

Consultation on detailed changes to the Key Information Set data collection for 2017 and approaches to presenting data on the successor to the Unistats website

| | |
|--------------------------------------|--|
| To | Registrars or higher education managers at higher education providers in England, Northern Ireland and Scotland, including Alternative Providers of higher education |
| Of interest to those responsible for | Student data and information; Planning; Admissions and marketing |
| Publication date | June 2016 |
| Enquiries to | Catherine Cameron, tel 0117 931 7476, email c.cameron@hefce.ac.uk |

Executive summary

Purpose

1. The purpose of this document is to update the sector on work the UK higher education funding bodies are carrying out to implement the outcomes of the review of Unistats and the Key Information Set (KIS), and to seek feedback on aspects of this, particularly the detailed changes to the KIS data collection we are proposing for 2017.

Key points

2. The consultation is being carried out jointly by the higher education funding bodies in England, Northern Ireland, Scotland and Wales, and by the Higher Education Statistics Agency (HESA), as we work in partnership on the UK-wide KIS data collection.

3. This document consists of three sections:

Section 1 – Proposals for a revised data collection to replace the current KIS collection.

Section 2 – Proposals for aggregation of student-related data on the successor to the Unistats website.

Section 3 – An update on the investigation of possible approaches to presenting survey data on the successor to the Unistats website.

4. The ‘Consultation on changes to the National Student Survey, Unistats and information provided by institutions’ (HEFCE 2015/24), published as part of the Review of information about learning and teaching, and the student experience, included proposals to remove the following information from the KIS:

- proportion of time spent in scheduled learning and teaching and independent study
- proportion of assessment by method
- tuition fee information
- accommodation costs.

5. There was broad support for our proposals, so we are removing these elements from the collection. As we have previously consulted on this point, we are not seeking associated views from the sector in this consultation.

6. In parallel with considering the impact of these changes on the data collection, we have taken the opportunity to review the remaining data items. We are therefore also proposing to make some changes to these data items based on feedback from users (students and their advisers) and the sector. Our view is that the changes we propose will make the information more useful to prospective students, without adding burden for providers. We are consulting on these changes to ensure that they are meaningful and practical for the full range of provision across the sector.

7. We have also reviewed our approach to aggregating data. At the moment, in an attempt to meet the publication thresholds we apply for student-related data, we sometimes combine data over years and subject areas. This may be necessary where a course is new or does not have sufficiently large numbers for us to publish data for the current year at course level. We are proposing to retain this approach, and this document outlines our rationale for doing so. We are also inviting comments on this proposal.

8. Finally, we have been investigating the alternative approaches to presenting National Student Survey and Destination of Leavers from Higher Education survey data that we might adopt on the successor to the Unistats website. This document provides an opportunity for us to update the sector about this work, and the possible approaches we may take.

9. We intend to publish a specification for the data collection, which will replace the current KIS collection by the end of August 2016. We will consider the data output and content required for the new website later this year.

Consultation responses

10. The Higher Education Public Information Steering Group, the funding bodies and HESA will consider the outcomes of the consultation in late summer 2016.

11. In considering the responses, we will commit to read, record, and analyse the views of every response to this consultation in a consistent manner. For reasons of practicality, a fair and balanced summary of responses rather than the individual responses themselves will usually inform any decision made. In most cases the merit of the arguments is likely to be given more weight than the number of times the same point is made. Responses from organisations and representative bodies with a high interest in the area under consultation, or likelihood of being affected most by the proposals, are likely to carry more weight than those with little or none.

12. In our analysis, we will explain how the consultation responses were considered in our decision. Where we have not been able to respond to a significant material issue, we will usually explain the reasons for this.

13. We will publish our analysis of consultation responses alongside the new data specification. We anticipate publishing this by early September 2016.

14. This consultation is being conducted jointly by the four higher education funding bodies and HESA, but HEFCE will hold the response data. Information provided in response to a request, invitation or consultation from HEFCE may be made public, under the terms of the

Freedom of Information Act or of an appropriate licence, or through another arrangement. Such information includes text, data and datasets. The Freedom of Information Act gives a public right of access to any information held by a public authority defined within the Act, in this case HEFCE. It applies to information provided by individuals and organisations, for example universities and colleges. HEFCE can refuse to make such information available only in exceptional circumstances. This means that data and information are unlikely to be treated as confidential except in very particular circumstances. Further information about the Act is available at www.ico.org.uk.

Action required

15. Responses to this consultation should be made online by **noon on Tuesday 26 July 2016**, using the online response form which can be accessed alongside this document on the HEFCE website at www.hefce.ac.uk/lt/unikis/consultations.

16. To ensure that the changes we make to the data collection have taken account of the full range of higher education provision, we would welcome responses from across the sector, including those of alternative providers that have yet to submit KIS data.

17. We recognise that these proposals may be of interest or relevance to different areas in institutions, but would encourage the submission of a single institutional response. We will therefore be grateful to institutions for collating responses internally where necessary.

Background

18. In 2015, the four UK higher education funding bodies carried out a review of Unistats and the Key Information Set (KIS)¹. The findings of this review were reflected in late 2015 in the proposals in 'Consultation on changes to the National Student Survey, Unistats and information provided by institutions' (HEFCE 2015/24), a consultation with the sector as part of the wider Review of information about learning and teaching, and the student experience². The key proposals relating to Unistats and the KIS were:

- a. To retain a website which would act as an authoritative source of national data for students and their advisers.
- b. In designing this website, to ensure that it would reflect diverse student information needs and help students to identify the information they might want to consider in making their decision, and assist them with finding this information.
- c. To transfer publication of learning and teaching information to institutions, to reflect findings of student demand for detailed, nuanced information and the limitations of our current approach of presenting summary metrics.
- d. To ask institutions to publish detailed information about their provision, in line with Competition and Markets Authority guidance on consumer law³. The funding bodies would support them in this by providing good practice guidance on how this might look for different types of provision, to facilitate consistency and therefore ease of comparability for students.

19. Consultation responses broadly supported the proposals. Seventy-four per cent of respondents agreed with the respective roles we identified for funding bodies and institutions, and 69 per cent agreed that the proposed changes would improve the accessibility of information and ensure the data we provide is meaningful to students. The main reservations were around removing central, comparable information, and a significant number of respondents suggested that we enforce the guidance we proposed providing to institutions on information to be provided on their own websites. An analysis will shortly be published as 'Summary of responses to consultation on changes to the National Student Survey, Unistats and information provided by institutions'⁴.

20. The funding bodies' respective decision-making bodies approved our recommendation to proceed with the changes proposed as a result of the outcomes of the consultation, but

¹ The four UK higher education funding bodies are the Department for the Economy in Northern Ireland (DfE), which has taken on the functions of the former Department for Employment and Learning, the Higher Education Funding Council for England (HEFCE), the Higher Education Funding Council for Wales (HEFCW) and the Scottish Funding Council (SFC). 'Report on the review of the Key Information Set and Unistats' (HEFCE 2015/27) is available on the HEFCE website at www.hefce.ac.uk/pubs/year/2015/201527/.

² Consultation on changes to the National Student Survey, Unistats and information provided by institutions' (HEFCE 2015/24) is available at www.hefce.ac.uk/pubs/year/2015/201527/. Further information is available at <http://www.hefce.ac.uk/lt/roiconsult/>.

³ Available at www.gov.uk/government/publications/higher-education-consumer-law-advice-for-providers.

⁴ We anticipate the analysis being available by the beginning of July 2016 at www.hefce.ac.uk/pubs/.

noted the need to consider how we might ensure students can easily access information relating to accommodation costs.

21. We are currently taking forward three main streams of work to implement the outcomes of the review:

- development of a revised data specification and approaches to presenting data on a new website
- development of a new website
- development of guidance for institutions on providing information in areas relevant to decision-making about undergraduate study.

22. In developing the guidance for institutions under the third of these work streams, we will seek to respond to concerns about the loss of comparable information by promoting consistency of presentation. We will also work closely with the National Union of Students and the Competition and Markets Authority on its development.

23. We are also looking at a branding and communications strategy for the new website.

24. This document focuses on the first of these work streams. Before issuing it, we invited individuals from a range of institutions to attend a focus group meeting to advise us on our proposals. This was held on Wednesday 25 May 2016, and a note of the meeting, including those who attended, is attached at Annex A.

Section 1: Proposals for a revised data collection

25. In addition to making changes to take account of the outcomes of the Review of Unistats and the KIS, we have considered whether further changes would be desirable to make the information more useful to students. This section explains the detailed changes we are proposing. Those who are not familiar with the current data specification may find it helpful to refer to the guidance on the Higher Education Statistics Agency (HESA) website⁵.

26. A summary of current KIS input fields, and or whether we intend to remove or retain them, is attached at Annex B. In paragraphs 27 to 59 we have outlined our proposed approach to various aspects of the data collection, and would welcome feedback on each of these areas.

27. As a result of the review there will no longer be the concept of a Key Information Set. This reflects findings about individual preferences for information, and is consistent with some key information being provided on institutional websites in a different format from the summary metrics which we currently publish. We intend to give the data collection the same name as the website on which we publish the output; we are considering whether that will continue to be 'Unistats'.

Updating the dataset

28. The Unistats dataset is updated weekly, and we propose that this continues so that any changes can be reflected in a timely manner.

29. We have received positive feedback on the introduction of delegated sign-off, so we propose to continue with the current approach to signing off the data: the head of institution

⁵ See www.hesa.ac.uk/C16061.

must sign off the initial dataset for publication, but may then delegate authority for the sign-off of subsequent amendments.

Coverage of the record

30. We propose to retain the same criteria for courses to be included in the record: a record should be returned for each undergraduate course to which students can apply for entry in the following academic year, where these courses last more than one year when studied full-time. The only types of course resulting in a masters level qualification which should be included are integrated masters, such as MEng programmes.

31. The funding bodies intend to explore extending the coverage of the National Student Survey (NSS) to incorporate shorter courses and, if this happens, we will extend coverage of the collection to include these. We also plan to consider further publishing information about higher and degree apprenticeships.

Return of data for sub-contracted provision

32. The current approach to presenting data on Unistats where a provider is in a sub-contractual relationship with another (otherwise known as ‘franchised’ provision) is for the courses to be listed under the teaching institution. Responsibility for returning data for sub-contracted provision is as follows:

- a. Where all of a teaching institution’s higher education provision is franchised from one other institution, that other institution should return the data.
- b. Where a teaching institution has provision franchised from multiple institutions, the teaching institution needs to return its own data. (Otherwise it would be difficult for us to identify when there was a complete dataset for that institution, and therefore when to aggregate the data to allow it to be published.
- c. Where students study at two or more institutions as part of the same course, the registering institution completes the return.

33. We believe that continuing to present course information by teaching institution is the most helpful approach for prospective students, as it makes it clear to them where they will study. We also intend to continue to aggregate data to attempt to reach publication thresholds. To enable us to do this, we propose to retain the current approach to the return of data for collaborative provision.

KISTYPE

34. The concept of a KISTYPE was introduced to enable institutions whose provision allows subjects to be combined flexibly, to return their data, particularly data on learning and teaching methods and assessment, in a less burdensome way⁶. It allows for the return of subject-level information which can be grouped together for advertised combinations of subjects where it is not possible or desirable to return a full KIS record (a KISTYPE of 1) for an advertised programme.

35. We are proposing discontinuing the KISTYPE and associated fields. Given that we are no longer gathering learning and teaching and assessment information, our view is that it

⁶ Further explanation of types of KIS is available on the HESA website at www.hesa.ac.uk/index.php?option=com_studrec&task=show_file&mdl=16061&href=KISTYPES.html.

should be possible to return all provision under what is currently KISTYPE 1, so KISTYPE is no longer required.

Sandwich year, year abroad and foundation year

36. It is possible to indicate in the KIS whether a course has an optional or compulsory sandwich year, year abroad or foundation year. The current guidance is that this should only be returned where the placement or year abroad equates to a full single course stage.

37. We propose retaining these fields, as they are used widely by students when considering which courses they may be interested in. We have also considered whether it would be helpful to make prospective students aware of opportunities for shorter placements or periods of study overseas, and would welcome feedback on this.

38. We note the advice from the focus group we convened earlier this year that clear guidance would be needed to ensure consistent use of these fields by institutions.

Course title

39. When the KIS was first introduced, course titles were captured in free text without any validation. Institutions took various approaches to how they constructed these, sometimes making it challenging for users of the data to identify comparable courses. In response to this, we introduced separate fields to capture qualification aim, honours and characteristics such as sandwich year in a structured way. We also introduced validation to prevent the repetition of this information in the title field in inconsistent ways.

40. User testing and feedback from institutions indicate that, while the introduction of the additional fields has been beneficial, there are concerns about inconsistencies between the titles used for courses on institutional websites and how they appear on Unistats, particularly where there are multiple courses with the same title on Unistats. We therefore propose retaining the structured fields (qualification aim, honours and characteristics), as this will allow users to filter course data and ensure some consistency of presentation. We also propose, however, to amend the guidance to allow institutions to return a title which reflects how the course is named on their website, and to remove the data validation which prevents the inclusion of text such as 'year in industry' or 'year abroad' to enable this.

41. We will ensure that the search functionality of the new website is developed to ensure courses with complex titles are not less easily found than others. We will also monitor course titles and adjust validation where we find the changes we make have had unintentional consequences, in terms of allowing institutions to include content we do not consider appropriate.

Distance learning

42. The DISTANCE data item currently records whether a course is only available by distance learning. Our observation, however, is that users of are also interested in knowing when a course can optionally be pursued by distance learning. We propose to introduce an additional value for this item to indicate that it is possible but not mandatory to study a course fully through distance learning.

Accreditation information

43. The data collection currently allows institutions to indicate accreditations awarded to the course and recognition or endorsement by public, statutory and regulatory bodies. They do this by selecting codes from the accreditation table, which results in statements appearing on the Unistats website describing the accreditation.
44. We have identified widespread issues with the accuracy of accreditation information on Unistats, but accreditation is an important factor for a significant number of students who use it as an initial filter for the courses they consider, so we intend to retain this information. We have sought feedback from institutions about the reasons for these data quality issues, and are investigating ways in which we could improve quality assurance of this information and assist institutions in selecting the correct accreditation statements for their courses.
45. We have considered the possibility of collecting this information from accrediting bodies, but do not believe this will be feasible because of the need to link to KISCOURSE. We also believe the responsibility for providing this information should remain with institutions, as it is a legal requirement and therefore it is most appropriate that they retain control of its provision and responsibility for its accuracy.
46. We do not propose any changes to the way accreditation information is captured in the data collection: we plan to continue to record a code for accreditation type, which will link to a statement to be published on Unistats and a URL linking to information on the accrediting body's website. We have considered a number of areas where it has been suggested that it might be desirable to make changes, notably:
- a. Whether it is appropriate to represent different types of accreditation, endorsement and recognition in different ways (for example, different indicators for accreditations which are essential for professional practice and for endorsements from bodies that provide validation that content meets employer expectations).
 - b. Whether it is desirable to change the guidance on accreditations which are due for renewal. We currently advise that accreditations due for renewal in the intervening period between data collection and the start of the course should be included.
47. In terms of the first point, we believe the additional complexity this would introduce would not be helpful to students, and that it would be preferable to ensure that the wording of the statement is very clear about the nature of the accreditation. We will also be able to explain the different types of accreditation displayed more fully on the site, as it will provide guidance on such areas.
48. In terms of the second point, our view is that it would be disadvantageous to institutions not to permit them to display an accreditation which can be reasonably expected to be renewed. Similarly, user experience testing to date suggests that qualifying information in any way reduces the level of trust placed in it, so we do not propose to introduce an additional data item to indicate when an accreditation is due for renewal.
49. The list of accrediting bodies used for the KIS was initially drawn from a list put together by the Higher Education Better Regulation Group, but there has been an annual process for consideration of bodies which wish to be added to the list. This has resulted in some bodies appearing on the list which do not meet the current eligibility criteria, as these

have evolved over time. Over the coming months, we will review the eligibility criteria and the process for assessing whether bodies meet these. We will then carry out an exercise to establish whether the bodies on the list meet the new criteria and remove those which do not. This will provide us with a robust basis for the future management of this information.

Course length

50. As we are no longer collecting information on scheduled learning and teaching, there is no data item we can use to convey the length of the course to students. We are proposing retaining the NUMSTAGE field to capture the length of the course in years so that this can be displayed. We recognise, however, that this may be more complex for part-time provision, which can be studied at variable levels of intensity. We therefore propose recording the length of the course in years for full-time courses in the NUMSTAGE field, but displaying a message on the website for part-time courses referring students to the institutional website for details of the length of the course.

Accommodation information

51. The proposals consulted on last year included the removal of accommodation information from the data collection. This was considered relatively onerous for institutions to produce, and not applicable to all. Our conclusion was that, given that this data is available from other sources, continuing to collect and publish it centrally had questionable value.

52. There was strong feeling among some respondents to the consultation, however, that this should be retained.

53. The evidence we gathered from students during the review indicated that the information provided on Unistats is not as helpful or as easily understandable as it might be. Having given this issue further consideration, and discussed it with the Higher Education Public Information Steering Group and the focus group we convened, we have concluded that the best approach to presenting this information is to provide guidance to institutions on supplying information about institutional accommodation on their own websites, which students can easily access through a URL provided in the ACCOMURL field. Where appropriate, they can also provide links to where students can find information about the costs of private accommodation.

Employability

54. The record currently contains the EMPLOYURL field, which captures a link to employability information on the institution's website. We are reviewing our approach in this area, and will be revising the guidance to reflect our expectations of what information should be provided by institutions, but we propose retaining these fields.

New courses

55. Currently on Unistats, we display messages where we have no data for a particular item or where data has been aggregated. The message when we cannot publish any data is:

'There are not enough data available for this item to give specific information for the course. It may be because the course size is too small, or because it is a new course.
This does not reflect on the quality of the course.'

56. These messages are the same whether the reason we cannot publish data is because relevant publication thresholds have not been met or because the course is new. We currently use 'new' to describe a course that has either not yet recruited or not been running long enough to have generated all the data for Unistats; for example, there may be no Destination of Leavers from Higher Education (DLHE) survey data if a course has not yet had a graduating cohort.

57. Feedback received from users indicates that these messages are insufficiently clear, and it would be helpful if we could differentiate between instances where sample sizes are small and where there is no data (or aggregated data is displayed) for other reasons.

58. We are proposing to ensure that the reason aggregated data or no data appears is clear to users by:

- a. Displaying a message to indicate where publication thresholds have not been met or, for aggregated data, have only been met at that level of aggregation.
- b. Displaying a message to indicate where the course is new (or has not been run in that format previously). We intend to do this where there are no HESA or Individualised Learner Record course records linked to the KIS course as part of the return.
- c. Displaying a message to indicate where data is not available because the course has not been running sufficiently long to generate it; for example, when a course has not yet had a cohort participate in the NSS or DLHE survey. We will determine this from the NSS, DLHE and other data sources, as applicable.

Maintaining links to institutional websites

59. The record currently contains a number of URL fields which link to particular areas of information on institutional websites. We propose retaining these and – given the change in emphasis to presenting information on institutions websites and the role of Unistats in helping prospective students find this easily – it is particularly important that these links are accurate. We often find issues with broken or incorrect links during data audits, and have been considering possible ways of improving data quality in this area.

60. We propose adding a facility to the new website so that users can report incorrect links, and will ask institutions to nominate a contact to whom this information can be passed. We will monitor reports of incorrect links and, if a significant number are reported at a single institution, will follow this up with a nominated contact. We would welcome suggestions for other steps we could take to facilitate maintenance of the website links included in the return. We have considered automated solutions, but links are often to incorrect pages rather than producing errors, so we do not think this type of solution will resolve the issue.

Consultation question 1a

Do you agree with our proposed approach to each of the areas outlined in the consultation document?

Consultation question 1b

Do you have comments on any other areas of the data collection?

Section 2: Proposals for aggregating student-related data on the successor to Unistats

61. The research carried out for the review of Unistats and the KIS indicates that prospective students are most interested in information about courses, and we therefore intend to continue to present course-level information on the successor to Unistats.

Current approach

62. To ensure the data we publish is statistically robust and meets data protection standards, we apply publication thresholds for student-related data, such as NSS and DLHE survey responses. User testing has found that an absence of data can be viewed negatively by users, and prospective students would rather see some data than none, so where these publication thresholds are not met by the current year's data for a course, we attempt to meet them by aggregating data across years and subject area. We aggregate data in the following order until data that meets the thresholds is achieved:

- course level, most recent two years
- Unistats subject level 3, most recent year (108 subjects)
- Unistats subject level 3, most recent two years
- Unistats subject level 2, most recent year (42 subjects)
- Unistats subject level 2, most recent two years
- Unistats subject level 1, most recent year (21 subjects)
- Unistats subject level 1, most recent two years.

If thresholds are still not met, we cannot publish any data.

63. In 2015, we consulted on and implemented a reduction in the headcount or full person equivalent threshold from 23 to 10, and we intend to retain this threshold⁷. As part of the implementation of the outcomes of the review of Unistats and the KIS, however, we have been considering whether our current approach to aggregation is the most appropriate.

Consideration of alternative aggregation approaches

64. We have considered whether years and subjects are the most appropriate measures to use in our aggregation methodology, or whether alternative approaches could be taken. We have evaluated various other ways in which we could aggregate, such as department or faculty, campus, location, and modules. Our guiding principle has been that the way we combine data should seek to reflect as closely as possible the experience of students on the course for which the aggregated data is being published. For this reason, we concluded that aggregation by campus or location would be too broad. Using module would be more likely to group students in a representative way, but this data is not available for all courses at all institutions, so we could not adopt this approach. Similarly, while using department or faculty could group students in a representative way, this is not available to us in the student data.

⁷ The consultation document and analysis of responses are available on the HEFCE website at www.hefce.ac.uk/lt/unikis/aggregation/.

We therefore came to the conclusion that the most reasonable approach is to retain an approach based on subject groupings, as these act as a proxy for department or faculty.

65. We have therefore considered whether our current subject groupings, which use the Joint Academic Coding System (JACS) subject codes as their basis, is still appropriate⁸. The Higher Education Data and Information Improvement Programme is currently developing a new subject coding system, the Higher Education Classification of Subjects (HECOS), which will replace JACS. Many of the responses to our 2015 consultation on publication thresholds and subject grouping⁹ proposed that HECOS should be considered when it was developed. A default hierarchy for grouping subjects within HECOS to allow for aggregation of data has recently been proposed¹⁰. Our next step will be to consider the appropriateness of this for use in our aggregation methodology, or whether an alternative hierarchy of HECOS would be better. We hope to confirm the subject hierarchy we will use for aggregation when the revised data specification is published.

66. We have also considered whether we should combine data over more than two years, particularly as the Teaching Excellence Framework technical consultation proposes using three years to calculate metrics¹¹. The purpose of publishing data on Unistats (and its replacement) is to aid prospective students in making decisions about courses they wish to study, and our view is that adding a third year reduces the relevance of the data to an unacceptable level. The likelihood increases that provision could have changed in the intervening period, or that the aggregated data fails to reflect changes that institutions make in response to NSS results. We therefore propose continuing to aggregate data over two years only.

Consultation question 2a

Do you agree that continuing to aggregate data by subject and across two years is the most appropriate approach given the purpose of publishing it?

Aggregation of 2017 NSS data

67. Alongside the Unistats review, we have also been reviewing the NSS, and proposals for changes for 2017 and beyond were included in HEFCE 2015/27. There was broad support for the proposals and it is expected that, while the new survey will retain many of the existing questions in their current format, other questions will be reworded and reordered and new questions introduced. We provide data for all NSS questions on the Unistats website, and need to consider how this data may be aggregated within the new website.

68. The 2017 questionnaire is in the final stages of development and it is likely that the first few questions of the NSS will retain their order. The rest of the survey, however, will differ from 2016 in that it will be a mixture of new, revised or reordered questions. It is therefore possible that wording for a particular question could remain the same, but its place in the

⁸ See <https://www.hesa.ac.uk/jacs3>.

⁹ See <https://www.hefce.ac.uk/lt/unikis/consultations>.

¹⁰ See https://hediip.ac.uk/subject_coding/.

¹¹ See <https://www.gov.uk/government/consultations/teaching-excellence-framework-year-2-technical-consultation>. This proposes to use metrics averaged over three years in addition to metrics for individual years.

survey changed, in which case the response might differ as the student could have been influenced by the content of previous questions.

69. To inform the development of the 2017 survey, we have carried out a 2016 pilot and are proposing to analyse the pilot data to establish which of the existing questions (including reworded questions) could legitimately be aggregated. This can be achieved by using the first group of consistent questions as a control group and then assessing either the consistency per question between the pilot and main survey, or whether the relationships between questions are different between the two surveys.

70. We need to decide an approach to publishing NSS data for the first year of the new survey, given that we will have a mixture of questions, some suitable for aggregation over two years and some either not amenable to aggregation or with only one year of response data. Possible approaches could be:

a. Not to aggregate any NSS response data over two years where aggregation is necessary to meet thresholds, and only aggregate across subjects. This would mean we would publish less course-level data, but responses would be aggregated in the same way for all questions for each course. This would be easier to explain to users than a mixture of aggregation approaches. We would also publish less data overall, as responses that could be publishable if aggregated at the highest level of the subject hierarchy and also over two years would not be published.

b. To aggregate questions over two years, where questions are suitable for aggregation and we have last year's data, but not to attempt to aggregate responses to other questions across subjects. This would mean that for some courses we were only able to publish responses to a subset of the NSS questions (those which could be aggregated over two years).

c. To aggregate questions over two years where questions are suitable for aggregation and we have last year's data, and to attempt to aggregate responses to other questions across subjects. This would mean that for some courses the data displayed would result from aggregating responses over two years for some questions, and from aggregating across subject areas for other questions. We would publish more data than with other options, but it would be challenging to ensure that users understood they were being presented with a mixture of course-level and subject-level data.

Consultation question 2b

Which approach to aggregation of NSS data do you favour us adopting in 2017? Provide a rationale for your response.

Section 3: Data Presentation

71. When consulted on a reduction in publication thresholds for student-related data, we sought expert statistical advice on whether, and to what extent, this would increase data volatility. We also asked about any additional considerations we should be aware of. In addition to providing specific advice on the effect of a reduction, the statistician we consulted advised that, given the extent of the uncertainty in the data values and the high incidence of overlapping confidence intervals between courses, we should make the level of uncertainty in the data

clearer to students (regardless of whether we reduced publication thresholds)¹². We had also observed in user testing that students sometimes concluded that NSS results, for example, were better for one course than another when the values were not statistically different. As an interim measure, we developed a video which we embedded on Unistats to explain how sample size and response rates could affect confidence in the data, but we committed to investigating alternative data presentations as part of the work to implement the outcomes of the review.

72. We have been working with the user research consultants for Unistats, Fluent Interaction, to test various data representations of NSS data with prospective students and parents of applicants. Paragraphs 72 to 75 present a summary of what we have found.

Confidence intervals

73. We tested various visual representations of confidence intervals around percentage results. None of these tested well, and even when the concept was explained to participants, they either struggled to understand it or did not think it helpful to have it displayed.

Benchmarking

74. We tested a number of representations of relative position against a benchmark for the subject concerned. Participants generally found this easy to understand and it appeared to reduce the likelihood of them making invalid comparisons: where values were different, but the same in terms of position in relation to the benchmark, they were less quick to think one better than the other than we observed with strictly numeric representations. However, participants did not respond well to a presentation that was purely a statement of the position relative to the benchmark (for example, 'Above', 'Below' or 'Typical'); and wanted to see the percentage value in addition to this.

Sample sizes

75. In the various presentations we tested, we included sample size and the number of respondents. Participants understood that the larger the sample size and the higher the response rate, the more confidence they could have in the data and the more likely it was to be representative.

Distribution of responses

76. We included a breakdown of responses in some of the visualisations we tested (that is, the number of respondents who provided each of the five answers on the NSS response scale). These were received positively and appeared to add value by providing an additional way for students to differentiate between similar survey results.

Conclusions

77. In the past, confidence intervals have been displayed for NSS results. This approach was changed as they were not well understood. It was clear from the testing that, displaying the

¹² The confidence interval is the statistical tool that we used to measure the level of uncertainty in data. This is a numeric range with a minimum and maximum bound, within which the true value of an unknown value is likely to fall. This is useful where a hypothetical value has been generated using a sample, providing an estimate of the value for the population. Generally, the smaller the difference between the minimum and the maximum of the confidence interval, the greater the certainty that the generated value represents the true value (or 'population value'). Overlapping confidence intervals indicate that values are not significantly different (meaning that we cannot be sure that they are different).

confidence interval around the value for survey data, while it might meet the need to be more transparent about the level of uncertainty, would not be helpful to prospective students, as it is likely to either confuse them or cause them to interpret the confidence interval as something else, for example a range of responses or grades.

78. We therefore intend to carry out further testing of visual representations which aim to ensure that students do not make invalid comparisons based on small differences in data values, and thereby to improve understanding of the uncertainty in survey results.

79. If we were to decide that indicating position relative to the sector benchmark was desirable, we would carry out benchmarking for all questions in a comparable way to how we currently benchmark the overall satisfaction question in the NSS (Question 22) in the public dataset¹³.

Consultation question 3

Do you have any comments on the approaches we are exploring for presenting survey data, for example the use of benchmarking?

¹³ See www.hefce.ac.uk/lt/nss/faq/#Q2.

Abbreviations

| | |
|--------------|---|
| DfE | Department for the Economy in Northern Ireland |
| DLHE | Destination of Leavers from Higher Education survey |
| HECOS | Higher Education Classification of Subjects |
| HEFCE | Higher Education Funding Council for England |
| HEFCW | Higher Education Funding Council for Wales |
| HESA | Higher Education Statistics Agency |
| JACS | Joint Academic Coding System |
| KIS | Key Information Set |
| NSS | National Student Survey |
| SFC | Scottish Funding Council |