a second letter, based on two criteria: a low response to online correspondence based on learnings from the 2013 UES, or lower than expected results based on 2013 UES outcomes.

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.

A key focus of the 2014 UES was working collaboratively with institutions, wherever possible, to maximise participation rates in the survey. Feedback was sought from institutional Survey Managers regarding the timing and targeting of follow-up and reminder activity. Moreover, many institutions undertook supplementary activities to promote the 2014 UES and encourage student participation. The most commonly employed methods were notifications on learning management systems, emails from the Vice-Chancellor, social media posts, institutional websites and internal staff emails.

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 2014 UES.

Throughout June and early July 2014, GCA-SRC contacted Vice-Chancellors at the Table A and B institutions and invited them to participate in the 2014 UES. Concurrently, an introductory email was sent to institutional survey staff to make them aware of the survey. This allowed survey staff to prepare for their anticipated involvement in the project and provided the opportunity for early discussions on privacy policies, additional items and populations, and their ability to prepare population and sample details. A UES administration guide was also prepared and sent to institutions to assist them in undertaking the UES at their institution. Promotional materials required to facilitate the pre-survey engagement were distributed to institutional contacts prior to the commencement of data collection fieldwork. Moreover, the second day of GCA's annual two-day Survey Manager Information Forum (SMIF), held in mid-July, was dedicated to the 2014 UES.

Regular updates were provided to institutions throughout the data collection phase of the project. As was the case in 2013, the online survey included real-time reporting functionality, which allowed institutions to monitor the progress of data collection and engage with the project management team regarding targeted non-response follow-up activities, as appropriate.

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. The analytic unit for the 2013 and 2014 UES implementations is the course. In the 2013 data set and again in 2014, 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 99,112 students who completed the 2014 UES, 9,210 (9.3 per cent) furnished a valid response for their second course element, resulting in 108,322 valid responses.

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

While the overall institutional response rate remains an important measure of survey administration effectiveness, there was a shift in the 2014 UES from overall response rates to stratum-level response rates. Institutions were given targets for each subject area and encouraged to promote student engagement and participation at this level (see Section 2.3.2).

The overall response rates achieved for each institution in 2014 are presented in Table 4, along with the response rate achieved in 2013. In spite of the more challenging response rate targets, more than half of the institutions improved upon their response rate from 2013, yielding a national response rate of 30.1 per cent (up from 29.3 per cent in 2013).

Table 4. UES response rates, 2013 and 2014

University	2013	2014	University	2013	2014
University of Divinity	50.5	50.4	Deakin University	29.2	30.1
Bond University	32.8	42.8	The University of Melbourne	34.5	29.6
Central Queensland University	36.0	38.6	The University of Sydney	30.3	29.6
The University of Queensland	32.5	38.6	Macquarie University	26.3	29.5
The University of Adelaide	41.4	38.4	Federation University Australia	22.1	29.3
Charles Darwin University	40.5	37.3	University of Wollongong	23.5	29.3
University of the Sunshine Coast	29.2	37.3	Curtin University	26.1	28.1
The University of New England[a]	32.9	37.0	University of Canberra	24.4	27.8
Monash University	39.7	36.9	The University of New South Wales	27.0	27.7
James Cook University	29.0	36.5	The University of Notre Dame Australia[a]	26.0	27.1
University of Tasmania	33.0	35.7	Griffith University	23.5	26.8
Murdoch University	30.6	35.6	Victoria University	17.9	26.8
Charles Sturt University	32.3	35.4	La Trobe University	33.0	26.7
University of Southern Queensland	25.2	35.0	University of Technology, Sydney	28.2	25.7
The Australian National University	29.3	33.5	Queensland University of Technology	29.4	25.0
Edith Cowan University	29.3	33.4	RMIT University	20.8	25.0
Flinders University	35.2	32.9	University of Western Sydney	26.6	24.2
Southern Cross University	24.4	32.4	Swinburne University of Technology	25.5	22.6
The University of Western Australia	39.7	30.8	Australian Catholic University	23.7	20.9
University of South Australia	25.2	30.8			
The University of Newcastle	34.0	30.3	**Total**	**29.3**	**30.1**

[a] Institution sent email invitations to their own students in 2014.

In some cases where institutions achieved lower response rates in 2014 compared with 2013, this is a direct result of the revised sampling approach, which aimed to reduce oversampling of students from the larger strata. In total the 2014 UES achieved 99,112 completed surveys, fewer than the 100,225 achieved in 2013 due to the revised sampling approach.

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 5 (subgroup) and Table 6 (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, especially stage of studies, indigenous status, disability status and study mode.

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 and 2013, 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. Students who were the first in their family to attend university were somewhat under-represented in the sample of secured responses, but not to the same degree as gender.

Table 5. 2014 UES response characteristics and population parameters by subgroup

Group	Subgroup	UES sample		In-scope population	
		n	%	n	%
Stage of studies	Commencing	66,474	61.4	239,631	61.4
	Later year	41,848	38.6	150,418	38.6
Gender	Male	36,498	33.7	167,592	43.0
	Female	71,824	66.3	222,457	57.0
Indigenous	Aboriginal or Torres Strait Islander	1,264	1.2	4,660	1.2
	Not Aboriginal or Torres Strait Islander	105,871	98.8	381,136	98.8
Home language	English	81,557	78.2	288,988	76.5
	Other	22,671	21.8	88,599	23.5
Disability	Disability reported	5,657	5.2	16,733	4.3
	No disability reported	102,665	94.8	373,316	95.7
Study mode	Internal	93,029	85.9	342,544	87.8
	External/multi-modal	15,293	14.1	47,505	12.2
International	Domestic student	96,994	89.5	342,872	87.9
	International student	11,328	10.5	47,177	12.1
First in family	First in family	28,547	47.9	112,663	53.0
	Not first in family	31,022	52.1	99,961	47.0
Total		**108,322**	**100**	**390,049**	**100**

The sample also closely matches the in-scope population in terms of subject area (see Table 6). The largest difference between the sample and population was observed in relation to the business and management subject area (2.7 percentage points), which was the only difference 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 (10.3 per cent), business management (7.5 per cent), nursing (7.0 per cent), natural and physical sciences (6.4 per cent), and health services and support (5.4 per cent). These five subject areas together constitute more than a third of the entire sample.

Table 6. 2014 UES response characteristics and population parameters by subject area

Group	Subgroup	UES sample		In-scope population	
		n	%	n	%
Natural and Physical Sciences	Natural & Physical Sciences	6,986	6.4	27,694	7.1
	Mathematics	414	0.4	1,092	0.3
	Biological Sciences	1,993	1.8	5,561	1.4
	Medical Science & Technology	3,024	2.8	8,417	2.2
IT	Computing & Information Systems	3,390	3.1	12,470	3.2
Engineering and Related Technologies	Engineering – Other	3,771	3.5	14,289	3.7
	Engineering – Process & Resources	572	0.5	1,930	0.5

Table 6. (continued)

Group	Subgroup	UES sample		In-scope population	
		n	%	n	%
	Engineering – Mechanical	707	0.7	2,466	0.6
	Engineering – Civil	947	0.9	3,452	0.9
	Engineering – Electrical & Electronic	814	0.8	2,582	0.7
	Engineering – Aerospace	477	0.4	1,454	0.4
Architecture and Building	Architecture & Urban Environments	1,984	1.8	7,074	1.8
	Building & Construction	491	0.5	2,524	0.6
Agriculture and Environmental Studies	Agriculture & Forestry	680	0.6	1,820	0.5
	Environmental Studies	1,113	1.0	3,271	0.8
Health	Health Services & Support	5,861	5.4	22,064	5.7
	Public Health	1,311	1.2	4,170	1.1
	Medicine	1,988	1.8	5,821	1.5
	Nursing	7,621	7.0	29,045	7.4
	Pharmacy	1,096	1.0	2,812	0.7
	Dentistry	542	0.5	1,383	0.4
	Veterinary Science	526	0.5	1,482	0.4
	Physiotherapy	870	0.8	2,672	0.7
	Occupational Therapy	1,056	1.0	2,814	0.7
Education	Teacher Education – Other	2,192	2.0	6,835	1.8
	Teacher Education – Early Childhood	2,025	1.9	5,942	1.5
	Teacher Education – Primary & Secondary	5,203	4.8	18,480	4.7
Management and Commerce	Accounting	1,667	1.5	6,278	1.6
	Business Management	8,152	7.5	39,734	10.2
	Sales & Marketing	875	0.8	3,688	0.9
	Management & Commerce – Other	4,677	4.3	19,039	4.9
	Banking & Finance	922	0.9	3,731	1.0
Society and Culture	Political Science	651	0.6	1,963	0.5
	Humanities inc History & Geography	11,105	10.3	41,500	10.6
	Language & Literature	599	0.6	1,550	0.4
	Social Work	2,377	2.2	6,266	1.6
	Psychology	4,757	4.4	13,579	3.5
	Law	4,049	3.7	13,501	3.5
	Justice Studies & Policing	965	0.9	3,713	1.0
	Economics	1,222	1.1	4,990	1.3
	Sport & Recreation	225	0.2	965	0.2
Creative Arts	Art & Design	3,192	2.9	12,270	3.1
	Music & Performing Arts	1,405	1.3	4,274	1.1
	Communication, Media & Journalism	3,733	3.4	13,084	3.4
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	95	0.1	308	0.1
Total		**108,322**	**100**	**390,049**	**100**

3.3 Weighting

In the 2012 UES, weighting was undertaken to ensure that reported results were representative of the overall population. In 2013, weighting was trialled to correct the serious gender imbalance in the sample of secured responses, but was found to have no substantial impact on the results at a national level. Given the serious under-representation of males in the 2014 UES sample (see Table 5), corrective weighting was again trialled. To facilitate this, post-stratification weights by gender, subject area and stage of studies were computed separately for each institution.[10] This resulted in a total of 3,754 non-zero weighting strata.[11] Weights ranged in size from 0.7 to 56.0. The mean weight was 3.7 and the median 3.1.

Table 7. Comparison of raw and weighted percentage satisfied scores by subgroup

Group	Subgroup	Quality of entire educational experience		Quality of teaching		Teaching Quality	
		Raw	Weighted	Raw	Weighted	Raw	Weighted
Stage of studies	Commencing	83	82	83	82	84	83
	Later year	77	76	77	76	79	78
Gender	Male	78	78	78	77	80	79
	Female	82	82	82	82	83	83
Indigenous	Aboriginal or Torres Strait Islander	81	81	83	82	83	82
	Not Aboriginal or Torres Strait Islander	80	80	81	80	82	81
Home language	English	82	81	82	81	83	82
	Other	76	75	77	76	79	78
Disability	Disability reported	78	78	79	79	80	80
	No disability reported	81	80	81	80	82	81
Study mode	Internal	81	80	81	80	82	81
	External/multi-modal	79	79	80	79	81	81
International	Domestic student	81	81	81	80	82	82
	International student	74	74	76	75	78	78
First in family	First in family	83	83	83	83	85	84
	Not first in family	83	83	82	82	84	83
Total		**81**	**80**	**81**	**80**	**82**	**81**

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. As in 2013, two questionnaire items and one focus area were selected for this analysis: the quality of the 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. The results are presented in Table 7 (subgroup) and Table 8 (subject area).[12]

[10] 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.

[11] When calculating the weights, 11 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 7 and 8.

[12] 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.

Table 8. Comparison of raw and weighted percentage satisfied scores by subject area

Group	Subgroup	Quality of entire educational experience		Quality of teaching		Teaching Quality	
		Raw	Weighted	Raw	Weighted	Raw	Weighted
Natural and Physical Sciences	Natural & Physical Sciences	84	82	85	83	85	84
	Mathematics	79	79	78	78	81	81
	Biological Sciences	86	86	86	86	86	86
	Medical Science & Technology	86	86	86	86	87	87
IT	Computing & Information Systems	74	74	72	71	75	75
Engineering and Related Technologies	Engineering – Other	77	77	74	73	77	76
	Engineering – Process & Resources	73	71	70	69	75	73
	Engineering – Mechanical	74	74	68	68	72	72
	Engineering – Civil	78	78	72	71	76	76
	Engineering – Electrical & Electronic	78	78	74	76	79	79
	Engineering – Aerospace	76	76	73	74	76	77
Architecture and Building	Architecture & Urban Environments	75	75	73	73	78	78
	Building & Construction	76	76	70	71	74	76
Agriculture and Environmental Studies	Agriculture & Forestry	83	83	81	81	83	83
	Environmental Studies	83	82	84	82	85	85
Health	Health Services & Support	82	82	83	82	84	84
	Public Health	80	78	80	78	80	78
	Medicine	79	79	77	76	80	79
	Nursing	78	77	79	79	80	79
	Pharmacy	81	81	81	81	80	81
	Dentistry	73	73	68	68	75	76
	Veterinary Science	83	85	86	87	85	86
	Physiotherapy	88	87	86	85	88	88
	Occupational Therapy	87	85	85	84	88	87
Education	Teacher Education – Other	78	79	79	79	80	79
	Teacher Education – Early Childhood	81	81	83	82	83	83
	Teacher Education – Primary & Secondary	82	82	81	81	81	81
Management and Commerce	Accounting	78	77	77	77	79	78
	Business Management	78	78	77	76	78	78
	Sales & Marketing	80	79	78	78	80	80
	Management & Commerce – Other	77	77	75	75	78	77
	Banking & Finance	75	75	73	72	74	72
Society and Culture	Political Science	85	85	82	82	84	83
	Humanities inc History & Geography	83	83	85	85	86	86
	Language & Literature	86	85	87	86	89	88
	Social Work	81	81	82	82	84	84
	Psychology	84	84	87	87	87	88
	Law	82	81	83	83	85	84
	Justice Studies & Policing	81	80	82	81	83	82
	Economics	74	73	72	73	75	76
	Sport & Recreation	79	78	80	80	82	83
Creative Arts	Art & Design	79	79	79	79	82	82
	Music & Performing Arts	81	80	85	84	85	84
	Communication, Media & Journalism	83	83	83	83	84	84

Table 8. (continued)

Group	Subgroup	Quality of entire educational experience		Quality of teaching		Teaching Quality	
		Raw	Weighted	Raw	Weighted	Raw	Weighted
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	91	92	86	87	84	86
Total		**81**	**80**	**81**	**80**	**82**	**81**

It is evident from Tables 7 and 8 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 consistent with the results obtained in 2013 and is presumably related to the fact that the sample of secured responses reflects the in-scope population on most characteristics and subject area in particular. It was decided to analyse the data without applying weights. All results presented in this report, aside from those in Tables 7 and 8, are based on unweighted data.

3.4 Stratum-level precision

One of the major methodological improvements for the 2014 UES was the change in focus from the institution level to the stratum level (subject areas within institutions) for both sampling and response maximisation (see Section 2.3.2). The original intention of these methodological refinements was to reduce gender bias by targeting male-dominated subject areas for more intense and targeted response maximisation activities.

While the national response rate increased in 2014 relative to 2013, gender bias did not decrease (see Section 3.2). The main positive outcome from the stratum-level response maximisation was a general increase in the number of strata that met the desired level of precision (see Section 2.3.2). Table 9 shows that a combined total of 477 additional strata achieved the desired level of precision across the five focus areas in 2014, with more than 100 additional strata meeting precision targets in relation to each of the skills development, teaching quality and student support focus areas.

Table 9. Strata meeting desired level of precision,[a] 2013 and 2014

Focus area	2013		2014		Change
	n	%	n	%	
Skills Development	614	59.0	715	68.0	9.0
Learner Engagement	479	46.0	551	52.4	6.4
Teaching Quality	629	60.4	731	69.6	9.2
Student Support	405	38.9	522	49.7	10.8
Learning Resources	638	61.3	723	68.8	7.5
Total strata	**1,041**		**1,051**		

[a] ±7.5 percentage points at a 90 per cent level of confidence.

3.5 Precision of national estimates

Because the 2014 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 90 per cent confidence intervals around the point estimates. These

confidence intervals have been calculated as 1.645 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 G.

Table 10. Percentage satisfied results by subgroup with 90 per cent confidence intervals

Group	Subgroup	Quality of entire educational experience[a]	Quality of teaching[a]
Stage of studies	Commencing	82.6 (82.4,82.9)	82.5 (82.3,82.7)
	Later year	77.0 (76.7,77.3)	77.3 (77.0,77.6)
Gender	Male	77.9 (77.6,78.2)	77.5 (77.2,77.9)
	Female	81.7 (81.5,81.9)	82.0 (81.8,82.2)
Indigenous	Aboriginal or Torres Strait Islander	80.6 (79.1,82.2)	82.6 (81.1,84.1)
	Not Aboriginal or Torres Strait Islander	80.4 (80.3,80.6)	80.5 (80.3,80.7)
Home language	English	81.8 (81.6,82.0)	81.7 (81.5,81.9)
	Other	75.9 (75.5,76.3)	76.6 (76.2,77.0)
Disability	Disability reported	78.2 (77.5,79.0)	79.0 (78.3,79.7)
	No disability reported	80.6 (80.4,80.8)	80.6 (80.4,80.8)
Study mode	Internal	80.6 (80.4,80.8)	80.6 (80.4,80.8)
	External/multi-modal	79.4 (79.0,79.9)	80.0 (79.5,80.4)
International	Domestic student	81.2 (81.0,81.4)	81.1 (80.9,81.3)
	International student	74.2 (73.6,74.8)	75.6 (75.0,76.2)
First in family	First in family	83.2 (82.9,83.5)	83.4 (83.1,83.7)
	Not first in family	83.2 (82.9,83.5)	82.4 (82.1,82.7)
Total		**80.5 (80.3,80.6)**	**80.5 (80.4,80.7)**

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

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 90 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 4.1 and 4.8 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. It is important to note that 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 11. Percentage satisfied results by subject area with 90 per cent confidence intervals

Group	Subgroup	Quality of entire educational experience[a]	Quality of teaching[a]
Natural and Physical Sciences	Natural & Physical Sciences	83.7 (83.1,84.4)	85.0 (84.4,85.6)
	Mathematics	79.0 (76.4,81.6)	78.4 (75.8,81.0)
	Biological Sciences	86.2 (85.2,87.3)	86.2 (85.2,87.3)
	Medical Science & Technology	86.0 (85.1,86.8)	86.2 (85.4,87.0)
IT	Computing & Information Systems	74.0 (72.9,75.0)	71.6 (70.6,72.7)

Table 11. (continued)

Group	Subgroup	Quality of entire educational experience[a]	Quality of teaching[a]
Engineering and Related Technologies	Engineering – Other	77.3 (76.3,78.2)	73.9 (72.8,74.9)
	Engineering – Process & Resources	72.5 (69.9,75.1)	70.4 (67.8,73.0)
	Engineering – Mechanical	73.7 (71.4,76.0)	68.4 (65.9,70.8)
	Engineering – Civil	78.2 (76.3,80.1)	71.6 (69.6,73.7)
	Engineering – Electrical & Electronic	77.5 (75.5,79.5)	74.4 (72.3,76.4)
	Engineering – Aerospace	75.9 (73.2,78.5)	72.7 (70.0,75.5)
Architecture and Building	Architecture & Urban Environments	75.2 (73.8,76.5)	73.0 (71.6,74.4)
	Building & Construction	76.4 (73.5,79.2)	70.1 (67.0,73.1)
Agriculture and Environmental Studies	Agriculture & Forestry	82.9 (81.1,84.8)	81.0 (79.0,83.0)
	Environmental Studies	82.8 (81.3,84.3)	84.0 (82.5,85.4)
Health	Health Services & Support	82.3 (81.6,83.0)	82.8 (82.1,83.5)
	Public Health	79.8 (78.3,81.3)	79.9 (78.4,81.4)
	Medicine	79.1 (77.9,80.3)	76.8 (75.6,78.1)
	Nursing	77.8 (77.2,78.5)	79.3 (78.7,80.0)
	Pharmacy	80.7 (79.2,82.3)	81.3 (79.8,82.8)
	Dentistry	72.7 (70.2,75.1)	67.5 (64.9,70.1)
	Veterinary Science	82.9 (80.7,85.1)	86.3 (84.3,88.3)
	Physiotherapy	87.5 (86.0,89.0)	86.2 (84.6,87.8)
	Occupational Therapy	86.6 (85.3,88.0)	85.4 (84.0,86.8)
Education	Teacher Education – Other	78.3 (77.1,79.5)	78.8 (77.6,80.0)
	Teacher Education – Early Childhood	81.4 (80.2,82.5)	82.8 (81.7,83.9)
	Teacher Education – Primary & Secondary	81.6 (80.8,82.3)	81.2 (80.5,82.0)
Management and Commerce	Accounting	78.3 (76.9,79.8)	77.0 (75.6,78.5)
	Business Management	78.2 (77.5,78.9)	76.8 (76.1,77.5)
	Sales & Marketing	79.9 (77.9,81.8)	78.2 (76.2,80.2)
	Management & Commerce – Other	77.3 (76.4,78.1)	74.9 (74.0,75.8)
	Banking & Finance	75.2 (73.1,77.2)	72.5 (70.4,74.6)
Society and Culture	Political Science	84.5 (82.5,86.4)	82.3 (80.3,84.3)
	Humanities inc History & Geography	82.7 (82.2,83.2)	85.4 (84.9,85.9)
	Language & Literature	85.5 (83.6,87.3)	86.6 (84.8,88.4)
	Social Work	81.1 (80.0,82.1)	82.3 (81.3,83.3)
	Psychology	83.9 (83.2,84.6)	86.7 (86.0,87.3)
	Law	81.7 (80.8,82.5)	83.3 (82.5,84.1)
	Justice Studies & Policing	80.7 (78.9,82.5)	82.4 (80.7,84.2)
	Economics	73.7 (71.9,75.5)	72.2 (70.4,74.0)
	Sport & Recreation	78.7 (74.7,82.6)	79.9 (76.0,83.8)
Creative Arts	Art & Design	79.0 (78.0,80.0)	79.3 (78.2,80.3)
	Music & Performing Arts	81.0 (79.6,82.4)	85.0 (83.7,86.2)
	Communication, Media & Journalism	82.6 (81.7,83.5)	83.2 (82.3,84.0)
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	90.5 (86.4,94.6)	86.3 (81.5,91.1)
Total		**80.5 (80.3,80.6)**	**80.5 (80.4,80.7)**

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

4. Key results from the 2014 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 first adopted for the 2013 UES Report in the interest of parsimony of reporting and explanation, and is maintained here for consistency.

Considering first the overall results, there is much variation in percentage satisfied results. These ranged from 86 per cent in relation to the learning resources focus area, down to 61 per cent for the learner engagement focus area. Encouragingly, a relatively large proportion of students indicated satisfaction with the quality of teaching provided by their institution and their skills development (82 and 81 per cent, respectively). In terms of the student support provided by their institution, 73 per cent of survey respondents 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 scores by subgroup

Group	Subgroup	Skills Development	Learner Engagement	Teaching Quality	Student Support	Learning Resources
Stage of studies	Commencing	79	60	84	76	88
	Later year	85	63	79	68	81
Gender	Male	78	61	80	71	84
	Female	83	61	83	74	86
Age group	Under 25	82	65	82	72	86
	25 to 29	80	53	80	72	83
	30 to 39	79	46	82	76	84
	40 and over	79	42	84	78	84
Indigenous	Aboriginal or Torres Strait Islander	82	55	83	76	87
	Not Aboriginal or Torres Strait Islander	81	61	82	73	86
Home language	English	82	61	83	73	86
	Other	80	60	79	70	84
Disability	Disability reported	77	59	80	77	83
	No disability reported	82	61	82	72	86
Study mode	Internal	82	65	82	72	86
	External/multi-modal	81	40	81	74	84
International	Domestic student	82	62	82	73	86
	International student	80	57	78	70	84
First in family	First in family	80	59	85	77	89
	Not first in family	79	63	84	75	88
Previous university experience	Current university	80	58	84	74	86
	Another university	78	53	84	76	86
	New to higher education	80	62	84	76	89
Total		**81**	**61**	**82**	**73**	**86**

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.

Considering male and female students, most differences in percentage satisfied results were fairly 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 male and female students. No difference between males and females was observed in relation to learner engagement.

In relation to study mode, internal students were vastly more likely to be satisfied with their level of engagement than those studying externally or by mixed mode, with 25 percentage points between the groups. The differences in relation to the other four focus areas were relatively small.

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 40 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, as previously identified, 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 1.2 per cent of the sample (see Table 5), were less likely than their non-Indigenous classmates to be satisfied with their level of engagement, a result which is of some concern. They were, however, somewhat more likely to be satisfied with the student support provided by their institution. Differences between Indigenous and non-Indigenous students in relation to the other three 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.

Students who spoke English as their main language at home were more likely than those from a non-English speaking background to be satisfied with every aspect of their educational experience. Aside from teaching quality, however, which saw four percentage points separate 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. These differences were largest 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 five 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, with the exception of skills development, these differences were relatively small.

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 15 percentage points separating the subject areas (medical science and technology with the highest results, and economics with the lowest), followed by the teaching quality (17 percentage points), skills development (19 percentage points) and learning resources (20 percentage points) focus areas. The widest range is observed for learner engagement, with 30 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 considerably less likely to express satisfaction in relation to skills development and teaching quality than students in other subject areas.[13] Another example can be seen for dentistry students in relation to teaching quality and learning resources. These two results underscore the fact that broad disciplinary aggregations hide much of the detail that is relevant to schools, faculties and academic departments.

Table 13. Percentage satisfied scores by subject area

Group	Subgroup	Skills Development	Learner Engagement	Teaching Quality	Student Support	Learning Resources
Natural and Physical Sciences	Natural & Physical Sciences	80	63	85	74	88
	Mathematics	76	57	81	78	85
	Biological Sciences	85	67	86	75	90
	Medical Science & Technology	84	69	87	79	91
IT	Computing & Information Systems	75	58	75	71	83
Engineering and Related Technologies	Engineering – Other	79	66	77	71	86
	Engineering – Process & Resources	82	75	75	71	81
	Engineering – Mechanical	76	69	72	67	81
	Engineering – Civil	80	72	76	67	82
	Engineering – Electrical & Electronic	78	67	79	73	88
	Engineering – Aerospace	79	65	76	69	82
Architecture and Building	Architecture & Urban Environments	81	69	78	67	73
	Building & Construction	79	60	74	69	82
Agriculture and Environmental Studies	Agriculture & Forestry	76	62	83	73	87
	Environmental Studies	84	64	85	76	87
Health	Health Services & Support	84	63	84	74	88
	Public Health	81	64	80	75	88
	Medicine	88	76	80	73	82
	Nursing	86	61	80	75	86

[13] 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 generally closer to that of management and commerce students than that of society and culture students.

Table 13. (continued)

Group	Subgroup	Skills Development	Learner Engagement	Teaching Quality	Student Support	Learning Resources
	Pharmacy	84	67	80	72	85
	Dentistry	85	57	75	70	76
	Veterinary Science	84	75	85	73	86
	Physiotherapy	90	76	88	78	88
	Occupational Therapy	90	74	88	78	91
Education	Teacher Education – Other	80	59	80	72	86
	Teacher Education – Early Childhood	86	55	83	75	87
	Teacher Education – Primary & Secondary	82	61	81	73	86
Management and Commerce	Accounting	76	50	79	71	83
	Business Management	79	58	78	71	85
	Sales & Marketing	83	65	80	70	86
	Management & Commerce – Other	76	57	78	70	85
	Banking & Finance	73	49	74	68	84
Society and Culture	Political Science	82	68	84	74	86
	Humanities inc History & Geography	80	56	86	72	85
	Language & Literature	80	55	89	75	85
	Social Work	85	53	84	76	84
	Psychology	83	56	87	76	89
	Law	86	57	85	69	85
	Justice Studies & Policing	79	46	83	74	84
	Economics	72	55	75	64	82
	Sport & Recreation	81	65	82	66	86
Creative Arts	Art & Design	80	64	82	71	81
	Music & Performing Arts	81	71	85	71	79
	Communication, Media & Journalism	83	66	84	71	87
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	84	61	84	76	93
Total		**81**	**61**	**82**	**73**	**86**

While confidence intervals are not shown in Table 13, it is important to interpret the results with respect to the remarks made in Section 3.5 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.

4.2 University experience perceptions over time

Table 14 compares results from the 2014 UES with those from 2013. 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, there were two notable methodological changes between 2013 and 2014, specifically the adoption of a stratum-level approach to sampling and response maximisation, and the removal of module rotation. At least some of the differences in results between the two collections under examination may be attributable to these changes. Finally, as noted in Section 1.2, one questionnaire item was removed from the student support focus for the 2014 UES. Although the

2013 focus area results have been recalculated without this item, it is possible that respondents may have approached the other questions in this focus area differently due to the exclusion of this item.

Table 14. Percentage satisfied results by subject area, 2013 and 2014

Group	Subgroup	2013[ab]					2014[a]				
		SD	LE	TQ	SS	LR	SD	LE	TQ	SS	LR
Natural and Physical Sciences	Natural & Physical Sciences	77	58	83	60	87	80	63	85	74	88
	Mathematics	73	53	80	64	87	76	57	81	78	85
	Biological Sciences	82	62	84	62	87	85	67	86	75	90
	Medical Science & Technology	80	63	82	62	87	84	69	87	79	91
IT	Computing & Information Systems	72	58	74	58	81	75	58	75	71	83
Engineering and Related Technologies	Engineering – Other	76	63	71	57	81	79	66	77	71	86
	Engineering – Process & Resources	79	69	73	57	82	82	75	75	71	81
	Engineering – Mechanical	76	61	70	54	78	76	69	72	67	81
	Engineering – Civil	78	66	71	55	81	80	72	76	67	82
	Engineering – Electrical & Electronic	73	63	72	58	80	78	67	79	73	88
	Engineering – Aerospace	77	63	71	55	82	79	65	76	69	82
Architecture and Building	Architecture & Urban Environments	77	61	75	52	71	81	69	78	67	73
	Building & Construction	72	53	70	53	81	79	60	74	69	82
Agriculture and Environmental Studies	Agriculture & Forestry	73	56	77	62	85	76	62	83	73	87
	Environmental Studies	79	61	84	64	84	84	64	85	76	87
Health	Health Services & Support	80	61	81	59	86	84	63	84	74	88
	Public Health	84	61	83	63	87	81	64	80	75	88
	Medicine	85	73	75	58	78	88	76	80	73	82
	Nursing	84	55	75	61	84	86	61	80	75	86
	Pharmacy	84	65	80	60	85	84	67	80	72	85
	Dentistry	82	62	72	55	76	85	57	75	70	76
	Veterinary Science	85	71	84	57	81	84	75	85	73	86
	Physiotherapy	87	77	86	65	90	90	76	88	78	88
	Occupational Therapy	85	70	82	61	90	90	74	88	78	91
Education	Teacher Education – Other	81	54	78	60	85	80	59	80	72	86
	Teacher Education – Early Childhood	84	49	82	62	84	86	55	83	75	87
	Teacher Education – Primary & Secondary	82	60	80	58	84	82	61	81	73	86
Management and Commerce	Accounting	74	45	76	60	82	76	50	79	71	83
	Business Management	77	54	76	57	83	79	58	78	71	85
	Sales & Marketing	78	57	75	55	81	83	65	80	70	86
	Management & Commerce – Other	75	52	75	59	82	76	57	78	70	85
	Banking & Finance	71	45	71	56	83	73	49	74	68	84
Society and Culture	Political Science	76	52	82	57	81	82	68	84	74	86
	Humanities inc History & Geography	78	52	84	58	83	80	56	86	72	85
	Language & Literature	77	51	88	67	85	80	55	89	75	85
	Social Work	83	48	82	60	80	85	53	84	76	84
	Psychology	82	55	86	64	86	83	56	87	76	89
	Law	83	54	81	57	84	86	57	85	69	85
	Justice Studies & Policing	74	44	75	56	81	79	46	83	74	84

Table 14. (continued)

Group	Subgroup	2013[ab]					2014[a]				
		SD	LE	TQ	SS	LR	SD	LE	TQ	SS	LR
	Economics	68	50	71	50	81	72	55	75	64	82
	Sport & Recreation	85	60	78	59	89	81	65	82	66	86
Creative Arts	Art & Design	78	60	80	56	79	80	64	82	71	81
	Music & Performing Arts	78	71	82	59	78	81	71	85	71	79
	Communication, Media & Journalism	81	61	82	57	85	83	66	84	71	87
Food, Hospitality and Personal Services	Tourism, Hospitality & Personal Services	81	58	83	57	91	84	61	84	76	93
Total		**79**	**57**	**79**	**59**	**83**	**81**	**61**	**82**	**73**	**86**

[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 2013 and 2014 results (see Section 4.2). Moreover, the 2013 results on the student support focus area will differ to those presented in the 2013 UES National Report, as these have been recalculated without a questionnaire item omitted from the 2014 UES (see Section 1.2).

At the national level, higher percentage satisfied results were observed across all five focus areas in 2014. The largest difference in results between years was seen in relation to the student support focus area, with 14 percentage points separating 2013 (59 per cent) and 2014 (73 per cent). A year-on-year difference of this magnitude is most likely the result of changes to the questionnaire and survey method and, as will be discussed in Section 4.3, appears to be related to unusually large increases on several questionnaire items in this focus area.

The next largest difference was observed in relation to learner engagement (4 percentage points), followed by teaching quality and learning resources (each with 3 percentage points). Two percentage points separated 2013 and 2014 in relation to skills development. 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 2013 and 2014 collections.[14]

4.3 Results on individual questionnaire items

Table 15 presents percentage satisfied results for the 46 individual survey items underpinning the five UES focus areas, stratified by stage of studies. Results from the 2013 UES are presented to facilitate comparisons over time. When reading the results in Table 15, the previously discussed caveats on comparing the 2013 and 2014 UES collections should be borne in mind (see Section 4.2). Detailed response category percentages are presented in Appendix H.

[14] A clear potential source of differences between 2013 and 2014 is differential response rates at the stratum level. To investigate this, weights were computed by diving 2013 stratum counts by those from 2014. These weights were then applied to the 2014 sample so that the distribution of responses across strata would reflect that of the 2013 UES. This was found to have only a trivial impact on the 2014 results, which provides some evidence that the annual differences reported in Table 15 are not strongly related to differences in stratum level response rates between years. That being said, if late responders to the UES differ systematically to early responders, reducing the response requirements across many strata in 2014 could have introduced bias, which would not be addressed by the analysis described in this note.

Table 15. Percentage satisfied results for UEQ items by stage of studies, 2013 and 2014

Focus area	Item	2013[ab]			2014[a]		
		C	LY	Tot.	C	LY	Tot.
Skills Development	Developed critical and analytical thinking	66	73	69	68	76	71
	Developed ability to solve complex problems	53	64	58	57	67	61
	Developed ability to work effectively with others	57	64	60	59	67	62
	Developed confidence to learn independently	67	75	70	69	77	72
	Developed written communication skills	57	68	61	58	71	63
	Developed spoken communication skills	47	58	52	50	61	54
	Developed knowledge of field studying	76	78	77	78	80	79
	Developed work-related knowledge and skills	60	62	60	62	65	63
Learner Engagement	Felt prepared for your study	56	61	58	63	69	66
	Had a sense of belonging to your university	52	47	50	54	50	53
	Participated in discussions online or face-to-face	54	59	56	57	63	59
	Worked with other students as part of your study	60	64	62	62	67	64
	Interacted with students outside study requirements	45	47	46	46	48	46
	Interacted with students who are very different from you	54	52	53	55	53	55
	Been given opportunities to interact with local students	56	54	55	58	57	58
Teaching Quality	Study well structured and focused	68	61	65	71	64	68
	Study relevant to education as a whole	71	68	70	74	70	72
	Teachers engaged you actively in learning	62	59	61	65	62	64
	Teachers demonstrated concern for student learning	59	56	57	61	59	60
	Teachers provided clear explanations on coursework and assessment	63	61	62	66	62	65
	Teachers stimulated you intellectually	68	65	67	70	68	69
	Teachers commented on your work in ways that help you learn	49	49	49	51	52	52
	Teachers seemed helpful and approachable	70	68	69	73	71	72
	Teachers set assessment tasks that challenge you to learn	77	72	75	79	75	77
	Quality of teaching	81	76	79	83	77	81
	Quality of entire educational experience	82	76	79	83	77	81
Student Support	Experienced efficient enrolment and admissions processes	68	63	66	73	70	72
	Induction/orientation activities relevant and helpful	51	42	48	60	50	56
	Received support from university to settle into study	52	40	47	62	51	58
	Administrative staff or systems: available	63	56	61	65	60	63
	Administrative staff or systems: helpful	61	53	58	62	56	60
	Careers advisors: available	47	42	45	49	45	47
	Careers advisors: helpful	47	41	44	49	44	47
	Academic or learning advisors: available	61	57	59	63	59	62
	Academic or learning advisors: helpful	64	59	62	66	62	65
	Support services: available	54	49	52	56	53	55
	Support services: helpful	54	51	53	57	55	56
	Offered support relevant to circumstances	28	24	26	48	43	46
	Received appropriate English language skill support	20	16	18	38	31	35
Learning Resources	Quality of teaching spaces	86	80	84	88	82	86
	Quality of student spaces and common areas	78	70	75	80	73	78
	Quality of online learning materials	86	82	84	88	84	86
	Quality of computing/IT resources	83	77	80	85	79	83
	Quality of assigned books, notes and resources	81	76	79	82	77	80
	Quality of laboratory or studio equipment	85	77	82	87	80	84
	Quality of library resources and facilities	88	84	87	89	87	88

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

[b] Due to changes in methodology, care should be taken when comparing 2013 and 2014 results (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 (88 per cent), along with the quality of teaching spaces and online learning materials (both with 86 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 81 per cent), which were also amongst the highest-rated items. Many of the lowest results were associated with the student support and learner engagement focus areas, which may be of some concern to institutions. In relation to student support, only 35 per cent of respondents indicated that they received appropriate English language support, whilst fewer than half believed that they had been offered support relevant to their circumstances (46 per cent), and that careers advisors were available

and helpful (each with 47 percent). In relation to student engagement, only 46 per cent reported interacting with students outside of study requirements.

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 (13 percentage), spoken communication skills (11 percentage points) and the ability to solve complex problems (each with 10 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 11 percentage points separating them. A similar result was observed in relation to enrolment and admissions processes (10 percentage points). Given that these experiences would still be fresh in the minds of commencing students, these are hardly surprising results. 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 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 10 percentage points separating the lowest and highest results. Conversely, 37 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.

Although these same broad trends were also observed in 2013, several notable differences were observed between years in relation to the magnitude of the results. In particular, students who completed the 2014 UES were more likely than those who completed the 2013 survey to indicate that they were offered support relative to their circumstances (20 percentage points), received appropriate English language skill support (17 percentage points) and support provided by the university to settle into study (11 percentage points); all items within the student support focus area. By way of comparison, the average increase across questionnaire items between 2013 and 2014 was four percentage points. It is highly likely that these large differences are related to the removal of the university services item from the 2014 UEQ and the impact this modification had on how students approached the remaining items. The “offered support” and “English language” items, for instance, were directly adjacent to the now-omitted university services item on the 2013 UEQ and recorded much higher proportions of “not applicable” responses in 2014. The removal of module rotation in 2014 may also be a contributing factor.

Interestingly, no item obtained a higher result in 2013 than 2014. Given the previously discussed methodological differences between the two collections, these results should be interpreted with caution. They do not necessarily reflect genuine improvements in practice.

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, 90 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.

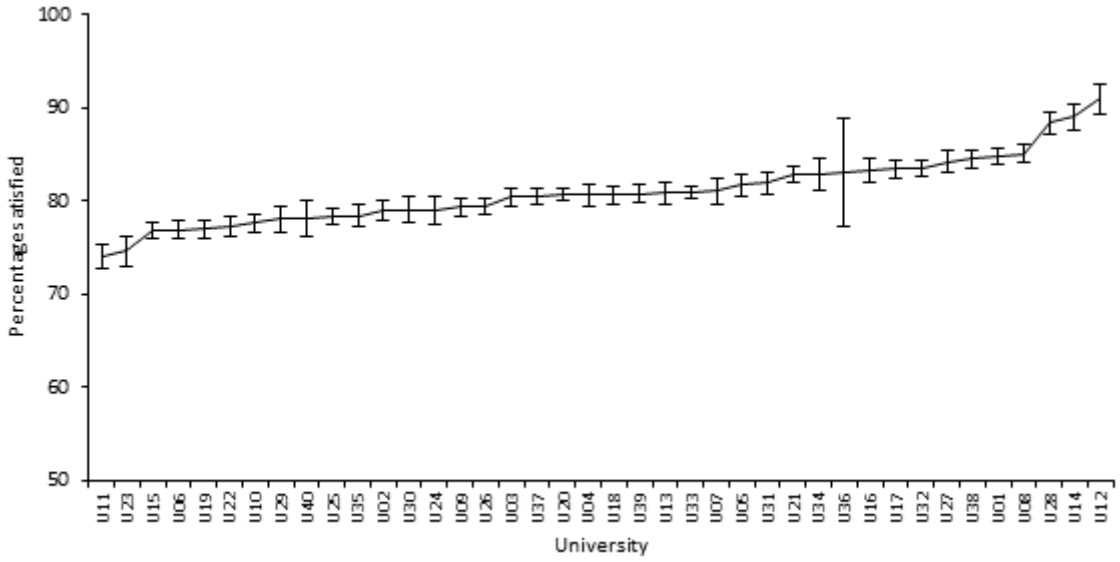


Figure 1. Percentage satisfied results on the quality of entire educational experience

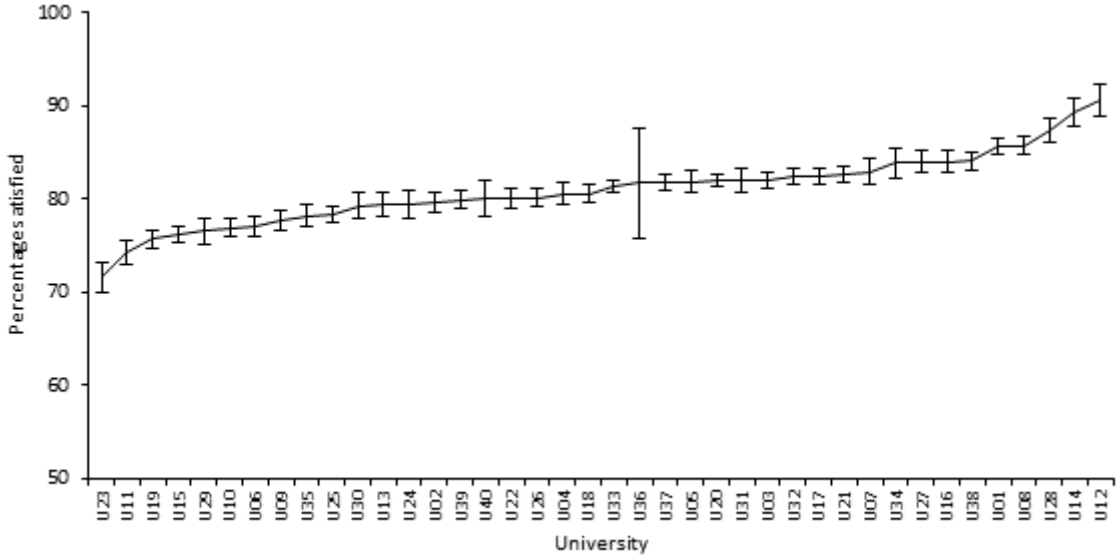


Figure 2. Percentage satisfied results on the quality of teaching

[15] For example, "U01" represents the same institution in Figures 1 and 2. Moreover, it also represents the same institution in the 2013 and 2014 editions of this report.

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. Indeed, there is a strong correlation in the ranking of institutions in both figures and the slopes of the lines are similar; 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.

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 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 473,633 students from 716 colleges and universities completing the 2014 NSSE.[17]

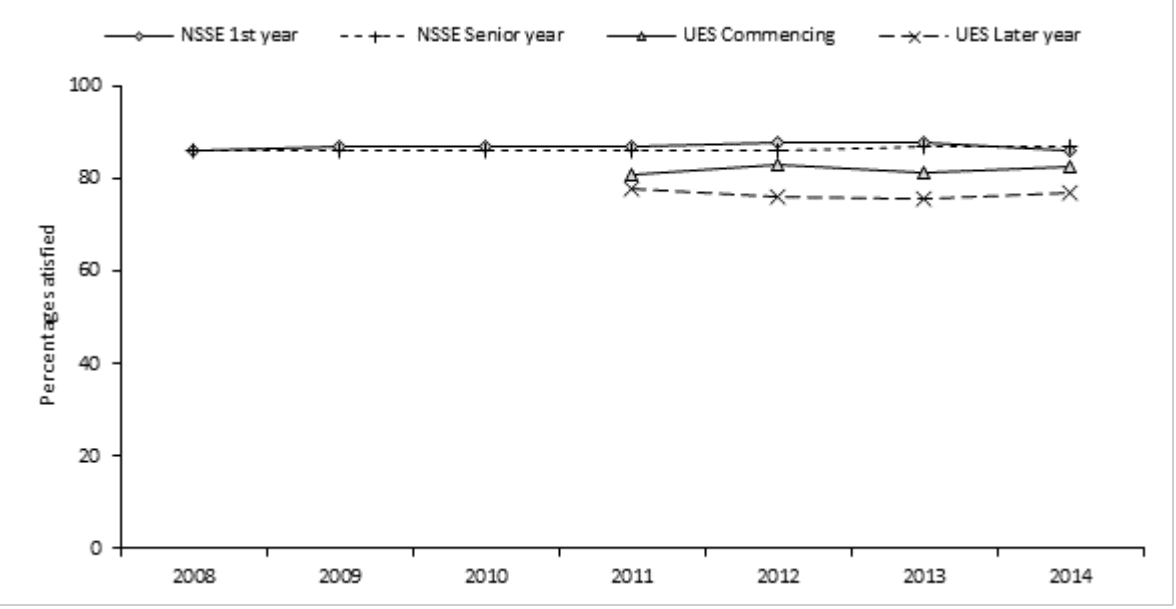


Figure 3. Entire educational experience rated positively, UES and NSSE, 2008 to 2014

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 4.2 concerning changes to the

[16] "How would you evaluate your entire educational experience at this institution?"
[17] Indiana University. (2014). *About NSSE*. Retrieved 16 Dec., 2014, from http://nsse.iub.edu/html/about.cfm

UES collection methodology should also be considered in relation to this figure. It is also critical to note that, while the 2012, 2013 and 2014 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.

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.

In 2014, four CEQ scales were administered to a small sample of UES respondents to facilitate benchmarking with the UK National Student Survey (NSS), which contains several questions with similar wording. Most notably, both the CEQ and NSS have an overall satisfaction item with near-identical wording,[18] measured on a five-point Likert-type response scale. The NSS, administered mostly to final year undergraduates, is run across all publicly funded higher education institutions in England, Wales, Northern Ireland and Scotland,[19] reducing the potential for non-random selection inherent in the NSSE.

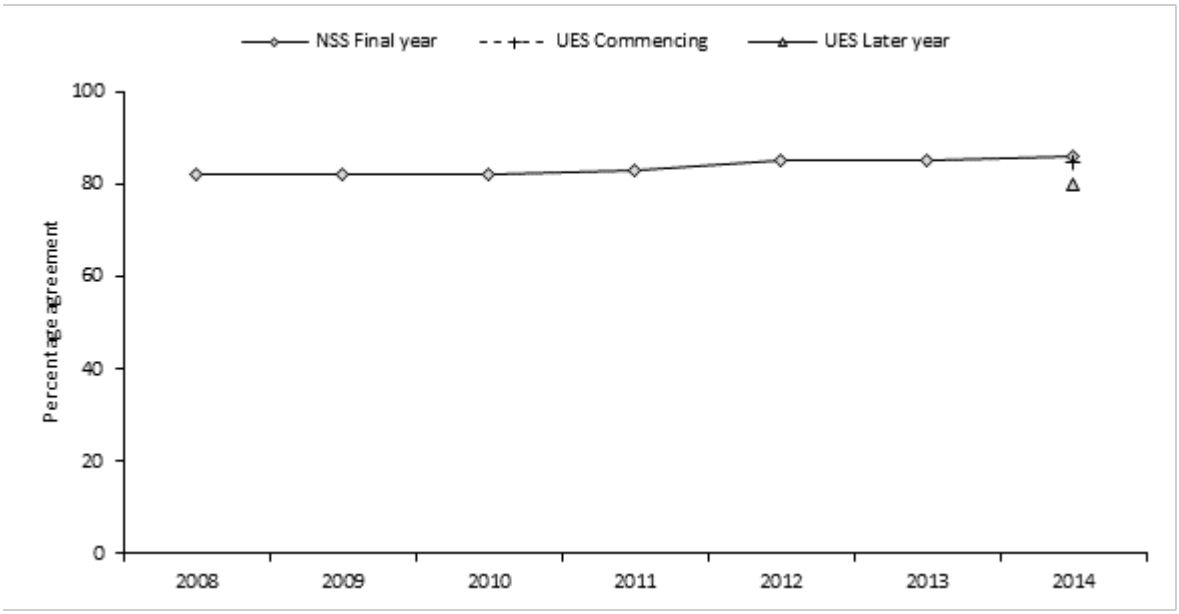


Figure 4. Overall satisfaction with course quality, UES CEQ and NSS, 2008 to 2014

[18] "Overall, I am satisfied with the quality of the [this] course."
[19] HEFCE. (2013). *The National Student Survey*. Retrieved 16 Dec., 2014, from http://www.thestudentsurvey.com/the_nss.html

Figure 4 presents the percentage of NSS and UES CEQ respondents who were satisfied with the quality of their course. Comparing final/later-year students, it can be seen that UK students are more likely to express satisfaction with the quality of their course, with around six percentage points separating the two groups. Given the large number of responses to both surveys,[20] this difference is likely to be statistically significant; however it does not account for potential differences in the composition of the respective undergraduate student populations, nor methodological differences between the two surveys. It is interesting, however, that both surveys show Australian-enrolled students to be less likely to be satisfied with their higher education experience than their overseas counterparts. Also of interest in this figure is the extent to which the NSS overall satisfaction results are consistent over time.

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 2014. The results of this question are presented by student subgroup in Table 16. Overall, 17 per cent of respondents indicated that they had considered leaving, the same proportion as in 2013.

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	18	Disability	Disability reported	24
	Later year	15		No disability reported	16
Gender	Male	16	Study mode	Internal	17
	Female	17		External/multi-modal	18
Age group	Under 25	16	International	Domestic student	17
	25 to 29	19		International student	14
	30 to 39	21	First in family	First in family	19
	40 and over	21		Not first in family	16
Indigenous	ATSI	27	Previous university experience	Current university	19
	Not ATSI	17		Another university	17
Home language	English	18		New to higher education	18
	Other	14	**Total**		**17**

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 three 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 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

[20] In all, 3,057 and 2,506 commencing and later-year students, respectively, responded to the CEQ item. The level of statistical precision achieved for this item was ±1.1 and ±1.3 percentage points at a 90 per cent confidence level for commencing and later-year students, respectively, at the national level.

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.

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 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 study mode, gender or previous university experience.

The percentage of students considering leaving their university in 2013 is plotted against (self-reported) average grades in Figure 5. 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 2014.

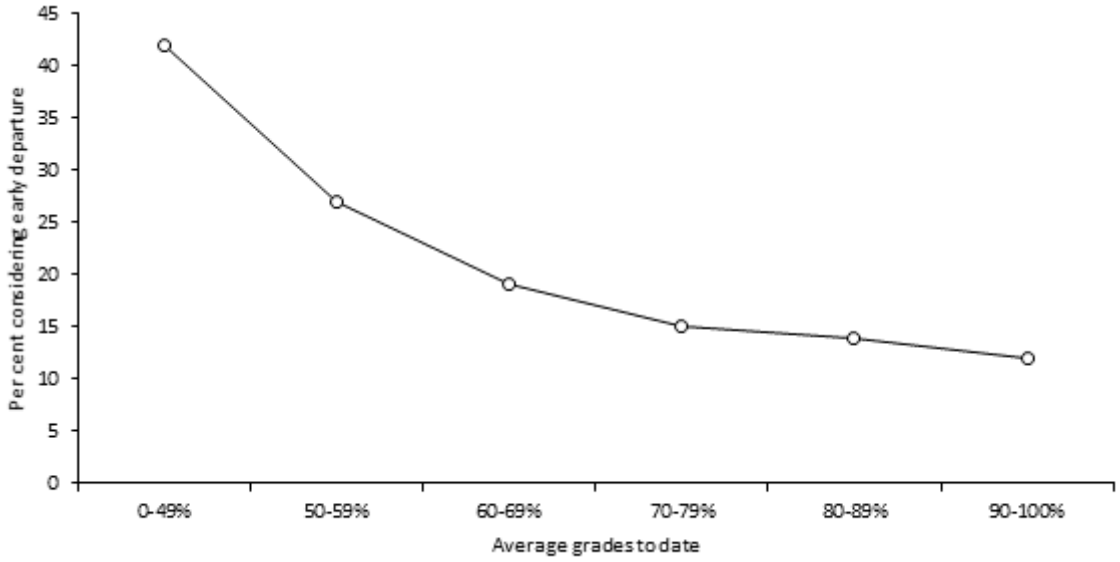


Figure 5. Percentage of students considering early departure by average grades to date

Students who expressed a serious consideration of leaving their university in 2014 were then asked to indicate, from a list of 30 possible reasons, why they considered doing so. These are summarised in Table 17, along with equivalent results from the 2013 UES. 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), difficulties relating to finances and workload, and study/life balance (each with 28 per cent), unspecified personal reasons (24 per cent) and the need to do paid work (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.

Table 17. Selected reasons for considering early departure, 2013 and 2014

Departure reason	Per cent of those considering departure		Departure reason	Per cent of those considering departure	
	2013	2014		2013	2014
Health or stress	31	31	Other	13	15
Financial difficulties	29	28	Gap year / deferral	12	11
Workload difficulties	28	28	Commuting difficulties	11	11
Study / life balance	29	28	Academic exchange	10	10
Expectations not met	24	25	Fee difficulties	9	9
Personal reasons	24	24	Other opportunities	9	9
Need to do paid work	24	23	Social reasons	8	8
Boredom/lack of interest	21	22	Travel or tourism	7	7
Career prospects	20	21	Institution reputation	7	7
Change of direction	20	20	Administrative support	7	6
Need a break	20	19	Moving residence	5	5
Family responsibilities	18	18	Standards too high	6	5
Quality concerns	15	15	Graduating	5	5
Academic support	15	15	Government assistance	4	3
Paid work responsibilities	15	15	Received other offer	3	3

Encouragingly for institutions, the most common (arguably) institutional factor indicated by students was that their expectations had not been met (25 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 (22 per cent), career prospects (21 per cent), a change in direction (20 per cent) and a need to take a break (19 per cent). These results are broadly consistent with the 2013 UES.

4.7 Results for non-university higher education students

As discussed in Section 1.2.1, the 2014 UES project included a pilot of the UES questionnaire and methodology on students of 15 NUHEIs under the name “Student Experience Survey”. Because the 2014 SES was a pilot project involving a small convenience sample of NUHEIs, only high-level aggregate results are presented for this cohort. These results are presented in Table 18, along with benchmark results for university students.

Table 18. Summary results for non-university higher education institutions and universities

	Focus areas					Questionnaire items	
	Skills Development	Learner Engagement	Teaching Quality	Student Support	Learning Resources	Quality of entire educational experience	Quality of teaching
NUHEIs	88	71	89	82	80	85	86
Universities	81	61	82	73	86	81	81

When comparing results for NUHEI and university students there are several important caveats to consider. First, only 15 of the approximately 130 NUHEIs currently operating in Australia were participants in SES pilot and these were not selected in a random fashion. Hence, they may not be representative of the non-university higher education sector as a whole. Second, the NUHEIs in the sample tend to teach a narrower range of subject areas than the universities (23 compared with 45,

respectively). Finally, the demographic characteristics of the two samples differ in several important respects. In particular, NUHEI students are more likely than their peers from universities to be international students, speak a language other than English at home and be the first in their family to enrol in higher education. Differences in results between NUHEI and university students may be attributable, at least in part, to these factors.

As shown in Table 18, NUHEI students tended to be more likely than university students to indicate satisfaction with their higher education experience. Only in relation to the learning resources focus area did a larger proportion of university students than NUHEI students indicate satisfaction. The largest differences between NUHEIs and universities across the five focus areas were observed in relation to learner engagement and student support, with 10 percentage points separating the two student cohorts. The smallest differences across focus areas were observed in relation to skills development (6 percentage points favouring NUHEIs) and learning resources (6 percentage points favouring universities). Four percentage points and six percentage points separated NUHEI and university students regarding satisfaction with the quality of their entire educational experience and the quality of teaching, respectively.

APPENDICES

Appendix A: University Experience Questionnaire (UEQ)

Table 19. 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 20. Learner Engagement items

Stem	Item	Response scale
At university during 2014, 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 2014, 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 2014, 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 21. 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 2014, 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 2014, 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 22. Student Support items

Stem	Item	Response scale
At university during 2014, 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 2014, 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 2014, 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 2014, 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 2014, 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?	
	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 23. 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 24. 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: Analysis of the psychometric properties of the revised Student Support focus area

B.1 Introduction

As discussed in Section 1.2, a psychometric assessment of responses to the 2013 UES identified one item in the student support focus area that was answered unpredictably by respondents and whose inclusion in the calculation of scores was likely to degrade the quality of measurement.

This appendix summarises a re-assessment of the student support focus area in light of the omission of the problematic item from the 2014 UES. As in 2013, the Rasch measurement model was used to assess the following aspects:

- How well the items in the focus area seemed to work 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.

B.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 B.7, 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. Both of these are undesirable and degrade the quality of derived measures.

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

B.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 25.

Table 25. Selected Rasch model outputs for item assessment

Statistic	Meaning	Ideal range
Infit mean square	This is the mean of the squared residuals, [21] 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.
Person separation index	This indicates how well the set of items is able to distinguish between the persons measured	Values below 2 imply that the instrument may not be sensitive enough to distinguish between high and low scoring persons.
Item separation index	This indicates how well the sample of persons enables the item locations to be determined	Values below 3 imply that the person sample is not large enough to confirm the hierarchy of item difficulty.

B.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.

B.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.

[21] 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.

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.

B.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.[22] 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.

The following sections summarise these aspects of quality for the student support focus area, namely item fit, category thresholds, and targeting.

B.3 Fit to the Rasch model

The Rasch summary statistics are shown in Table 26. The most difficult item for respondents to endorse was *englang* ("To what extent have you received appropriate English language skill support?") and the easiest was *effenrolm* ("To what extent have you experienced efficient enrolment and admissions processes?"). For these two items, the proportion of respondents indicating "Very much" was 17% and 37%, respectively.

Table 26. Item statistics for Student Support

Item	Measure	Mean square		Point-measure correlation	Item discrimination	Missing values (%)	Measure (2013)
		Infit	Outfit				
uniservices	-	-	-	-	-	-	1.53
englang	0.91	1.78	1.76	0.62	0.26	64.24	1.28
carhelp	0.35	0.88	0.90	0.69	1.10	60.21	0.12
offsup	0.35	1.38	1.36	0.67	0.67	35.15	0.81
caravail	0.34	0.82	0.84	0.69	1.14	59.38	0.11
supavail	0.11	0.89	0.88	0.68	1.10	67.07	-0.11
suphelp	0.06	0.95	0.93	0.67	1.06	68.02	-0.15
indorien	-0.16	1.11	1.18	0.59	0.80	10.82	-0.23
admhelp	-0.17	0.80	0.80	0.68	1.21	16.97	-0.49
supsettle	-0.20	0.93	0.98	0.65	1.02	0.22	-0.19
acdavail	-0.22	0.67	0.68	0.70	1.33	28.04	-0.54
admavail	-0.29	0.75	0.76	0.67	1.23	16.84	-0.59
acdhelp	-0.32	0.73	0.72	0.69	1.30	28.43	-0.64
effenrolm	-0.77	1.27	1.33	0.54	0.70	0.20	-0.91

The high mean square and low discrimination values for *englang* suggest that it was answered in a very unpredictable way by UES 2014 respondents—whether or not they received English language support seemed to be unrelated to their rating of other items in this focus area. In 2013, *englang* was the next worst fitting item after the now-omitted *uniservices*.[23] Including it in the calculation of student support scores risks lowering the quality of measurement for this focus area. Since *englang*

[22] Bond, T.G., & Fox, M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. (2nd ed.) Mahwah, NJ: Erlbaum.
[23] The infit mean square value in 2013 was 1.54 and the item discrimination was 0.73.

was answered by only a minority of students, the omission of the misfitting (but widely answered) *uniservices* would likely have a larger net effect on the quality of focus area scores.

The item measures for 2013 are shown as the right-most column in Table 26, including that for *uniservices*. The Rasch model centres item difficulties on zero by default during the estimation procedure, so some variation from 2013 to 2014 in item positions should be expected. Given the size of the cohort, though, large changes in the relative position of items should not occur. For the student support focus area, item locations have changed the most for *offsup* and *englang*. Overall, items now measure a narrower range of the underlying trait compared to 2013.

The high degree of item-level "missingness" suggests that only selected aspects of student support were relevant to respondents. Only two items—*supsettle* ("To what extent have you received support from your university to settle into study?") and *effenrolm*—seemed to apply to the group as a whole. Three additional items—*indorien* ("To what extent have you felt induction/orientation activities were relevant and helpful"), *admavail* ("To what extent have you found administrative staff or systems to be available") and *admhelp* ("To what extent have you found administrative staff or systems to be helpful")—appeared to be relevant to a large majority of respondents.

Table 27 shows the number of student support items answered by respondents. Fewer than 13% of respondents answered all items and 51% of respondents skipped five or more items.

Although the Rasch model makes use of all available data, estimates for datasets with a significant extent of missing responses will be associated with higher standard errors.[24] To improve measurement precision for this focus area, then, it would be best to re-visit its theoretical basis and develop items that are of relevance to a much greater proportion of respondents.

Table 27. Number of Student Support items answered by UES respondents[a]

Items answered	Percentage of respondents	Cumulative percentage of respondents
1	< 0.01	< 0.01
2	0.81	0.82
3	4.14	4.96
4	3.25	8.21
5	9.70	17.91
6	7.11	25.02
7	15.84	40.86
8	10.44	51.30
9	13.69	64.98
10	7.69	72.67
11	8.99	81.67
12	5.51	87.17
13	12.83	100.00

[a] This table only includes records with at least one completed item for this focus area.

The person and item separation indices are presented in Table 28, along with Cronbach's alpha, for both 2013 and 2014. The indices show how precisely the instrument enables persons and items to be located on the measurement scale. For both years, the large number of respondents ensures that item

[24] Linacre, J.M. (2014). *Winsteps® Rasch measurement computer program user's guide*. Beaverton, Oregon: Winsteps.com

locations are well established. The discrimination among persons is only marginally adequate, however, and somewhat worse in 2014. The degree of missingness among items will be a major contributor to reduced person separation.

Table 28. Person and item separation indices for Student Support

Summary statistic	2014	2013
Person separation index (> 2)	2.07	2.2
Item separation index (> 3)	78.58	140.3
Cronbach's alpha	0.92	0.91

B.4 Ordered category thresholds

Figure 6 shows the probabilities of respondents selecting the various response categories for the student support items, all of which use a common five-category response scale (see Appendix A). Respondents whose position is low on the latent trait (on the left end of the x-axis) are most likely to select "Not at all" to these items. By contrast, respondents with high levels of the trait (on the right end of the x-axis) are most likely to select "Very much". Respondents in the middle will select one of "Very little", "Some" or "Quite a bit", depending on their precise location on the trait.

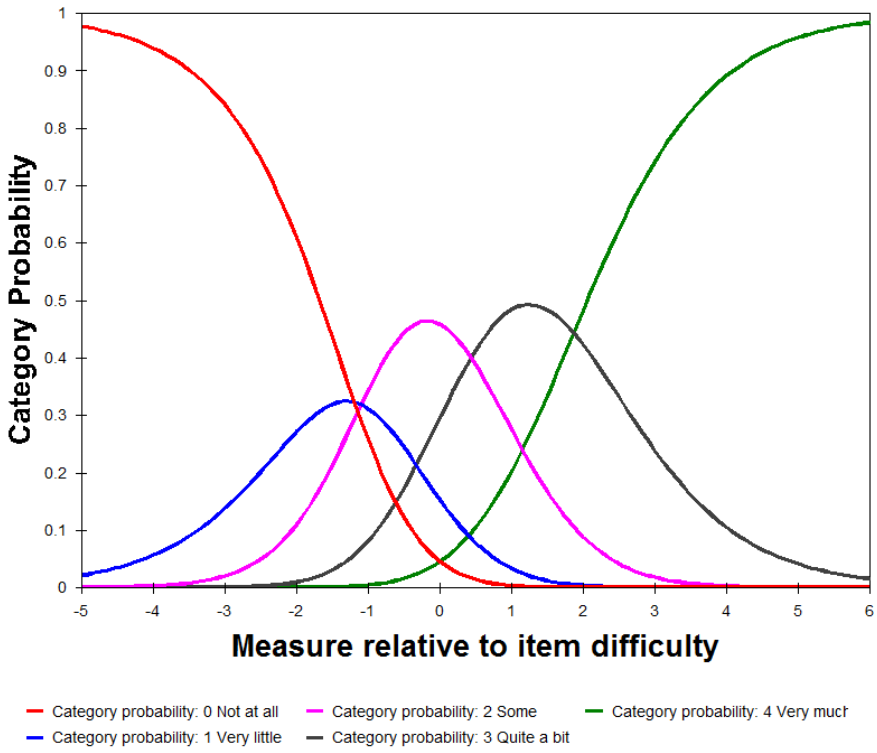


Figure 6. Category probabilities for Student Support

An important Rasch parameter is the "threshold" between two categories, shown in the figure where two probability curves intercept. For this scale, these thresholds are -1.20 logits (where the red "Not at all" curve intersects with the blue "Very little" curve), -1.10 logits (Very little/Some), 0.43 logits (Some/Quite a bit) and 1.87 logits (Quite a bit/Very much). The general guide for a five-category response scale is that thresholds should advance by at least 1.0 logits—gaps smaller than this indicate

a category that represents too narrow a segment of the latent variable, or a concept that is not well understood by respondents.[25] The two lowest thresholds here advance by only 0.1 logits so combining "Not at all" and "Very little" into a single category may be merited.

B.5 Item and person targeting

The final aspect of quality to be assessed is how well matched the items were to the persons being measured. Since the item difficulties and person measures are on the same scale, a simple side-by-side plot of the two sets of parameters will verify the targeting (see Figure 7).

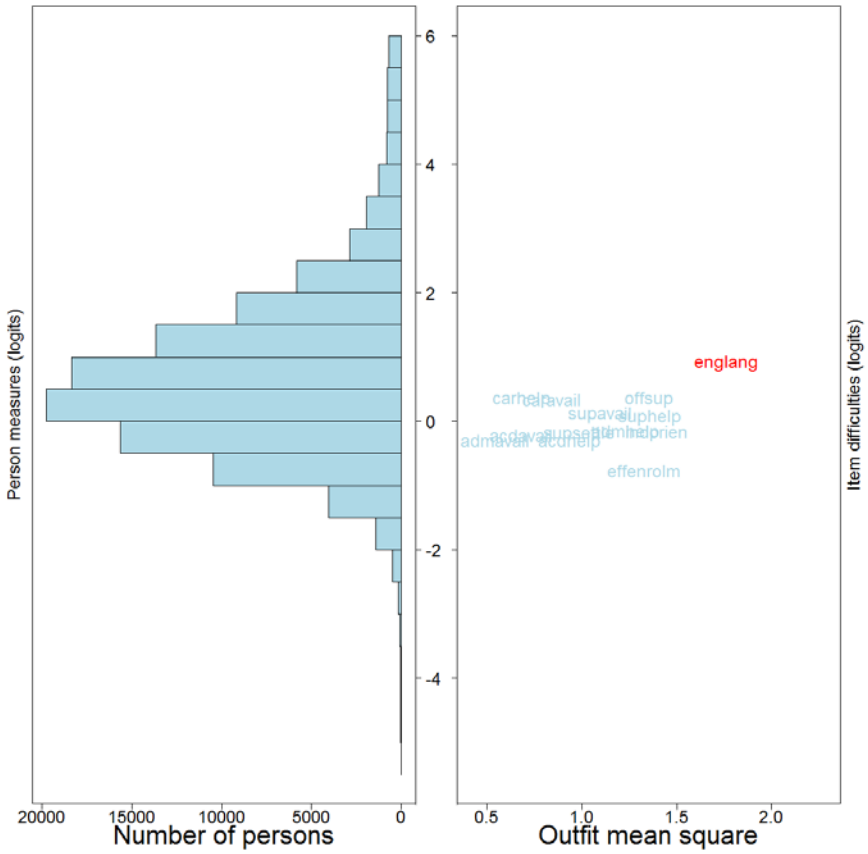


Figure 7. Person-item map for Student Support

This graph shows the distribution of person measures on the left and item difficulties on the right. The left side is organised so that persons with higher measures on student support are at the top and those with lower measures are at the bottom.[26] The items on the right side are arranged from the most difficult to endorse at the top to the easiest at the bottom, with the horizontal dimension representing item fit. The misfitting item (*englang*) is highlighted.

[25] Linacre, J.M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement*, *3*(2), 103-122. Retrieved from http://www.jampress.org/JOM_V3N2.pdf

[26] Note that the location of items has been adjusted slightly to minimise the extent of overlapping labels, so their positions may not exactly match the measures in the table of item statistics.

When items and persons are well matched, the two plots will show significant overlap, as they do here. Evident, however, is that the persons cover a greater range of the underlying trait than do the items—this means that respondents at the tails of the distribution (those who are low on the trait and those who are high) are measured with limited precision.[27] Measurement would be much improved for these respondents if there were additional items, some easier to endorse and some harder to endorse than the current items, that were a closer match to the respondents' locations on the trait.

A starting point would be to remove one or two items with low discrimination from the existing set and replace them with items that target the respondents with low precision. For instance, given that a number of respondents achieved the maximum score on this focus area,[28] it may be useful to develop items that are harder to endorse so these students can be separated.[29]

B.6 Summary

This appendix has presented a summary of results from a psychometric assessment of the student support focus area in the 2014 UES, following the omission of the *uniservices* item the instrument. The Rasch measurement model was used to calculate person scores and item difficulties and to determine how well items worked together to measure an underlying unidimensional attribute.

The key findings were as follows:

- There remained one misfitting item in the focus area (*englang*, "To what extent have you received appropriate English language skill support?").
- Person separation declined slightly from 2013 and there was a lack of items targeted at students at the tails of the underlying trait, resulting in low measurement precision for these respondents.
- There was very little progression between the difficulties of the two lowest rating scale categories ("Not at all" and "Very little") suggesting that respondents did not distinguish between the two.

Recommendations for future development of the student support focus area are as follows:

1. Based on its item fit statistics and on the large extent of missing data, the *englang* item should be omitted from the calculation of focus area scores and potentially from the instrument altogether.
2. The apparent lack of applicability of a number of items in the student support focus area requires an investigation into the theoretical foundation of its constituent items and potentially the development of replacement items that are more widely relevant to respondents.
3. The general lack of person separation and the low precision of estimates at the tails of the student distribution could also be remedied by the development of additional items.

[27] 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 0.03 logits with an accompanying standard error of 0.31 logits. Another respondent at the very top of the scale measured 5.70 logits, but this was associated with a standard error of 1.83 logits signifying much less certainty in their location on the underlying trait.

[28] In the field of educational testing, students who correctly answer all items on a test cannot be distinguished. Additional, harder items are required to separate these top performers.

[29] For more details on targeting of measurement instruments to respondents, see Bond, T.G., & Fox, M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. (2nd ed.) Mahwah, NJ: Erlbaum.

B.7 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.[30] 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:

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

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 $(-\infty, +\infty)$. The analysis reported here was conducted using the Winsteps software. Winsteps uses Joint Maximum Likelihood Estimation to solve the above equation.[31]

[30] Linacre 2014, *op. cit.*
[31] Wright, B.D., & Masters, G.N. (1982). *Rating Scale Analysis*. Chicago: MESA Press.

Appendix C: Course Experience Questionnaire (CEQ)

Table 29. CEQ items administered in the 2014 UES

Scale	Item[a]	Response scale
Good Teaching Scale	The staff put a lot of time into commenting on my work.	Strongly disagree / Disagree / Neither agree nor disagree / Agree / Strongly agree
	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.	

[a] R = Reverse coded for scoring purposes.

Appendix D: Analysis of the UEQ as administered to students from non-university higher education institutions

D.1 Introduction

As discussed in Section 1.2.1, the 2014 UES project included a pilot administration of the UES on students from non-university higher education institutions (NUHEIs) under the name of "Student Experience Survey" (SES). This appendix summarises a psychometric analysis of the UES items and focus areas as answered by NUHEI students as part of the SES pilot. The Rasch measurement model was used to assess the following aspects:

- How well the items in each focus area seemed to work 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, with a focus on comparing NUHEI and university respondents.

These aspects will be explored for each focus area in turn. Recommendations for refining the existing items and scales will be made.

The next section summarises three aspects of quality for each of the focus areas, namely item fit, category thresholds, and targeting. See Section B.2 in Appendix B for an overview of the analytical approach, including an explanation of these diagnostic outputs used in the assessment of quality.

D.2 Assessing item and focus area quality

D.2.1 Learner Engagement

This section summarises results for the learner engagement focus area, which contains the following items and response scale types.[32]

Variable	Item text	Scale
feelprepared	Felt prepared for study	Extent
interactdiff	Interacted with different students	Frequency
interactoth	Student interaction outside study	Frequency
oppploc	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

A summary of the item statistics for the learner engagement focus area is shown in Table 30. These are in order of item difficulty where *interactoth* ("How frequently have you interacted with students outside of study requirements?") was endorsed by relatively few students whereas *feelprepared* ("To what extent have you felt prepared for your study?") was endorsed by relatively many students.[33]

[32] Refer to Appendix A for a list of response scale categories.

[33] For instance, 23% of NUHEI students indicated that they interacted with other students "Very often" whereas 36% said they felt "Very much" prepared for their study. The first item was more difficult to endorse than the second.

There was no evidence of misfit for any of the items, with the various statistics remaining within acceptable bounds.

Also included in the table is the item difficulty for all university respondents (the last column, "Measure (UES)"). This enables differences in item difficulties for NUHEI and university students to be identified at a glance. More detailed checks for item differences are presented in Section D.3.

Table 30. NUHEI item statistics for Learner Engagement, with UES measures for comparison

Item	Measure	Mean square		Point-measure correlation	Item discrimination	Measure (UES)
		Infit	Outfit			
interactoth	0.62	1.01	1.00	0.68	1.04	0.64
interactdiff	0.22	0.90	0.92	0.63	1.11	0.31
workothers	0.09	0.80	0.78	0.69	1.28	-0.08
partidiscus	0.06	0.95	0.94	0.64	1.07	0.14
opploc	0.02	1.16	1.17	0.64	0.83	-0.32
sensebelong	-0.36	0.95	0.96	0.64	1.04	-0.10
feelprepared	-0.66	1.17	1.28	0.49	0.71	-0.60

Figure 8 shows the probabilities of NUHEI respondents selecting the various categories of the learner engagement items that use the "Extent" response scale (*opploc*, *sensebelong* and *feelprepared*). Respondents whose position is relatively low on the latent trait (on the left end of the x-axis) are most likely to select "Not at all" to these items. By contrast, respondents with high levels of the trait (on the right end of the x-axis) are most likely to select "Very much". Respondents in the middle will most likely select one of "Very little", "Some" or "Quite a bit", depending on their precise location on the trait.

An important Rasch parameter is the "threshold" between two categories, shown in the figure where two probability curves intercept. For this particular scale, these thresholds are -1.54 logits (Not at all/Very little), -0.76 (Very little/Some), 0.50 (Some/Quite a bit) and 1.79 (Quite a bit/Very much). The general guide for a five-category response scale is that thresholds should advance by at least 1.0 logits.[34] This is not the case for the two lowest thresholds and so combining adjacent categories may be merited.

As an aside, the uneven spacing of thresholds evident in this figure is in contrast to the assumption made when calculating focus area scores, namely that rating scale categories are equally-spaced cardinal numbers that can be used in arithmetic calculations. A particular strength of the Rasch model is that it avoids this arbitrary assignment of numbers to categories, instead estimating the most likely location on the underlying trait for persons, items and categories. It is recommended that focus area scores be calculated through the Rasch model in future. The maximum likelihood estimation procedures of the Rasch model also accommodate missing response in a much more robust way and also enable the calculation of standard errors.

[34] Linacre 1999, *op. cit.*

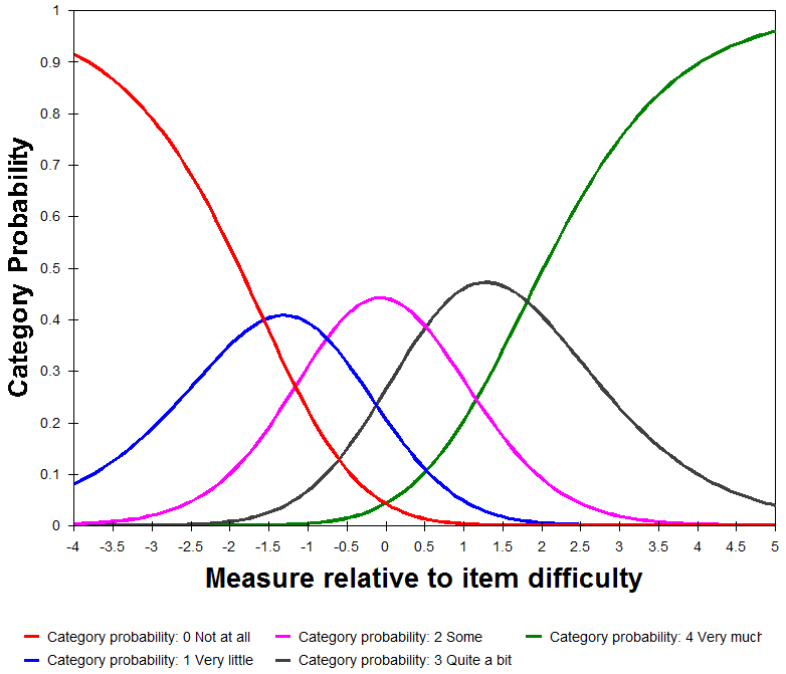


Figure 8. NUHEI category probabilities for "Extent" items, Learner Engagement

The corresponding graph for items using the "Frequency" scale (namely *interactoth*, *interactdiff*, *workothers* and *partidiscus*) is shown in Figure 9. For this scale, the thresholds are located at -1.73 logits (Never/Sometimes), 0.19 logits (Sometimes/Often) and 1.54 logits (Often/Very often). This progression meets the guidelines and so no collapsing of categories is indicated.

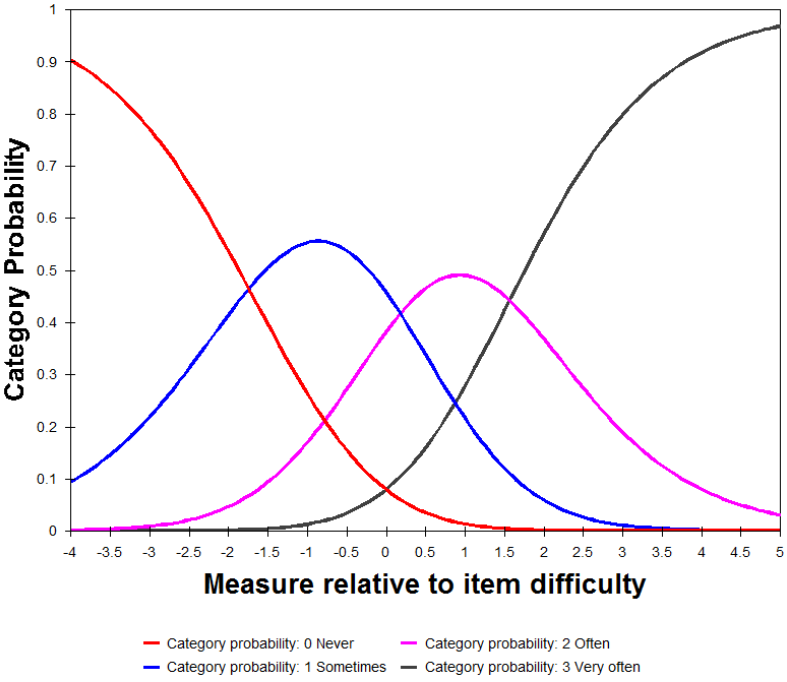


Figure 9. NUHEI category probabilities for "Frequency" items, Learner Engagement

The final aspect of quality to be assessed is how well matched the items were to the persons being measured. Since the item difficulties and person measures are on the same scale, a simple side-by-side plot—or Wright Map—of the two sets of parameters will verify the match (see Figure 10).

This graph shows the distribution of person measures on the left and the item difficulties on the right. The left side is organised so that persons with higher measures on the learner engagement focus area are at the top and those with lower measures are at the bottom. The items on the right side of the map are arranged from the most difficult to endorse at the top to the least difficult at the bottom.[35]

When items and persons are well matched, the two plots will show significant overlap, as they do here. Evident, however, is that the persons cover a greater range of the underlying trait than do the items—this means that the respondents at the tail of the distribution (those who are very low on the trait and those who are very high) are measured with limited precision.[36] 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 the respondents' locations on the trait.[37]

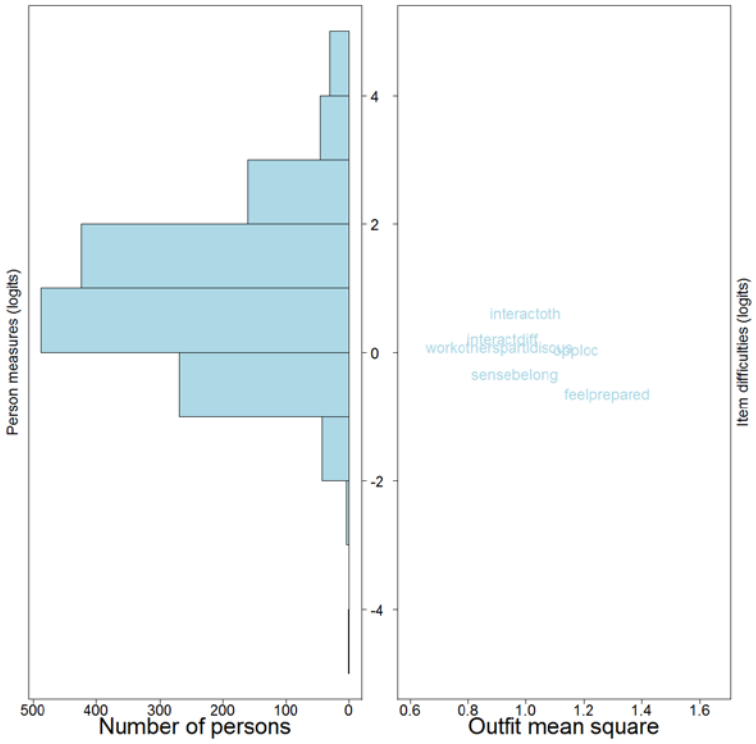


Figure 10. NUHEI person-item map for Learner Engagement

[35] Note that the location of items has been adjusted slightly to minimise the extent of overlapping labels, so the relative positions may not exactly match the measures in the table of item statistics.

[36] 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 0.16 logits with an accompanying standard error of 0.47 logits. Another respondent at the very top of the scale measured 4.84 logits, but this was associated with a standard error of 1.84 logits signifying much less certainty in their location on the underlying trait.

[37] For more details on targeting of measurement instruments to respondents, see Bond & Fox, *op. cit.*

D.2.2 Learning Resources

The learning resources focus area is measured by the following seven items.

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

A summary of fit statistics for this focus area is shown in Table 31, in order of most difficult to endorse (*qlstdspc*, "Student spaces and common areas") down to least difficult (*qltchspc*, "Teaching spaces"). No fit issues are evident for these items. Measures generally align well with those for the university respondents, apart from qltxtbook ("Assigned books, notes and resources") which was easier for non-university students to endorse than for university students.

Table 31. NUHEI item statistics for Learning Resources, with UES measures for comparison

Item	Measure	Mean square		Point-measure correlation	Item discrimination	Measure (UES)
		Infit	Outfit			
qlstdspc	0.47	1.07	1.07	0.73	0.92	0.41
qlcompit	0.34	0.96	0.96	0.74	1.04	0.15
qlequip	0.19	1.03	1.04	0.73	0.95	-0.08
qltxtbook	-0.11	0.94	0.96	0.71	1.06	0.40
qlonlmat	-0.21	1.01	1.02	0.70	0.98	-0.22
qllibres	-0.31	1.07	1.04	0.71	0.95	-0.51
qltchspc	-0.36	0.89	0.89	0.72	1.11	-0.15

The category probabilities for this scale are shown in Figure 11. The thresholds are located at -2.11 logits (Poor/Fair), -0.40 logits (Fair/Good), and 2.52 logits (Good/Excellent). The separation between these meets the guidelines so that no collapsing of categories is warranted.

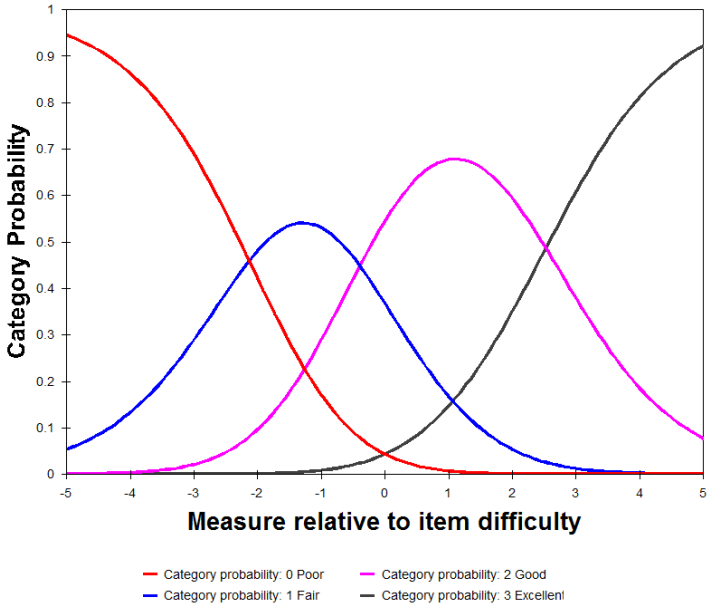


Figure 11. NUHEI category probabilities for Learning Resources

The person-item map for the learning resources focus area is shown in Figure 12. The issue of inadequate targeting is again evident, 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.

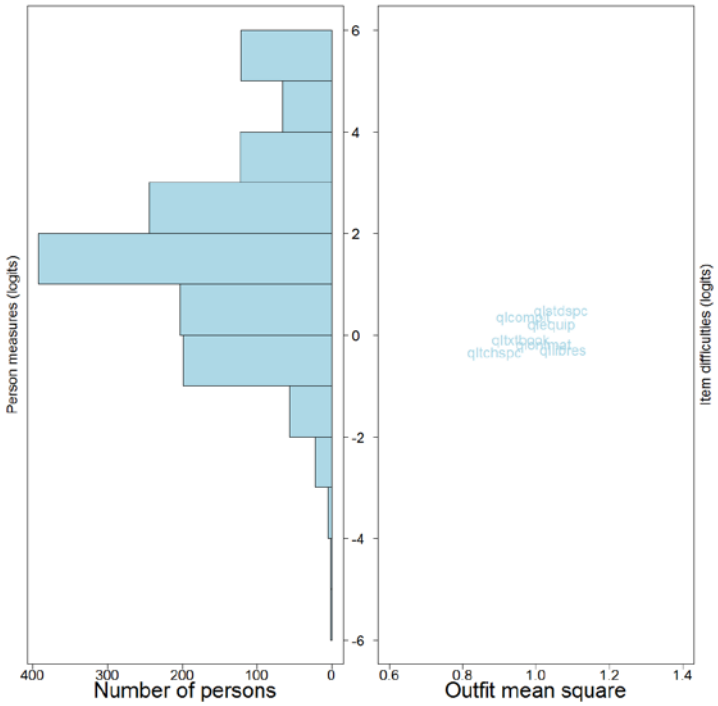


Figure 12. NUHEI person-item map for Learning Resources

D.2.3 Skills Development

The skills development focus area is constituted by the following eight items.

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 Rasch summary statistics are shown in Table 32. The most difficult item for respondents to endorse was *expprbslv* ("To what extent has your course developed your ability to solve complex problems?") and the easiest was *expknowl* ("To what extent has your course developed your knowledge of the fields you are studying?")—"Very much" was selected by fewer than a quarter of respondents for *expprbslv* but by almost half for *expknowl*.

Table 32. NUHEI item statistics for Skills Development, with UES measures for comparison

Item	Measure	Infit	Outfit	Point-measure correlation	Item discrimination	Measure (UES)
expprbslv	0.63	0.78	0.79	0.81	1.22	0.29
expwrkoth	0.46	1.32	1.33	0.74	0.69	0.24
expspeak	0.45	1.12	1.12	0.76	0.88	0.68
expwriting	0.10	1.04	1.05	0.75	0.95	0.15
expthink	-0.15	0.79	0.80	0.78	1.21	-0.30
expwrkskill	-0.21	0.95	0.94	0.76	1.06	0.16
expconfind	-0.30	0.98	0.99	0.76	1.02	-0.41
expknowl	-0.99	0.94	1.01	0.71	1.04	-0.83

The modelled probability curves are shown in Figure 13. For this graph, the thresholds are adequately spaced: -3.11 logits (Not at all/Very little), -1.47 logits (Very little/Some), 0.80 logits (Some/Quite a bit), and 3.78 logits (Quite a bit/Very much).

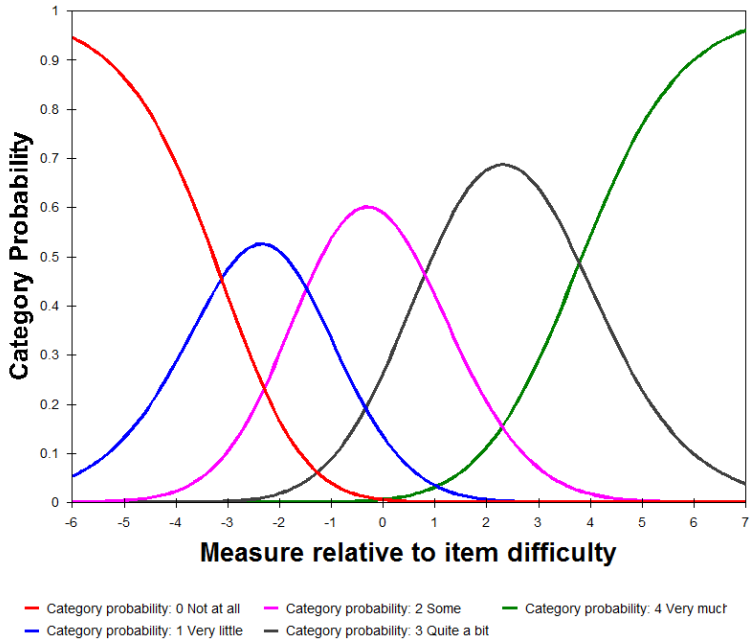


Figure 13. NUHEI category probabilities for Skills Development

The person-item map for NUHEI students on the Skills Development focus area is shown in Figure 14, where the need for additional items to distinguish between the top-scoring students is again apparent.

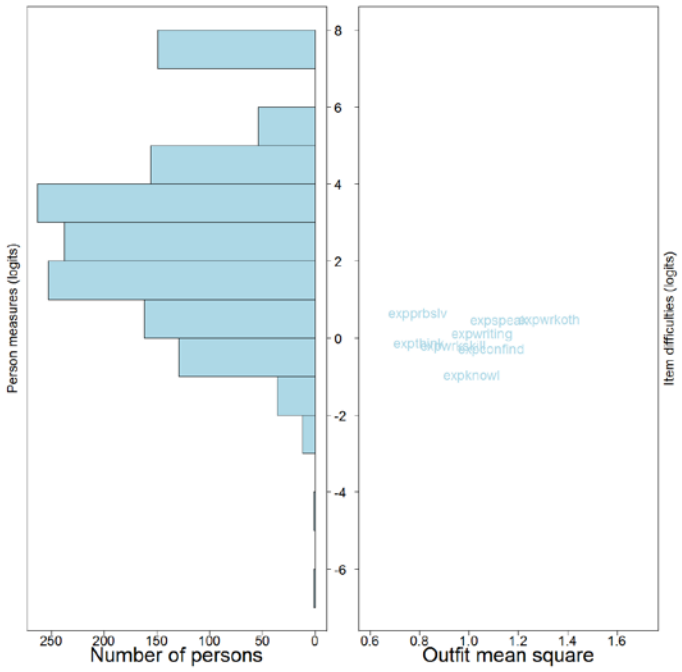


Figure 14. NUHEI person-item map for Skills Development

D.2.4 Student Support

The student support focus area in 2014 consists of the following thirteen items.

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

The Rasch summary statistics are shown in Table 33. The most difficult item for respondents to endorse was *caravail* ("To what extent have you found careers advisors to be available?") and the easiest was *effenrolm* ("To what extent have you experienced efficient enrolment and admissions processes?"). The high mean square and low discrimination values for *englang* ("To what extent have you received appropriate English language support?") suggest that it was answered in a very unpredictable way by NUHEI respondents. Including this item may degrade the quality of scores for this focus area, especially since fewer than half of all respondents provided a rating for the item. Despite the generally satisfactory Rasch statistics, the amalgamation into one focus area of items rating both the availability and helpfulness of services is problematic—calculating an aggregate score from such disparate, unevenly answered items has questionable validity and value.

Table 33. NUHEI item statistics for Student Support, with UES measures for comparison

Item	Measure	Mean square		Point-measure correlation	Item discrimination	Measure (UES)
		Infit	Outift			
caravail	0.64	0.90	0.94	0.70	1.04	0.34
suphelp	0.59	0.90	0.90	0.72	1.12	0.06
supavail	0.54	0.85	0.85	0.72	1.16	0.11
carhelp	0.51	1.01	1.01	0.68	0.96	0.35
englang	0.26	1.91	1.90	0.57	0.10	0.91
offsup	-0.01	1.35	1.34	0.65	0.72	0.35
admhelp	-0.14	0.87	0.84	0.69	1.16	-0.17
acdavail	-0.17	0.67	0.67	0.71	1.32	-0.22
admavail	-0.25	0.73	0.71	0.71	1.28	-0.29
acdhelp	-0.33	0.74	0.71	0.69	1.27	-0.32
indorien	-0.35	1.09	1.22	0.61	0.84	-0.16
supsettle	-0.59	0.91	0.99	0.64	1.07	-0.20
effenrolm	-0.72	1.32	1.34	0.56	0.70	-0.77

The category probability curves are shown in Figure 15. The thresholds are located at -1.40 logits (Not at all/Very little), -0.96 logits (Very little/Some), 0.51 logits (Some/Quite a bit) and 1.85 logits (Quite a bit/Very much). Based on this spacing, it may be useful to collapse the "Very little" category into one of the adjacent categories.

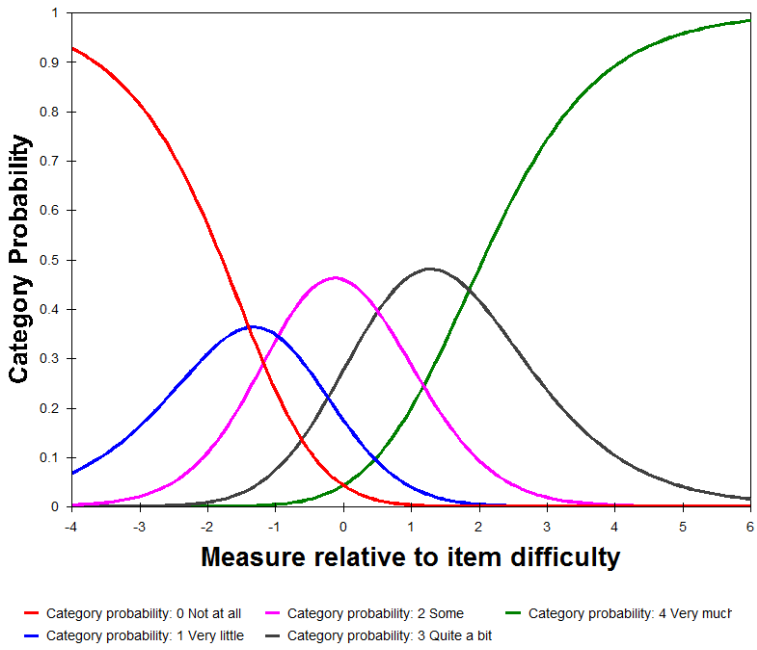


Figure 15. NUHEI category probabilities for Student Support

The person-item map for the student support focus area is shown in Figure 16. Note that the location of items has been adjusted slightly to minimise the extent of overlapping labels, so their relative positions may not exactly match the measures in Table 33. While the item difficulties cover a broader range than for other focus areas, a number of students are being measured only approximately by the present instrument. The misfitting item (*englang*) is highlighted.

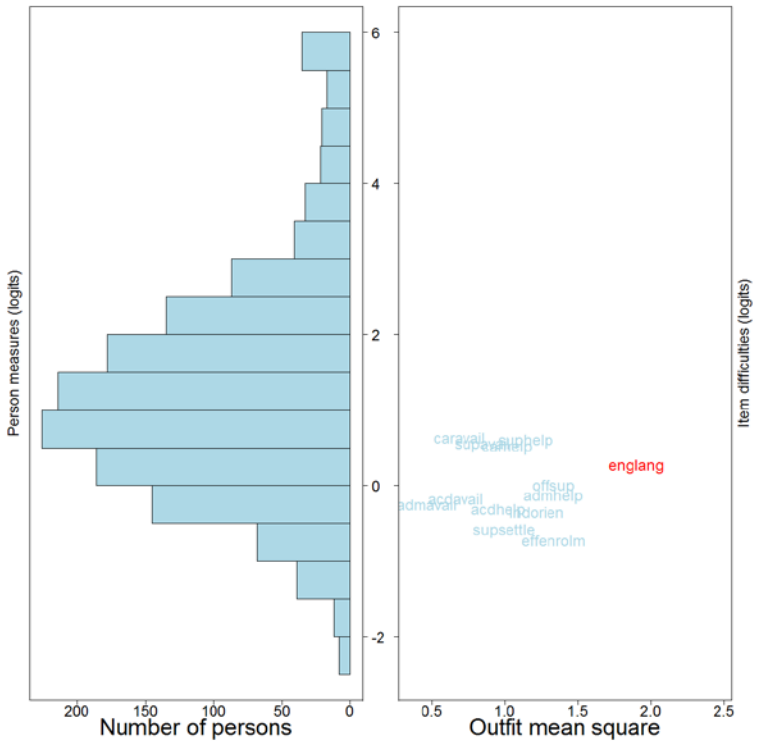


Figure 16. NUHEI person-item map for Student Support

D.2.5 Teaching Quality

The final focus area, teaching quality, is measured by the following eleven 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 Rasch fit statistics for the focus area are summarised in Table 34. The most difficult item to endorse was *qlovledu* ("Overall, how would you rate the quality of your entire educational experience this year?") and the easiest was *tchhelpapp* ("To what extent have the lecturers, tutors and demonstrators seemed helpful and approachable?"). While no items demonstrated signs of misfit, the inclusion of an overall rating item such as *qlovledu* among a mix of more specific items is potentially problematic. It is recommended that scores for this focus area be calculated without *qlovledu* in future.

Table 34. NUHEI item statistics for Teaching Quality, with UES measures for comparison

Item	Measure	Mean square		Point-measure correlation	Item discrimination	Measure (UES)
		Infit	Outfit			
qlovledu	0.44	1.03	1.11	0.72	0.95	0.30
tchfeedbck	0.38	1.17	1.18	0.74	0.81	0.85
tchclexpec	0.36	1.06	1.04	0.76	0.95	0.08
stdstruc	0.23	0.86	0.86	0.78	1.15	-0.03
qlteach	0.20	0.89	0.89	0.75	1.11	0.20
tchactiveng	0.07	0.79	0.80	0.79	1.22	0.12
tchconlrn	0.06	1.13	1.11	0.75	0.88	0.31
tchstimint	-0.08	0.87	0.86	0.78	1.14	-0.22
stdrelev	-0.46	1.05	1.09	0.72	0.94	-0.46
tchasschlng	-0.53	1.00	1.01	0.72	0.99	-0.67
tchhelpapp	-0.67	1.12	1.03	0.73	0.93	-0.50

The category probability curves for the nine items on the "Extent" scale are shown in Figure 17. The thresholds are located at -3.25 logits (Not at all/Very little), -1.54 logits (Very little/Some), 0.99 logits (Some/Quite a bit) and 3.81 logits (Quite a bit/Very much). These thresholds are adequately spaced so that no categories need to be combined for this focus area.

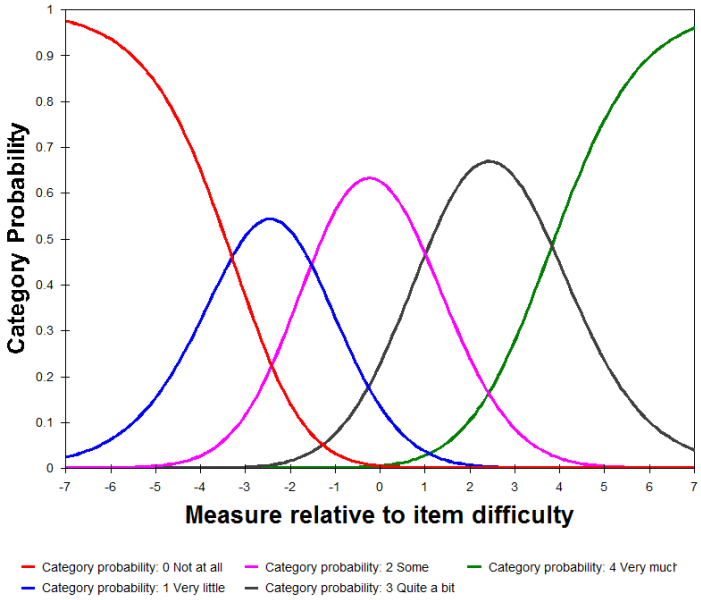


Figure 17. NUHEI category probabilities for "Extent" scale items, Teaching Quality

The category probabilities for the two items using the "Rating" scale are shown in Figure 18. The thresholds are located at -3.14 logits (Poor/Fair), -0.32 logits (Fair/Good) and 3.46 logits (Good/Excellent) and are adequately spaced.

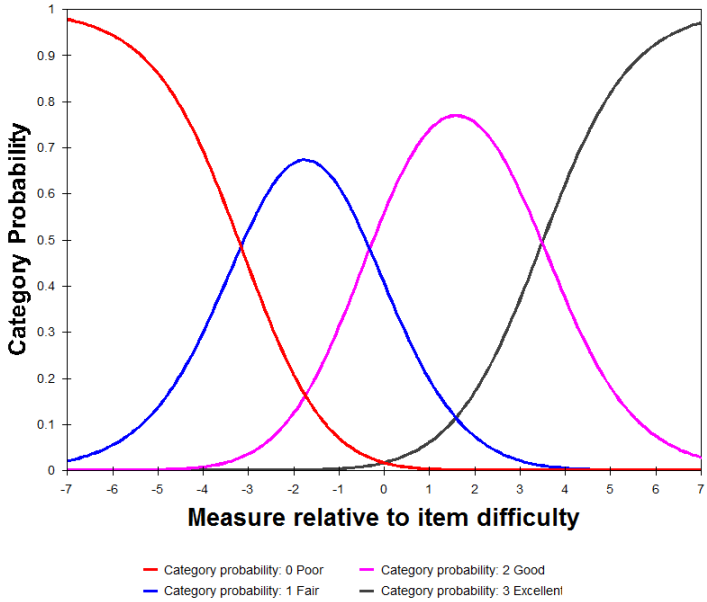


Figure 18. NUHEI category probabilities for Rating scale items, Teaching Quality

The person-item map in Figure 19 shows that items cover a very narrow range of the teaching quality trait compared to persons, suggesting that more items are needed to measure persons accurately for this focus area.

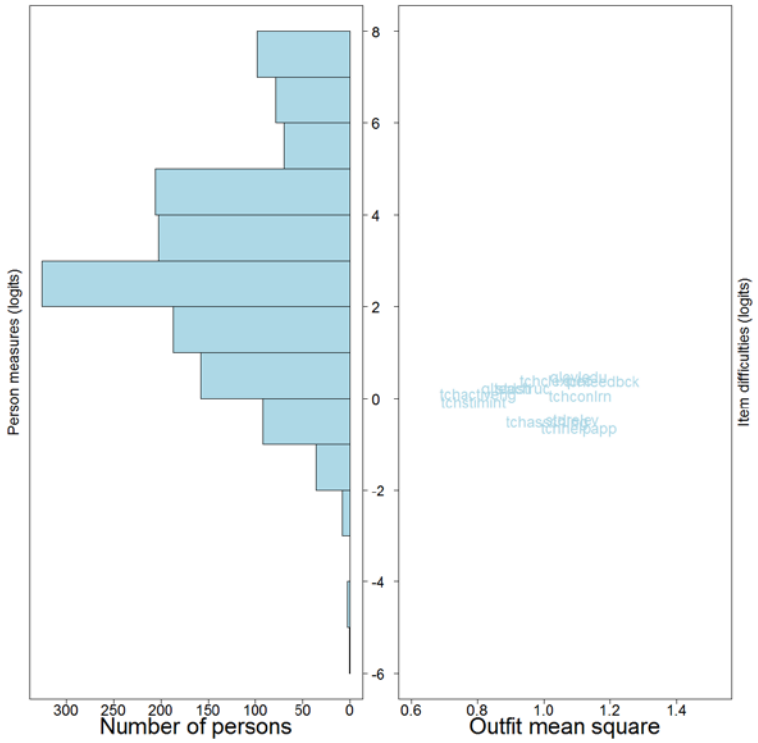


Figure 19. NUHEI person-item map for Teaching Quality

D.2.6 Focus area summaries

To complement the individual item results given in the preceding sections, Table 35 presents a summary of fit statistics as they apply to each of the five focus areas as a whole. The statistics presented are as follows:

- Person separation, which indicates how well a set of items is able to distinguish between the persons measured.
- Item separation, which indicates how well a sample of persons is able to separate the items on an instrument.
- Cronbach's alpha, which is an index of repeatability of the order of persons as measured by the set of items.

The separation indices are Rasch outputs and should be no less than 2 for persons and no less than 3 for items.[38] Cronbach's alpha, included for completeness, is a widely-reported correlation deriving from classical test theory and for which higher values are taken to be indicative of better reliability and consistency for a set of items.

Table 35. Overall NUHEI summary statistics, by focus area

Focus area	Person separation	Item separation	Cronbach's alpha[39]
Learner Engagement	1.79	10.39	0.78
Learning Resources	1.89	5.93	0.89
Skills Development	2.47	10.22	0.91
Student Support	2.15	10.83	0.94
Teaching Quality	2.76	7.40	0.95

The highlighted values in for learner engagement, learning resources and skills development indicate that the sets of items for these focus areas are not sensitive enough to distinguish between high (and low) scoring respondents. The other two focus areas do not have particularly strong person separation indices either, as was already evident from the person-item maps, so all five would benefit from the inclusion of more targeted items. A reasonable approach would be to omit the item with the lowest separation index for each focus area, and then to develop one or two new items that are targeted at the tails of the distribution.

The item separation indices are adequate and indicate that sufficient person responses were obtained to confirm the hierarchy of item difficulties with respect to the underlying trait.

[38] Linacre 1999, *op. cit.*

[39] Cases with missing data are dropped from the calculation of Cronbach's alpha so these values should be seen as only approximate. For a discussion on the limitations of this statistic for assessing instrument quality, refer to Sijtsma, K. (2009). On the use, the misuse and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107-120.

D.3 Differential item functioning

In the Rasch model, the probability of a respondent choosing a particular category for an item is related only to the difference between the person's and the item's positions on the underlying trait (see Section B.7 in Appendix B for details). It is not expected that systematic differences between an item's modelled and actual responses should occur for any particular subgroup of respondents.

Differential item functioning (DIF) investigates the items on an instrument, one at a time, for signs of interaction with respondent characteristics. The process calculates the item difficulty separately for each characteristic and identifies if one group of respondents scored higher than another group of respondents on an item, after adjusting for the overall scores of the respondents. The analysis cannot identify the reason for the discrepancies, but does flag items for closer scrutiny by item developers to ensure that including an item on a questionnaire does not "disadvantage" any particular group—that is, the items do not distort the measures of this group.

A DIF analysis was carried out on an extensive set of respondent characteristics. NUHEI respondents are compared with university respondents in Section D.3.1. DIF effects for subgroups of NUHEI respondents are investigated in Section D.3.2.

D.3.1 NUHEI and UES

In comparing NUHEI and UES item difficulties, it was desirable to minimise the impact on DIF of the heterogeneous nature of the UES cohort—namely, that any observed DIF effects should more likely be due to a different capability or understanding of item by NUHEI students than to NUHEI students being different from UES students in their basic characteristics. To achieve this, a subset of UES students was selected that matched as closely as possible the characteristics of the NUHEI students. The DIF analysis was then performed on the NUHEI students and the matched UES students.

For each NUHEI respondent, the closest matching UES student was identified. Where there were multiple matching students, one match was selected at random. The method is referred to as nearest neighbour matching, implemented in the *Matchit* package in R.[40] This process resulted in 1,494 NUHEI students and 1,494 matched UES students.

A difference in item calibrations of more than 0.5 logits is typically taken as suggestive of substantive DIF effects. For the NUHEI and matched UES students, only one item (*englang*, "To what extent have you received appropriate English language skill support") demonstrated a difference in item difficulty of this magnitude (see Table 36). The item with the next largest difference is included for comparison purposes (*suphelp*, "To what extent have you found support services such as counsellors, financial/legal advisors and health services to be helpful").

A positive DIF measure implies that the item was unexpectedly difficult for NUHEI students to endorse compared to matched UES students. A negative DIF measure implies that the item was unexpectedly easy for NUHEI students to endorse compared to matched UES students.

[40] Ho, D.E., Imai, K., King, G., & Stuart, E.A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, *42*(8), 1-28.

In this table, *suphelp* was endorsed less often by NUHEI respondents, compared to matched UES respondents, whereas *englang* was endorsed more often. Given the extent of misfit for *englang*, as described in Section D.2.4, its large DIF measure may not be relevant. For *suphelp*, though, questionnaire developers may like to consider if there are reasons for this unexpected difference between the two groups. If not, the item may be a candidate for revision or omission.

Table 36. Notable DIF effects, NUHEI versus matched UES students

Item	DIF measure (NUHEI—UES)
Student Support	
englang	-0.51
suphelp	0.43

D.3.2 NUHEI subgroups

A DIF analysis was also undertaken for subgroups of NUHEI respondents with respect to their demographic and enrolment characteristics. Notable item differences are presented for learner engagement items in Table 37, learning resources items in Table 38, skills development items in Table 39, student support items in Table 40, and teaching quality items in Table 41.

For example, Table 37 shows learner engagement items with large DIF effects for the different respondent characteristics. Respondents in the 30 to 39 years and the 40 years and over age groups were less likely to endorse *interactoth* ("How frequently have you interacted with students outside study requirements") than younger respondents. Similarly, respondents whose study location was external were less likely to endorse *workothers* ("How frequently have you worked with other students as part of your study") than internal students. So while older students and external students found these items more difficult to endorse than other students, it might be expected that they have less need or less opportunity to interact or work with other students, in which case the DIF effects are not surprising. The same may be said for some of the subject areas identified in this table. Careful scrutiny by item developers is required to determine if any items merit refinement or omission in future.

Table 37. NUHEI characteristics, items and DIF measures for Learner Engagement

Characteristic	Subgroup	Item	DIF measure
Age	30 to 39 years	interactoth	0.59
	40 years and over		0.82
Campus	External	workothers	0.49
Subject area	Accounting	workothers	-0.56
	Architecture & Forestry	interactoth	-1.05
		opploc	0.57
	Communication, Media & Journalism	partidiscus	-0.52
	Humanities inc History & Geography	workothers	0.50
	Music & Performing Arts	feelprepared	0.62
		workothers	-0.50
	Social Work	interactoth	0.64
	Teacher Education – Primary & Secondary	feelprepared	0.63
		interactoth	-0.62

Table 38. NUHEI characteristics, items and DIF measures Learning Resources

Characteristic	Subgroup	Item	DIF measure
Course type	Associate degree	qltxtbook	0.54
	Diploma	qlcompit	-0.53
Gender	Female	qllibres	0.51

Table 38. (continued)

Characteristic	Subgroup	Item	DIF measure
Subject area	Accounting	qlstdspc	-0.85
	Architecture & Forestry	qlcompit	-1.26
		qlequip	0.65
		qllibres	-0.61
		qlonlmat	0.75
	Art & Design	qlcompit	-0.65
		qlequip	-1.04
		qlonlmat	0.88
	Communication, Media & Journalism	qllibres	-1.04
	Humanities inc History & Geography	qlequip	0.58
	Music & Performing Arts	qllibres	-1.40
		qltxtbook	0.61
	Nursing	qlcompit	0.89
		qlequip	-1.34
		qltchspc	-0.63
	Social Work	qlonlmat	-0.69
		qltxtbook	-0.86
	Teacher Education – Early Childhood	qllibres	0.79
	Teacher Education – Other	qltxtbook	-0.62
	Teacher Education – Primary & Secondary	qllibres	-0.59

Table 39. NUHEI characteristics, items and DIF measures for Skills Development

Characteristic	Subgroup	Item	DIF measure
Age	30 to 39 years	expwrkoth	0.62
Attendance type	Part-time	expwrkoth	0.78
Campus	External	expspeak	0.61
		expwrkoth	0.64
Disability	Reported disability	expknowl	-0.66
		expspeak	0.62
Language spoken at home	Language other than English	expknowl	0.51
Subject area	Accounting	expthink	0.52
	Art & Design	expwriting	0.71
	Business Management	expknowl	0.58
	Computing & Information Systems	expknowl	0.62
	Humanities inc History & Geography	expknowl	-0.57
		expspeak	0.75
		expthink	-0.54
		expwriting	-0.81
		expwrkoth	0.79
	Music & Performing Arts	expwriting	0.80
		expwrkoth	-0.78
	Nursing	expwriting	0.51
	Teacher Education – Early Childhood	expprbslv	0.59
		expwrkskill	-0.55
	Teacher Education – Primary & Secondary	expspeak	-0.55
		expwrkoth	-0.53

Table 40. NUHEI characteristics, items and DIF measures for Student Support

Characteristic	Subgroup	Item	DIF measure
Age	40 years and over	effenrolm	-0.50
		englang	0.79
		offsup	0.55
Disability	Reported disability	offsup	0.50
Subject area	Accounting	offsup	0.60
	Communication, Media & Journalism	admavail	0.53
		effenrolm	0.54

Table 40. (continued)

Characteristic	Subgroup	Item	DIF measure
	Humanities inc History & Geography	englang	0.55
	Music & Performing Arts	indorien	0.60
		supsettle	0.52
	Social Work	effenrolm	-0.55
	Teacher Education – Other	offsup	-0.67

Table 41. NUHEI characteristics, items and DIF measures for Teaching Quality

Characteristic	Subgroup	Item	DIF measure
Disability	Reported disability	tchhelpapp	0.56
International	International student	qlovledu	0.51
		qlteach	0.60
		tchclexpec	-0.56
Previous university experience	Previously enrolled in a different course at the current university	tchhelpapp	0.51
Subject area	Accounting	qlteach	0.90
		tchactiveng	-0.59
	Architecture & Forestry	qlovledu	-0.56
	Art & Design	tchconlrn	0.50
	Communication, Media & Journalism	qlteach	-0.55
		stdstruc	0.80
		tchactiveng	-0.55
		tchfeedbck	-0.86
	Computing & Information Systems	qlovledu	0.70
		qlteach	0.73
		tchclexpec	-0.51
		tchfeedbck	-1.12
		tchstimint	0.56
	Humanities inc History & Geography	tchstimint	-0.50
	Music & Performing Arts	tchclexpec	0.62
	Teacher Education – Early Childhood	tchclexpec	0.75
	Teacher Education – Other	tchclexpec	0.64

D.4 Missing responses

The final analysis undertaken here is the extent to which items were skipped or could not be answered by NUHEI students, compared to university students. The Rasch model naturally accommodates such missing responses, but measures derived from items that are missing many responses may be unstable and associated with larger standard errors.[41] Neither of these outcomes are desirable and the validity of including affected items needs to be considered carefully.

The percentage of missing responses for each item is presented in Table 42 for NUHEI and for UES respondents. In this context, a response was missing if it was anything other than a rating scale category (that is, if it was “Not asked”, “Not applicable”, “Do not know answer/Refused item”, “Item skipped”, and so on).

[41] Linacre 2014, *op. cit.*

Table 42. Summary of missing responses, by NUHEI and UES students

Item	Missing responses (%)	
	NUHEI	UES
Learner Engagement		
feelprepared	0.41	0.20
interactdiff	0.48	0.16
interactoth	0.20	0.15
opploc	4.02	3.67
partidiscus	0.41	0.15
sensebelong	0.07	0.12
workothers	0.34	0.14
Learning Resources		
qlcompit	10.16	10.41
qlequip	33.40	31.10
qllibres	5.59	6.29
qlonlmat	4.98	3.56
qlstdspc	10.02	10.13
qltchspc	9.07	7.86
qltxtbook	4.70	4.35
Skills Development		
expconfind	0.95	1.32
expknowl	0.95	1.31
expprbslv	1.16	1.31
expspeak	1.02	1.32
expthink	0.95	1.30
expwriting	0.95	1.31
expwrkoth	0.89	1.32
expwrkskill	1.09	1.36
Student Support		
acdavail	12.47	28.04
acdhelp	12.82	28.43
admavail	4.50	16.84
admhelp	4.70	16.97
caravail	41.65	59.38
carhelp	42.67	60.21
effenrolm	0.20	0.20
englang	52.69	64.24
indorien	7.50	10.82
offsup	21.27	35.15
supavail	48.19	67.07
suphelp	48.94	68.02
supsettle	0.41	0.22
Teaching Quality		
qlovledu	0.14	0.07
qlteach	0.27	0.22
stdrelev	0.75	0.99
stdstruc	0.95	1.13
tchactiveng	0.89	0.76
tchasschlng	0.89	0.76
tchclexpec	0.95	0.72
tchconlrn	0.68	0.75
tchfeedbck	0.89	0.78
tchhelpapp	0.89	0.74
tchstimint	0.95	0.79

It is evident that a number of items did not apply or were not answered by many students, especially in the student support focus area. The extent of "missingness" in this focus area was less for NUHEI respondents than for UES respondents, however.

The item *qlequip* ("Laboratory or studio equipment") could safely be omitted from the learning resources focus area without significant impact on scores.

The extent of missingness for items in student support does bring into question the validity of the calculated scores and the aspects of student experience the focus area represents. Dropping affected items would profoundly influence the scores for this focus area, and this action is not recommended as a remedy. Instead, the theoretical basis of each item needs to be re-examined and its membership in the focus area confirmed. It may be possible to develop replacement items that are relevant to a larger proportion of the population and that capture the intent of the student support focus area. This is a longer-term undertaking, though, requiring qualitative research and pilot testing before administration to the whole population.

D.5 Summary

This appendix has presented a summary of results from a psychometric assessment of NUHEI responses to the 2014 University Experience Survey. The Rasch measurement model was used to calculate person scores and item difficulties across the five focus areas, and to determine how well items within each focus area seemed to be working together to measure an underlying unidimensional attribute.

The key findings were as follows:

- Most items seemed to meet the Rasch model's assumptions adequately. One item in the student support focus area (*englang*, “To what extent have you received appropriate English language skill support?”) exhibited poor fit statistics, however.
- A comparison of item responses between NUHEI respondents and “like” UES respondents revealed two items (*suphelp*, “To what extent have you found support services such as counsellors, financial/legal advisors and health services to be helpful”, and *englang*) that were answered very differently by NUHEI respondents. Within NUHEI students, some subgroups answered items in notably different ways, with most differences occurring across subject areas.
- Most focus areas manifested a lack of items targeted at students at the tails of the underlying trait, resulting in low measurement precision for these respondents.

Recommendations for future development and administration of the UES are as follows:

1. Based on its item fit statistics and on the large extent of missing data, the *englang* item should be omitted from the calculation of scores and potentially from the instrument altogether.
2. Items with notably different calibrations among subgroups of respondents should be investigated carefully to determine the reasons for the differences and to ensure that focus area scores are not being skewed by the inclusion of these items.
3. The *qlequip* item in the learning resources focus area should be considered for omission from the calculation of scores due to its high level of missing responses.
4. In the longer term:
    a. The apparent lack of applicability to students of a number of items in the student support focus area, for both university and NUHEI students, requires an investigation into the theoretical foundation of its constituent items and potentially the development of items that are more relevant to respondents.

b. The general lack of precision at the tails of the student distributions suggest that additional items are needed if it is desired to measure these respondents more accurately. The least sensitive items can be omitted on statistical grounds but the construction of new items requires a return to first principles to determine the attributes that will enable greater discrimination between high (and low) scoring students.

Appendix E: 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 20.

```
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 20. 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 21. 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 22.

```
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, admavailr, admhelpr, caravailr, carhelpr,
acdavailr, acdhelpr, supavailr, suphelpr, effenrolmr).
COMPUTE RESOURCES=MEAN.5(qltchspcr, qlstdspcr, qlonlmatr, qlcompitr, qltxtbookr, qlequipr, qllibresr).
```

Figure 21. 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 22. 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 23.

```
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 23. SPSS syntax used to compute item satisfaction variables

Appendix F: Promotional website tiles

Figure 24. UES 2014 promotional website tiles

Appendix G: Construction of confidence intervals

The 90 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 2014 UES, the formula for the confidence interval of a proportion is used.

$$95\%\ CI\ bound(\hat{p}) = 1.645 \times FPC \times SE(\hat{p}) = 1.645 \times \sqrt{\frac{N-n}{N-1}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

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 90 per cent confidence interval of each estimated proportion is then calculated as the proportion plus or minus its 90 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 H: Response category percentages

Table 43. 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	4
	Some	27	20	24
	Quite a bit	45	44	45
	Very much	23	32	26
Developed ability to solve complex problems	Not at all	1	1	1
	Very little	7	6	7
	Some	34	27	31
	Quite a bit	41	43	42
	Very much	16	24	19
Developed ability to work effectively with others	Not at all	2	2	2
	Very little	9	6	8
	Some	30	24	28
	Quite a bit	39	40	40
	Very much	19	27	22
Developed confidence to learn independently	Not at all	1	1	1
	Very little	5	4	5
	Some	24	18	22
	Quite a bit	43	42	42
	Very much	26	35	30
Developed written communication skills	Not at all	2	1	1
	Very little	8	6	7
	Some	32	22	28
	Quite a bit	40	41	41
	Very much	18	29	22
Developed spoken communication skills	Not at all	3	3	3
	Very little	13	9	11
	Some	34	27	32
	Quite a bit	34	38	36
	Very much	15	23	18
Developed knowledge of field studying	Not at all	0	1	1
	Very little	3	3	3
	Some	19	16	18
	Quite a bit	45	42	44
	Very much	33	37	35
Developed work-related knowledge and skills	Not at all	1	2	2
	Very little	7	8	8
	Some	29	26	28
	Quite a bit	42	39	41
	Very much	20	26	22

Table 44. Learner Engagement item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Felt prepared for your study	Not at all	1	1	1
	Very little	7	5	6
	Some	28	24	27
	Quite a bit	42	43	43
	Very much	21	26	23
Had a sense of belonging to your university	Not at all	2	3	3
	Very little	10	14	12
	Some	33	33	33
	Quite a bit	38	33	36
	Very much	17	16	17
Participated in discussions online or face-to-face	Never	8	7	7
	Sometimes	36	31	34
	Often	36	37	36
	Very often	21	26	23

Table 44. (continued)

Worked with other students as part of your study	Never	7	6	7
	Sometimes	31	26	29
	Often	37	35	36
	Very often	25	33	28
Interacted with students outside study requirements	Never	16	15	16
	Sometimes	38	37	38
	Often	27	27	27
	Very often	18	20	19
Interacted with students who are very different from you	Never	7	8	7
	Sometimes	38	39	38
	Often	36	35	36
	Very often	19	19	19
Been given opportunities to interact with local students	Not at all	3	4	4
	Very little	11	12	11
	Some	28	27	28
	Quite a bit	33	30	32
	Very much	25	27	26

Table 45. Teaching Quality item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Quality of entire educational experience	Poor	2	4	3
	Fair	15	19	17
	Good	57	55	56
	Excellent	26	22	25
Quality of teaching	Poor	2	4	3
	Fair	16	19	17
	Good	54	52	53
	Excellent	28	26	27
Teachers engaged you actively in learning	Not at all	1	1	1
	Very little	6	7	6
	Some	29	29	29
	Quite a bit	45	42	44
	Very much	20	20	20
Teachers demonstrated concern for student learning	Not at all	1	2	2
	Very little	7	9	8
	Some	30	30	30
	Quite a bit	41	39	40
	Very much	20	20	20
Teachers provided clear explanations on coursework and assessment	Not at all	1	1	1
	Very little	6	8	7
	Some	27	29	28
	Quite a bit	44	42	43
	Very much	22	20	21
Teachers stimulated you intellectually	Not at all	1	1	1
	Very little	5	6	5
	Some	25	26	25
	Quite a bit	45	43	44
	Very much	25	24	25
Teachers commented on your work in ways that help you learn	Not at all	2	3	3
	Very little	13	12	12
	Some	34	33	34
	Quite a bit	35	35	35
	Very much	16	17	16
Teachers seemed helpful and approachable	Not at all	1	1	1
	Very little	4	5	4
	Some	23	23	23
	Quite a bit	42	41	42
	Very much	30	30	30

Table 45. (continued)

Teachers set assessment tasks that challenge you to learn	Not at all	1	1	1
	Very little	3	4	3
	Some	18	20	19
	Quite a bit	50	47	49
	Very much	29	27	28
Study well structured and focused	Not at all	1	2	1
	Very little	4	6	4
	Some	25	29	26
	Quite a bit	50	46	49
	Very much	20	18	19
Study relevant to education as a whole	Not at all	0	1	1
	Very little	3	4	4
	Some	23	24	23
	Quite a bit	46	44	45
	Very much	27	27	27

Table 46. Student Support item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Received support from university to settle into study	Not at all	2	4	3
	Very little	8	13	10
	Some	28	32	30
	Quite a bit	36	30	34
	Very much	26	20	24
Experienced efficient enrolment and admissions processes	Not at all	2	3	2
	Very little	6	7	6
	Some	19	20	19
	Quite a bit	36	33	35
	Very much	37	37	37
Induction/ orientation activities relevant and helpful	Not at all	2	5	3
	Very little	9	14	11
	Some	29	31	30
	Quite a bit	34	30	33
	Very much	26	20	24
Administrative staff or systems: available	Not at all	1	2	2
	Very little	6	9	7
	Some	27	29	28
	Quite a bit	41	38	40
	Very much	24	22	23
Administrative staff or systems: helpful	Not at all	2	4	3
	Very little	8	10	9
	Some	27	29	28
	Quite a bit	38	35	37
	Very much	25	22	23
Careers advisors: available	Not at all	4	6	5
	Very little	13	16	14
	Some	34	34	34
	Quite a bit	31	29	30
	Very much	18	16	17
Careers advisors: helpful	Not at all	5	8	6
	Very little	12	15	14
	Some	34	32	33
	Quite a bit	30	27	29
	Very much	19	17	18
Academic or learning advisors: available	Not at all	1	2	2
	Very little	6	8	7
	Some	29	30	30
	Quite a bit	41	38	40
	Very much	22	21	22

Table 46. (continued)

Academic or learning advisors: helpful	Not at all	2	3	2
	Very little	5	7	6
	Some	27	28	27
	Quite a bit	40	37	39
	Very much	26	25	25
Support services: available	Not at all	3	5	4
	Very little	11	13	11
	Some	30	29	30
	Quite a bit	33	31	32
	Very much	23	23	23
Support services: helpful	Not at all	4	6	5
	Very little	10	11	11
	Some	30	28	29
	Quite a bit	32	30	31
	Very much	25	25	25
Offered support relevant to circumstances	Not at all	15	19	17
	Very little	11	13	12
	Some	26	25	25
	Quite a bit	23	20	22
	Very much	24	22	23
Received appropriate English language skill support	Not at all	31	38	33
	Very little	9	10	10
	Some	22	21	22
	Quite a bit	20	17	19
	Very much	18	14	17

Table 47. Learning Resources item response category percentages

Item and response categories		Commencing (%)	Later year (%)	All students (%)
Quality of teaching spaces	Poor	1	3	2
	Fair	9	13	11
	Good	45	47	46
	Excellent	45	37	42
Quality of student spaces and common areas	Poor	3	7	4
	Fair	13	18	15
	Good	43	44	44
	Excellent	40	32	37
Quality of online learning materials	Poor	1	2	2
	Fair	9	12	10
	Good	42	46	44
	Excellent	47	40	45
Quality of computing/IT resources	Poor	2	3	3
	Fair	11	15	12
	Good	46	48	47
	Excellent	41	34	38
Quality of assigned books, notes and resources	Poor	2	3	2
	Fair	13	17	15
	Good	50	52	51
	Excellent	35	28	32
Quality of laboratory or studio equipment	Poor	2	4	2
	Fair	10	13	11
	Good	43	46	44
	Excellent	45	37	42
Quality of library resources and facilities	Poor	1	2	1
	Fair	8	9	8
	Good	38	42	39
	Excellent	53	47	51

Appendix I: Subject area definitions

Table 48. 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)

Table 48. (continued)

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 J. Independent review of the 2014 UES

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

Dennis Trewin AO, FASSA Statistical Consultant

1. Terms of Reference

I was asked to review the 2014 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 the 2012 and 2013 UESs.

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 ten most important statistical outputs.

I have been active in international consulting in a number of countries. Over the last 4 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. I have also undertaken an efficiency and quality review for Statistics Denmark this year.

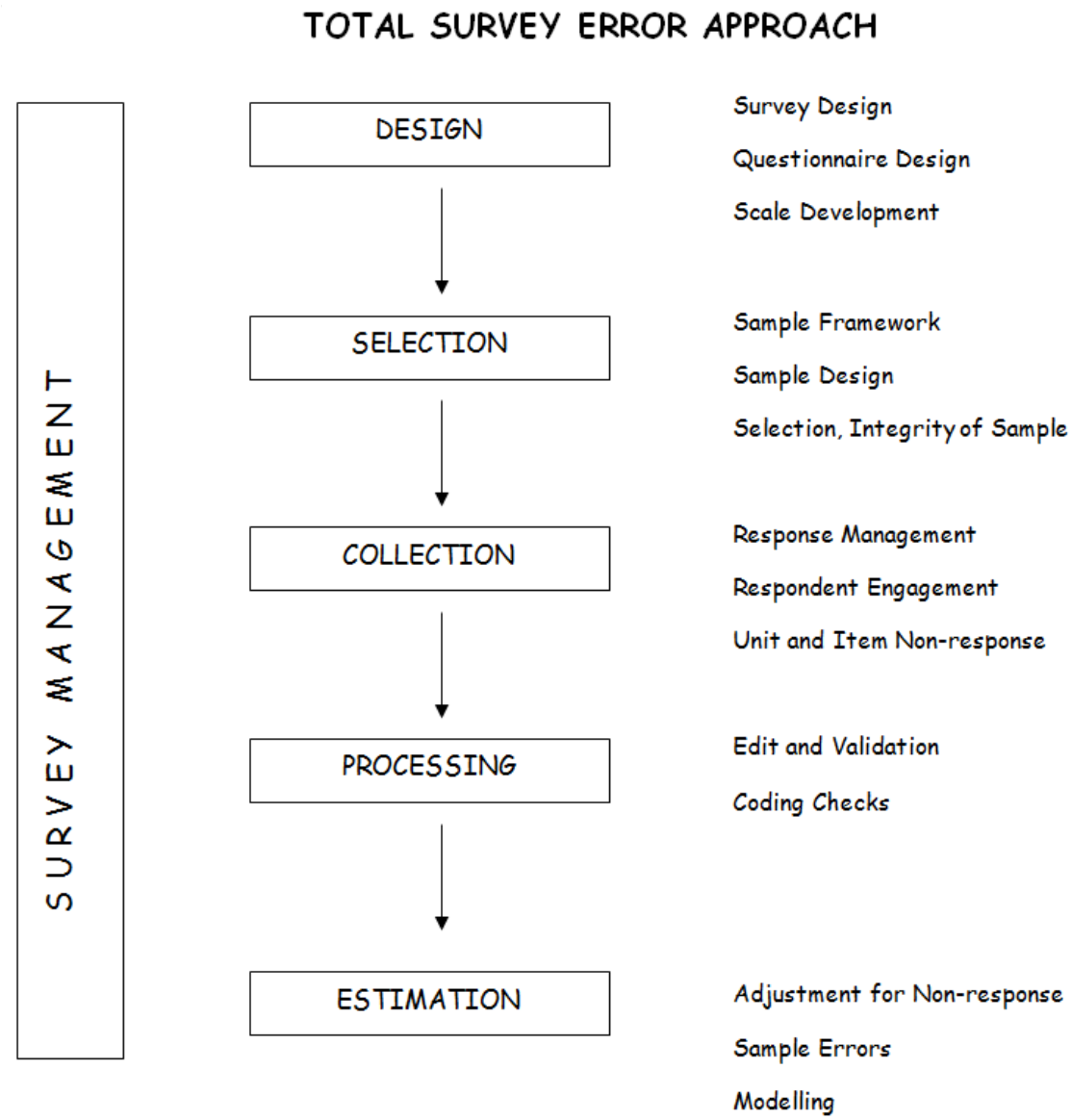
Other relevant external appointments are Past President of the International Statistical Institute, Past President of the International Association of Survey Statisticians, Chairman of the Policy and Advocacy Committee of the Academy of the Social Sciences of Australia, and Associate Commissioner of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

I have active associations with three Australian Universities. I don't believe 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 the 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 3 December 2014 where I had an opportunity to ask a range of questions about the survey processes and how effectively they worked in practice. I subsequently spoke to Sonia Whitely of SRC who was interstate on 3 December.

The following framework has been used for the preparation of previous reports and I have used it again for 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 only covers the Accuracy Dimension of the Quality assurance Framework used by the Australian Bureau of Statistics and many other national statistical agencies.



4. Summary of Findings

I have provided an overview of my assessment here. More details, using the framework above, are provided in the following parts of the Report. My assessment of whether the survey has met the requirements of the Department of Education or not are considered in the concluding paragraphs of this Section.

My assessment is that the main areas of risk to survey quality are sampling errors and non-response bias. The sample is large and significant sample errors are only really a concern at the most detailed level. Also, a well-thought through mitigation strategy was in place for other types of errors. Therefore, the focus of this Summary is on non-response bias.

A very professional approach has been taken to the design of the UES right through the various design stages. The 2014 UES has built on the previous UESs especially the experience learned during the 2013 UES. In particular, significant effort was devoted to maximising response (for a student population where response rates are traditionally low) and ensuring the achieved sample was representative. The latter was done by developing specific response rate targets for each institution by course type stratum. These are the so-called reportable strata. An overall response rate of 30.1% was achieved, whilst less than the target of 35%, this is good for a survey of this type and an improvement on previous rounds. Opportunities for further improvement should be explored in future surveys.

An interesting experiment was conducted in the pilot study of non-university higher education institutions which resulted in higher response rates for these institutions (48% compared with 30% for universities). The rolling incentives used in this pilot study to achieve this higher response rate appear to be well worth investigating for UES 2015.

Important steps taken to improve response and representation included (1) to engage as closely as possible with each University through an especially appointed survey manager, (2) a promotion and marketing campaign, (3) a planned and targeted reminder strategy, and (4) the use of incentives. These initiatives appear to be successful but no doubt could be further 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 through the use of target response rates for individual strata although some of these targets were very ambitious and probably impossible to achieve. As a consequence of the above steps, the sample was reasonably representative in respect of most of the key population characteristics. These include year of study, mode and type of attendance, Indigenous/Non-Indigenous, Disability and International/Domestic.

However, as in the previous year, the sample was not representative on gender (43% male in the population, 34% in the sample) and the response rates for Universities varied from 21% to 50% due to, in large part, the extent of student engagement at the University. However, the variability in response rates in Universities has reduced somewhat. As fourteen Universities have a response rate greater than or equal to 35%, it suggests to me that this level of response rate is achievable.

The response rates by field of study also varied somewhat but not by large amounts. For example, they were relatively high for Psychology and relatively low for Business Management and Natural & Physical sciences.

Analysis showed that gifts (shopping vouchers) did provide an incentive to respond but it seemed that it was more of an incentive for females than males. Different incentives for males should be investigated eg tickets for concerts or sports events. The other major initiative to increase response rates would be to obtain a higher level of co-operation from low performing Universities. As well as visiting the senior Executives of these Universities, it would be appropriate if the UES could be listed on the agenda of a Universities Australia meeting of the Vice-Chancellors or the Deputy Vice-Chancellors, Education.

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 universities where a sample was used have a smaller probability of selection than those universities where censuses were conducted. 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 sample. In the UES 2014, the unweighted estimates would be biased towards the responses of females, Universities with high response rates and Universities where censuses were conducted (mostly smaller universities). Based on the information I was shown:

- Females tend to have higher satisfaction levels than males,
- Not surprisingly, the Universities with higher response rates have higher 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 an upward bias in satisfaction which calculations by GCA have suggested are rather small. 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 but are likely to be larger. 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 be undertaken on the gender variable as the response rate for females is much higher than that for males and there are differences in their survey characteristics.

The estimates for fields of study for each University are also of special interest (ie the reportable strata). Not all these estimates are reliable. For some the effective sample size is low resulting in significant sampling errors. Also, response rates are very low in some cases and it is not clear that the sample is sufficiently representative to keep the risk of non-response bias to a low level. In fact the use of post-stratification weights by gender may be more important to the accuracy of the estimates at this more detailed level.

The focus on a representative sample and the use of weighting 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.

There are other potential sources of error such as those arising from poor questionnaire design or an inaccurate sample framework. I believe the risks from both these error sources are low.

There was no need for evaluation and testing of the survey instruments for UES 2014 as the UES 2013 questionnaire worked quite well and there were only minor changes. 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. If the questionnaire is too long, there will be a tendency to not answer the questions at the end. The average item non-response was 1.1% which is low compared with most surveys. It is about the same as UES 2013 (1.0%) but considerably lower compared with UES 2012 (7.7%).

The availability of the HEIMS data base as a sample framework provides greater assurance that the whole student population is being covered, simplifies the sample selection process, provides the survey administrator 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 is important to good response rates, especially the contact details. Furthermore, the use of HEIMS has enabled centralised data collection by SRC. Validity checks undertaken by SRC and the Universities suggest HEIMS is reasonably accurate in the sense that it is largely consistent with University records.

Has the survey met the client requirements?

The main survey requirements are:

(i) The general survey methodology used must be an initial approach to respondents via email providing links to an online survey instrument, followed by targeted communications to meet required response rates which may include hard copy mail outs.

(ii) The content is as for UES 2013 (with one question deleted) and supports the estimation of scales for 'Learner Engagement', 'Teaching Quality', 'Learning Resources', 'Student Support' and 'Skills Development'. The specified content also includes a number of specified analysis variables.

(iii) In addition, elements of the Course Experience Questionnaire (CEQ) must be administered to a sub-sample of students of sufficient size to achieve the accuracy required for international benchmarking.

(iv) Also, participating institutions must be provided with the opportunity to add institution specific items.

(v) The scope will include commencing and later year undergraduate degree students.

(vi) The UES must be conducted on the basis of a random stratified sample and the Department will provide HEIMS for that purpose.

(vii) A student engagement strategy which will support robust response rates should be designed and implemented.

(viii) The survey design must facilitate the publication of results based on the following stratification variables (institution, field of study, stage of studies).

(ix) At this level of stratification, estimates of percentage agreement for the 'quality of entire educational experience' item, in the teaching quality section of the questionnaire, are required at a confidence level of 90% for differences of + or – 5% (later modified to 7.5% with the approval of the Department).

(x) Statistical analysis should be used to determine whether any additional variables should be used for post-enumeration weighting to achieve representativeness.

(xi) A detailed publication of the results of UES2014 should be published.

(xii) Confidence intervals for the estimates should be estimated.

The survey design and operation met all the specified requirements for the survey with the exception of requirement (ix) as outlined above. Many of the reportable strata meet this requirement but some do not because the response rate did not reach the required level for the sampling errors to be sufficiently small. I would say that in most cases this would be because the response rate target was unreasonably high.

REPORT AGAINST ERROR FRAMEWORK

1. Survey Design

The survey design is very dependent on the survey objectives. These are specified in the contract with the Department of Education. The most relevant requirements are as follows.

- The general survey methodology used must be an initial approach to respondents via email providing links to an online survey instrument, followed by targeted communications to meet required response rates which may include hard copy mail outs.

- The content is as for UES 2013 (with one question deleted) and supports the estimation of scales for 'Learner Engagement', 'Teaching Quality', 'Learning Resources', 'Student Support' and 'Skills Development'. The specified content also includes a number of specified analysis variables.

- In addition, elements of the Course Experience Questionnaire (CEQ) must be administered to a sub-sample of students of sufficient size to achieve the accuracy required for international benchmarking.

- Also, participating institutions must be provided with the opportunity to add institution specific items.

- The scope will include commencing and later year undergraduate degree students.

- The UES must be conducted on the basis of a random stratified sample and the Department will provide HEIMS for that purpose.

- A student engagement strategy which will support robust response rates should be designed and implemented.

- The survey design must facilitate the publication of results based on the following stratification variables.

- At this level of stratification, estimates of percentage agreement for the 'quality of entire educational experience' item, in the teaching quality section of the questionnaire, are required at a confidence level of 90% for differences of + or – 5% (later modified to 7.5% with the approval of the Department).

- Statistical analysis should be used to determine whether any additional variables should be used for post-enumeration weighting to achieve representativeness.

- A detailed publication of the results of UES2014 should be published.

- Confidence intervals for the estimates should be estimated.

It is important to note 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, and for each course, if you want estimates of the same accuracy. This assumes similar response rates.

Did the design meet the above requirements? The design used for UES 2014 was consistent with the specified requirements for the survey. However, it did not meet requirement (ix) for all strata mainly because the response rates were too low for some strata. In my view, it was not possible to

completely meet this requirement because the target response rates were unrealistically high in some strata and some thought needs to be given to this prior to the design of UES 2015.

2. Questionnaire Design

There was no need for evaluation and testing of the survey instruments for UES 2014 as the UES 2013 questionnaire worked quite well and there were only minor changes. The main changes were (i) to eliminate one question from the 'student services' scale as psychometric tests showed it was not relevant to this scale and (b) the non-use of rotating panels because analysis showed the order of the panels did not matter. The questionnaire appears to have worked 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. If the questionnaire is too long, there will be a tendency to not answer the questions at the end. The average item non-response was 1.1% which is low compared with most surveys. It is about the same as UES 2013 (1.0%) but considerably lower compared with UES 2012 (7.7%).

Furthermore, item non-response did not vary much by item except for those students doing a double degree where fatigue might have been a factor. However, it was still only slightly higher for those students and still low compared with most surveys.

3. Scale Development

Prior to UES 2014 psychometric techniques were used to re-analyse the scales based on data from UES 2013. The same five scales were retained but one question was deleted from the 'student support' scale. This is consistent with a professional approach to scale development. It also supports consistency between the UES 2013 and UES 2014 surveys and would help analysis of changes between the two years. There is one exception. The 'student support' scale will vary between the two years because of the deletion of a question. This will be the reason for much of the change between 2013 and 2014 for this scale so movements cannot be reliably interpreted. I would suggest that this scale be recompiled for 2013 with the question deleted. Then a reliable movement between the two years can be estimated. Some thought needs to be given to how this data is to be presented.

4. Framework of Students

An important initiative in UES 2013 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 administrator 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, and accurate contact details, is also important for good response rates. Otherwise, non-contacts could be a major source of non-response.

One potential source of non-contact is selected students not opening the email. The University email address has been used on the expectation that students would use it during term when the survey was conducted. This seems like a reasonable assumption but it should be tested. This would be reasonably straightforward as it is possible to tell whether an email has been opened or not. If this number is reasonably large, consideration should be given to other ways of contacting these students, rather than mainly relying on university email.

5. Sample Design

The sample design was relatively 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 which is equivalent to the reportable strata for the output phase. This was appropriate given that each of these variables, and combinations of them, are disaggregations of particular interest and also because the stratification will assist the accuracy of the estimates. It also facilitates the compilation of weights that adjust for differences in the response rates and sampling fractions.

The design used for UES 2012 and UES 2013 specified that 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 the 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.

As I proposed in my last Report, the sample design should be revisited prior to the 2014 survey. The analysis of this approach suggested that this approach to determining the sample had a number of shortcomings. In general, large strata were over sampled and often achieved completed interviews in excess of 200. Consequently, students from large strata were substantially over-represented which also accentuated the gender imbalance because females were over-represented in the largest courses. Furthermore, the sample design did not take account of differential response rates across strata.

A different approach was taken to the sample design in UES 2014. It took account of the need to maximise the number of reportable strata. Target response rates were estimated for each reportable stratum although some of these were unrealistically high. These target response rates were used to tailor the response follow-up effort. Response rate targets for Universities varied considerably. They varied from 22% for Melbourne University (although it achieved a significantly higher response rate) to 68% for Bond University.

I am advised that the 35% target is actually 32.1% when calculated at the stratum level and aggregated. Rather than a broad target of 35%, the objective has for as many strata as possible to exceed their target response rate. This requires the support of the institutions. Furthermore, the focus of the survey is on results at the institution level and institution by field of study level. National estimates are a lower priority.

I agree that this sample design is more appropriate to meet the survey objectives. It has not been possible to provide estimates for all reportable strata but that is because it would have been necessary to obtain response rates that are unrealistically high for most of the strata that do not meet the confidence level objective. The sample errors for this sample design will be greater at the national level than what they would have been if the sample design had been optimised for national estimates but, as noted above, these are of lower priority.

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 and appeared to be reasonably accurate although some reformatting of street addresses was required to meet Australian Post standards. In fact, the sample was a Census for most Universities. If the sample fraction was high for a University, it was decided to conduct a Census to keep it simple. Also, two of the larger Universities decided to conduct a Census. The final outcome was that a sample was only used in eight of the largest Universities.

The Survey Managers in each University assisted with the selection and administration. The relationship with the survey managers was good in most Universities and the arrangements, involving joint responsibility, seem to have worked well.

7. Response Management

The overall response rate was 30.1% less than the target of 35% but slightly higher than 29.3% for UES 2013 and significantly higher than the response rate for UES 2012. It should also be noted that the non-response follow-up effort was targeted at strata where the response rate was significantly below the target response rate. This was the correct strategy but may not have maximised the overall response rate.

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. SRC have managed the reminder action (using a targeted approach) to maximise the number of reportable strata. This also improved the representativity of the sample. As a consequence, the achieved sample is reasonably representative of the population. It was the correct strategy but there may be some opportunities for improvement in UES 2015 as noted below.

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 Bethlehem 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).

The main means of increasing response rates were:

(i) A variety of activities at improving student awareness as well as promotion of the survey. These can also be reinforced at the University level eg lecturers mentioning the survey and its purpose in the classroom.
(ii) The explicit support of the University and the effectiveness of the appointed Survey Manager.
(iii) Reminders, both generic and targeted.
(iv) Incentives in the way of shopping vouchers.
(v) Follow-up using different communication modes such as hard copy letters to students who didn't reply to the original email invitation and sms where mobile telephone numbers were provided (17 out of 40 Universities)
(vi) Explicit Institutional efforts at strengthening response in some universities.

These seem appropriate but the effectiveness of all these strategies should be reviewed prior to the commencement of UES 2015. There is some data available to suggest reminders, incentives and the

hard copy letter are important but perhaps they could be better targeted. There is also anecdotal evidence that (vi) significantly increased response rates.

Reminders were definitely important in increasing the response rate and sample size. However, 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. It may be worthwhile to introduce further targeting of non-respondents to be followed up.

We know that males are under-represented but how to increase the male response rates? It is a global problem not just something that is peculiar to UES. Research has shown that females are more likely to engage in online activity characterised by communication and exchanging of information whereas males are more likely to engage in online activity characterised by seeking of information. I am not sure how to use this fact in encouraging increased male participation and could not find any relevant research. It is also known that incentives increase response rates. However, the shopping vouchers used for UES 2014 are more likely to increase female rather than male participation. It might be worth experimenting with different types of incentives for males eg tickets to concerts or sports events.

For the non-university, higher education pilot study, a rolling incentive program was used. A much higher response rate was achieved (47.9%). The rolling incentive program may have made a difference and this is certainly worth investigating.

There is considerable variation in response rates across universities. In general, the universities with low response rates were less engaged although in case of Swinburne there were administrative difficulties with the sample selection. Increasing engagement will require the support of the Vice-Chancellor or the Deputy Vice-Chancellor (Education) at these universities. This may require discussion at the University level and/or the regular meetings organised by Universities Australia. Fourteen universities achieved a response rate of 35% or more suggesting this is achievable for most universities.

The survey literature suggests a mixed mode of communication is the best way of increasing response rates. This is usually referring to the general population where it is more obvious that this is the case as different age cohorts will, on average, have different communication preferences. It is less obvious for student populations and the experience of UES 2014 should be analysed to see whether mail or mobile contacts have had an impact and perhaps might be used in different ways for UES 2015. It was noted that the letters seemed to have a very positive impact on response.

8. Engagement of Universities

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. The appointment of survey managers is an essential step.

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 at both the University and sector levels. If students are surveyed excessively, their co-operation is likely to diminish. Student surveys are initiated by both the Government and the University sectors. 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 30.1% was achieved, whilst less than the target of 35%, this is good for a survey of this type.

As discussed above, a number of 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 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 sample, 34% for the respondents and the response rates for Universities varied from 21% to 50% representing, in large part, the extent of student engagement at that level.

The data was not completely representative with respect to field of study. However, it was not too bad. Psychology students were over-represented. Students of management studies and physical and natural sciences were under-represented where there are more male students. For other fields of study, there were no significant differences.

What is the risk of non-response bias? Non-response bias will only exist if there are differential response rates among the population sub-groups AND the survey characteristics of those population sub-groups are different. I believe the achieved sample may be subject to some non-response bias largely because of the differential response rates for gender and field of study (to a much lesser extent) where we know the average survey characteristics are different. I don't believe the differential response rates for Universities will have much impact because the average survey characteristics are not that different.

In addition to the steps taken to ensure the representativeness of the sample, weighting procedures can mitigate against non-response bias. Although my assessment is that the non-response bias is likely to be relatively small, I would still recommend that weighting be undertaken. This would overcome any possible non-response bias problems with fields of study and universities. Furthermore, post-stratification weights based on gender would help to overcome any non-response bias problems from this source. Weighting is discussed in more detail below.

The average item non-response was 1.1% which is low compared with most surveys. This does not appear to be a significant issue for the reliability of the survey.

10. Edit, validation and imputation

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 and provided data files that were suitable for analysis. Furthermore a number of edits were built into the questionnaire.

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 mainly to check whether there were differences to those recorded on HEIMS. This impacted about 1.7% of records. From the methodological report, it appeared that this work was undertaken to a satisfactory standard.

12. Estimation, including adjustment of non-response

As mentioned above, the sample was not representative on gender (43% male in the population, 34% in the sample) and the response rates for Universities varied from 21% to 50%, due in large part to 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 by the use of weights. This is standard survey practice.

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. However, it will not adjust for any differences between respondents and non-respondents within strata. Post-stratification will 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. In fact, there are differences. Females tend to have a higher satisfaction levels so unweighted data would have an upward bias.

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 as well as some differences in their average survey characteristics. 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 is likely to be the case for UES as well. It was used as a stratification variable so there would be adjustments for any differences across fields of study if weighting was used.

Returning 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 2014, the unweighted estimates would be biased towards the responses of females, Universities with high response rates and those Universities where a Census was used. Based on the information I was shown:

- Females tend to have higher satisfaction levels than males,
- Not surprisingly, the Universities with lower response rates have lower satisfaction levels as the students are less engaged but the differences are not great 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. 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. They are also more likely to be significant at the reportable stratum level.

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 rarely used now.

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. They 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 publication – e.g. total population, field of study, University and field of study x University.

As the most important estimates are '% satisfied' it might be able to present graphically. As the underlying distribution is binomial there will be an approximate relationship between the percentage satisfied and the effective sample size. The '% satisfied' does not change that much so one approach could be to estimate sample errors (or confidence intervals) for different sample sizes for when '% satisfied' is 60%, 70%, 80% and 90%. These four alternative could be shown as four separate lines on a graph with sampling errors as the vertical axis and effective sample size as the horizontal axis. The approximate sample errors for other '% satisfied' can easily be interpolated from this graph.

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

Criteria need to be established for determining what data should be published. The key question is whether data for every reportable stratum should be published or not. Because of the small sample size, some reportable strata (ie University x Field of study x year of study) will have sampling errors higher than those specified in the Department of Education contract.

Furthermore, the non-response rate for some strata will be so high that the risk of non-response bias may be high. This may be mitigated to a large extent if separate weights are used at the post-stratification level for males and females.

I suggested that estimates for all reportable strata should 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.

A similar approach might be taken to those reportable strata subject to the risk of high non-response bias. There will be considerable overlap with those with high sampling errors mentioned in the previous paragraph. Ideally, the 'rule' should be based on a mixture of low response rates and the extent of gender representation. Some empirical work should be undertaken to determine the exact criteria. Alternatively a simple criteria based on response rates (eg less than 20%) might be used.

Of course, any cells that are confidential should be suppressed – less than 5 respondents may be a suitable guideline.

The publication should also provide readers with the 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. Non-response is clearly the non-sampling error of most significance.

Information on the design of the survey, survey variables and other meta data should be published especially to assist the more informed reader. This is the Methodological Report which was published for UES 2013. It was a good quality document that would have considerably assisted users

of the UES results. It is intended to do the same for UES 2014 and the draft Report I have seen is also a good quality and appropriate document.

Conclusions

The survey design and operation met all the specified requirements for the survey with the exception of requirement (ix) as outlined above. Many of the reportable strata meet this requirement but some do not because the response rate did not reach the required level for the sampling errors to be sufficiently small. I would say that in most cases this would be because the response rate target was unreasonably high.

The main risks to quality are sampling errors (at the reportable stratum level) and non-response bias both of which have been discussed in detail above. In particular, the relatively low response rate of 30.1% 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. However, it may be more significant in relative terms for smaller aggregates so more care should be taken in interpreting these estimates.

31 DECEMBER 2014