

Analysis of responses in relation to regulating student outcomes and setting numerical baselines

Purpose of this document

1. In November 2020, the Office for Students (OfS) published a consultation which set out proposals for a new approach to regulating quality and standards (we call this our 'phase one consultation').¹
2. In July 2021, we set out our more detailed proposals for regulating qualitative elements of quality and standards in a further consultation (we call this our 'phase two consultation').² Alongside this consultation we published our analysis of responses we had received to our November 2020 consultation.
3. In January 2022, we published more detailed proposals for our approach to regulating student outcomes (we call this our 'phase three consultation').³ This document provides our analysis of the responses received to the November 2020 consultation that relate to the regulation of student outcomes. It should be read alongside our main proposals.

¹ See www.officeforstudents.org.uk/advice-and-guidance/quality-and-standards/changes-to-our-approach/phase-one-strengthening-our-approach/.

² See www.officeforstudents.org.uk/advice-and-guidance/quality-and-standards/changes-to-our-approach/phase-two-high-quality-courses-and-reliable-standards/.

³ See www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/student-outcomes/.

Contents

Analysis of responses	3
Setting numerical baselines	3
Our response	6
Setting higher and more challenging baselines	9
Our response	10
Proposed indicators – general points made about all indicators	11
Our response	12
Proposed indicators - continuation	13
Our response	14
Proposed indicators - completion	15
Our response	16
Proposed indicators - progression.....	17
Our response	20
Proposed indicator - degree outcomes.....	22
Our response	23
Mode and level of study	23
Our response	26
Student characteristics.....	28
Our response	31
Small cohort sizes.....	31
Our response	32
Indicators for subjects and identifying pockets of performance.....	33
Our response	34
Partnerships.....	36
Our response	37
Indicators for transnational higher education courses	38
Our response	39
Considering data about higher education courses that are not eligible for funding by the OfS as context.....	39
Our response	40
Additional indicators.....	40
Our response	41
How baselines would be set for each indicator.....	41
Our response	43
Baselines for different student characteristics	44
Our response	45
Publication of data about providers' performance.....	45
Our response	46
Assessment of condition B3 and consideration of context.....	46
Our response	50
Assessment of initial condition B3 for providers seeking registration with the OfS.....	51
Our response	51
Definition of successful student outcomes	52
Our response	52
Coherence with our approach to access and participation	52
Our response	53
Coherence with TEF	53
Our response	54
Regulatory burden	54
Our response	55

Analysis of responses

Setting numerical baselines

4. The phase one consultation set out that, as described in the regulatory framework, assessment of condition B3 would be made in relation to a minimum level of absolute performance in order to protect students.⁴ It proposed that we would set numerical baselines for student outcomes and assess a higher education provider's absolute performance in relation to these.
5. A minority of respondents agreed with the proposed approach to set absolute numerical baselines. These respondents agreed that all students should expect the same minimum quality of education, including the outcomes delivered, and recognised that an assessment of absolute performance, rather than against benchmark values, was necessary if the Office for Students (OfS) was to avoid integrating disadvantage into the regulatory requirements by setting lower requirements for providers that recruit students from underrepresented groups.
6. There was, however, widespread support for the principles underpinning the proposals, to ensure that all students receive higher education that meets minimum quality and that gaps in outcomes between student groups reduce; and there was also considerable support for a more risk-based approach. However, the vast majority of respondents were not in favour of an approach that sets numerical baselines and assesses absolute performance as a way of achieving the OfS's policy aims.
7. The most common argument made against assessing providers based on absolute performance was that it would ignore external factors, which respondents argued were outside providers' control and could negatively affect outcomes. Student characteristics were the most mentioned factor, with respondents pointing out that there are currently differences in outcomes between demographic groups, and in relation to prior attainment. Respondents also commented on changes in students' personal circumstances, their choices about their education or career paths and their responsibility for achieving their own successful outcomes.
8. Respondents also suggested that an approach that assessed performance in absolute terms would not take account of known regional differences in outcomes and the local context in which providers operate.
9. Some respondents suggested that the established differences in outcomes between student groups are already being addressed through access and participation plans (APPs), and therefore to consider student outcomes through condition B3 represented a duplication.
10. Other respondents suggested that the process for setting minimum baselines and assessing providers' absolute performance is very challenging, given the variation in student outcomes between student groups, courses, and employment sectors and the considerable variation

⁴ See www.officeforstudents.org.uk/publications/consultation-on-regulating-quality-and-standards-in-higher-education/ and www.officeforstudents.org.uk/publications/securing-student-success-regulatory-framework-for-higher-education-in-england/.

between providers in scale of operation, courses, and UK-based and non-UK based education.

11. It was suggested by some respondents that the high continuation rates observed in UK higher education, compared with those seen internationally, demonstrates that the sector is too risk-averse in recruiting students from underrepresented groups or innovating provision, and that the focus on continuation may encourage this and reduce flexibility in higher education courses.
12. A large number of respondents suggested that assessing absolute performance would generate the unintended consequence of reduced access and participation for students from underrepresented groups because it would create disincentives for providers to recruit these students. This is because these students have worse outcomes, require more support to succeed and pose greater risks for provider performance than others. Respondents suggested that higher entry requirements and less flexible delivery models could become normalised.
13. Respondents suggested that providers may be less willing to admit students with disabilities because they are considerably less likely to progress to graduate-level employment or further higher study than students without a declared disability, or they may need more time to complete or to repeat as part of reasonable adjustments. One respondent noted that students with disabilities make up a higher proportion of those registered on courses in creative subjects, and the proposals may incentivise closures due to lower graduate progression rates or lower graduate salaries in these types of courses. It was also noted that students with physical disabilities may be less geographically mobile, leading to less choice of provider at the point of application.
14. Some respondents suggested that providers might be less willing to recruit mature students because their continuation and completion tend to be worse and mature students whose objective is study for pleasure, rather than to progress to graduate employment, would not contribute positively to a progression indicator.
15. Other respondents suggested that if providers became more risk-averse in admissions decisions and course provision there could be a particular impact on minority ethnic students, due to persistent attainment gaps at school level leading to less choice of provider at the point of application.
16. Respondents also suggested that an assessment of absolute performance would be detrimental to the diversity of providers, courses and geographical access to higher education. It was suggested that the proposed approach could discourage providers from innovative admissions practices or course offers, including flexible and modular study options aimed at widening participation and lifelong learning models that accommodate students' need to pause and return to higher study based on personal circumstances.
17. Respondents also commented that the proposals could lead providers to discontinue courses that would not perform well against the baselines and to adopt a risk-averse approach to developing course portfolios. For example, courses whose graduates do not typically achieve graduate-level employment might be discontinued. This could generate geographical gaps in

course availability, including through partnership provision, and reduced access for students who need to study in their home location or mature students who return to education.

18. Another common point made by respondents was a view that outcomes indicators will obscure or ignore other relevant aspects of 'quality'. It was suggested that a focus on performance against a numerical baseline would lead to providers focusing on achieving a narrow set of outcomes measures at the expense of a broader focus on quality.
19. Some respondents pointed out that successful outcomes against baselines may not in fact mean that a provider is providing a high quality student experience.
20. One respondent considered that the proposals suggest that the OfS has already reached a firm view on the application of numerical baselines for continuation, completion and progression, as well as on the approach that it will take to considering the drivers of 'outcomes' unrelated to the quality of the educational provision.

Impact on specific types of providers

21. Respondents suggested that the proposals would unfairly disadvantage providers with high proportions of students from disadvantaged backgrounds, or studying as part of transnational educational partnerships, because assessment would not consider 'value added' or students' prior levels of attainment and socioeconomic circumstances, which affect outcomes and which they argued are outside providers' control.
22. It was suggested that further education colleges that deliver higher education could opt to withdraw from higher education because of the proposals, expanding geographical gaps in delivery and disadvantaging students, particularly those from disadvantaged backgrounds, and stakeholders. We have taken this to mean that because some of the providers in the English higher education sector with the lowest rates of continuation, completion and progression are further education colleges, this might result in those providers finding it challenging to meet baselines and therefore facing regulatory intervention, and they might withdraw from higher education to avoid this.
23. Respondents also suggested that small and specialist providers that focus on widening participation for students from underrepresented groups could be discouraged by the reliance on absolute performance, because their outcomes cannot be compared to larger providers where underrepresented students will be part of a larger student body.

Alternatives suggested to absolute performance measures

24. A considerable number of respondents suggested that a benchmarking approach should be used to set baselines and assess student outcomes. They stated that benchmarking used in the Teaching Excellence and Student Outcomes Framework (TEF) demonstrates that it would be possible to set baselines for intervention that reflect a provider's benchmark, which could increase over time in response to successful intervention, while establishing a high absolute baseline that would exclude from regulatory intervention providers performing well above the baseline. The view was that this would resolve various issues about the role of contextual information in assessment.
25. Other arguments were that benchmarking is important to ensure that providers' operating contexts and missions, and regional differences between them, are considered; that this

approach would be more transparent; and that benchmarking would ensure that factors that may be outside providers' control, but which impact on student outcomes, are considered, including socioeconomic factors. Some respondents stated that benchmarking was essential for a meaningful, robust, and fair assessment of student outcomes.

26. A number of respondents suggested that approaches to maintain quality and standards must not have unintended consequences for access and participation and efforts to expand equality and diversity across the sector. It was suggested that the OfS should complete an equality impact assessment to consider how the proposals could impact the access and participation of students from disadvantaged backgrounds, by driving risk-averse provider admissions.
27. Other respondents suggested that if the OfS proceeds to assess provider performance in absolute terms, this should be focused on exceptional cases, given the diversity of the higher education sector, student groups, and established inequalities reflected geographically and between student groups.
28. Others thought the OfS should ensure that providers delivering courses to meet the needs of certain student groups, including specialist and strategically important courses that do not have high levels of demand, are not disincentivised, because otherwise this could damage the diversity and vibrancy of the higher education sector.
29. A further suggestion was that, if the OfS proceeds with the use of baselines that do not change to reflect differences between demographic groups, it should ensure that such assessment balances reliance on absolute performance with progress towards the commitments in access and participation plans.

Our response

30. Our phase one consultation set out our proposal to continue to have numerical baselines for condition B3 that apply to all providers' absolute performance, and that the assessment of condition B3 should heavily rely on this. We also proposed that we should continue to give appropriate consideration to contextual factors for an individual provider, as a way of ensuring that we have properly interpreted its absolute performance. This would allow us to assess whether there are any contextual factors that might contribute to performance being below a numerical baseline, and which might explain the absolute performance. We explained that our proposal meant that considering contextual factors in this way could mean that, if a provider's performance was below one or more numerical baselines, this could potentially be judged to be acceptable if contextual or other relevant factors accounted for such performance.
31. We have reflected on comments that context did not appear to be an integral part of our assessment. We now propose that our minimum requirements are tested through a combination of assessing:
 - the student outcomes a provider has delivered against numerical thresholds
 - the context in which those outcomes were delivered.

This is explained further in the phase three consultation, 'Consultation on regulating student outcomes'.⁵

This change in our approach has introduced the use of the term 'numerical thresholds', which we have used throughout this document in addition to our terminology of 'baselines' from the phase one proposals.

In this document:

- 'baseline' or 'numerical baseline' has been used to refer to the proposals set out in our phase one consultation, and our analysis of the responses we received
- 'numerical threshold' has been used to refer to the proposals we are setting out in phase three of our consultation.

32. We set out the factors that we proposed to consider, and were interested in views about other contextual factors we should take into account and the weight we should place on them. We have considered the responses to the consultation and have set out further detail of how we propose to consider contextual factors, including student characteristics, both in setting the numerical thresholds and in assessing an individual provider's performance. More details of our approach can be found in Proposals 3 and 5 in our separate consultation document.⁶ These proposals mean that we will take account of most of the factors respondents identified as relevant, including student characteristics (such as disability or socioeconomic background) and region of study, where we consider this appropriate. In proposing to take these factors into account, we consider that concerns about the impact on particular provider types will also be addressed.
33. We are proposing to make changes to the contextual factors that we will consider when assessing a provider's performance. These are set out in Proposal 5 of the phase three consultation document, and include the use of individual provider benchmarks in our consideration of provider context.
34. We recognise that there may be positive reasons for students leaving their courses or pursuing careers that are not managerial or professional, and this is also reflected in the approach proposed in our phase three consultation. We also consider that providers have a considerable influence over the outcomes students achieve and that factors beyond their control are not so extensive as to make it impossible to establish minimum requirements including numerical thresholds.
35. We have considered comments that suggested that established differences in outcomes between different student groups are already, and would more appropriately continue to be, addressed through APPs. As respondents' comments recognised, APPs in their current form

⁵ See www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/student-outcomes/.

⁶ See Proposals 3 and 5 in www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/student-outcomes/.

deal with gaps in performance between different groups, and providers will have different priorities and areas for improvement. Our view is that closing gaps in performance between different student groups is only meaningful once minimum requirements for quality are met, and APPs are not currently intended to secure a minimum expectation of performance for any group of students.

36. APPs are therefore not, in their current form, an appropriate mechanism to secure the policy objective that students from all backgrounds should expect providers to deliver a minimum level of outcomes. They are also only a requirement for providers that are registered in the Approved (fee cap) category of the OfS's Register of higher education providers and want to charge a higher fee limit. There is therefore a large number of registered providers that are not subject to regulation through APPs. A reliance on APPs would therefore not be sufficient to protect all students and improve outcomes at all providers.
37. The Secretary of State for Education has published new guidance on future APPs.⁷ We will have regard to that guidance, and our view is that the proposals set out in the phase three consultation are consistent with the Secretary of State's desire to ensure that more students from disadvantaged backgrounds successfully complete their course and progress to highly skilled employment.
38. We have considered the impact of our proposals on the incentives for providers to recruit and support students from underrepresented groups. In general, we do not consider that the impact of our proposals will be detrimental to the interests of those students, or potential students, because:
 - a. It is a consequence of our policy intention that those providers that do not achieve student outcomes that meet the minimum requirements we expect will face regulatory intervention. We maintain that it is appropriate that this approach should apply regardless of the characteristics of the students that a provider recruits. As our intervention is likely to be escalatory, providers will be given the opportunity to improve their outcomes over an appropriate period of time.
 - b. There is a direct benefit to students from underrepresented groups and future students if their provider takes action to improve the outcomes it delivers.
 - c. The impact of our proposed numerical thresholds would be that a relatively small proportion of students from underrepresented groups would be studying at providers that might be subject to regulatory action.
 - d. Our proposals in relation to the TEF would reward providers that demonstrate excellence in teaching and student outcomes for the particular types of student they recruit by considering, if a provider is judged to meet our minimum requirements, how well it performs against its benchmark.
 - e. Our requirements in APPs will complement the approach set out in B3.

⁷ See 'The future of access and participation', available at www.officeforstudents.org.uk/advice-and-guidance/regulation/guidance-from-government/.

39. We have noted the points made by respondents about the possibility that a provider might fail to meet a numerical threshold for its students despite having met the minimum requirements for the other ongoing conditions for quality and standards.⁸ We agree that this would be possible, and this is intentional. We have set out in the phase three consultation that we consider it necessary to act where a provider does not deliver positive outcomes for its students. This is because we consider student outcomes to be an essential part of our approach to regulating quality, but also because condition B3 has a particular function in securing value for money for students and the taxpayer.
40. We consider that our wider approach to regulating quality and standards will address the concerns raised by respondents that we have a narrow focus on data; the other B conditions ensure that there are regulatory requirements relating to other aspects of quality.
41. In relation to the comment that we have already reached a firm view on the use of numerical baselines, the phase one consultation was published at an early stage of policy development to allow us to gather views about our proposed approach to inform how we might develop more detailed proposals. We were open-minded about the extent of changes to the proposals that might be necessary as we developed our phase three proposals. We have made changes where we think these are appropriate, and have set out our responses to the consultation responses within this document and in the phase three consultation. We will carefully consider responses to the phase three consultation before any decisions are taken.

Setting higher and more challenging baselines

42. The phase one consultation proposed that more challenging baselines would be set that would apply to each indicator and all providers.
43. A small number of respondents supported the OfS proposal to set higher and more challenging baselines, because they thought this would better enable the OfS to deliver a risk-based approach to regulation and to maintain public confidence. It was suggested that setting baselines at too low a level would mean some providers will only seek to meet the baselines, which could lower the quality of provision.
44. Others commented that they were not clear that current baselines are too low and commented that any changes would need to remain proportionate for all types of provider to achieve and be fairly assessed against. Other respondents commented on the potential impact that higher baselines might have.
45. It was suggested that there might be a disproportionate impact on further education colleges and small providers, and that assessing performance against higher baselines could pose an unfair risk to providers that are providing high quality education as assessed through conditions B1, B2 and B4 but are catering to predominantly disadvantaged students.
46. As with other aspects of the proposals, some respondents suggested that higher baselines might result in a change to providers' recruitment behaviour, leading to their becoming more selective in their recruitment to avoid regulatory intervention, which would disadvantage

⁸ Our proposals for regulating these other aspects of quality and standards were set out in our phase two consultation.

students from underrepresented groups. Respondents also suggested that some subject areas would be placed at greater risk of 'low' performance than others, and that this would increase the possibility of course closure, which would narrow the range of subjects of offer.

47. Another respondent commented that raising the baseline for continuation from, for example, 75 to 90 per cent would lead to a loss of recognition of the achievement of the 75 per cent of students who succeed. If a provider were to be deregistered based on a failure to achieve the higher baseline, that would prevent the 75 per cent of students from having the opportunity to succeed.
48. Some respondents were concerned about the introduction of more challenging baselines while the sector is recovering from the coronavirus pandemic.

Our response

49. We are of the view that some baselines used during initial registration were too low. For example, the baseline for the percentage of students progressing to graduate employment or further study was 35 per cent, fewer than half of students who complete their studies and well below the average sector performance. We therefore consider that there is considerable scope for at least some of the numerical thresholds we set to be higher than the baselines we used during the initial registration period. We have set out in our phase three consultation the full range of factors which we propose taking into account when setting numerical thresholds. Further information on this can be found in Proposal 3.
50. Our phase three consultation document sets out the impact that we expect these numerical thresholds to have in general. We have also identified the impact for each of the proposed numerical thresholds on providers and students from underrepresented groups in a separate document.⁹
51. We have set out in our phase three consultation how we propose to take account of the statistical uncertainty in small cohort sizes and consider that this will address points about the potential impact on small providers including further education colleges. We consider that beyond this the proposed numerical thresholds will affect any provider of any kind (regardless of its size or corporate form) that has performance below our minimum requirements.
52. We have considered the comment that a numerical threshold of, for example, 75 per cent would lead to a loss of recognition for the 75 per cent who succeed in favour of the protection of the 25 per cent who do not, and that if this resulted in deregistration this would deprive the 75 per cent of the opportunity of success. We have set out in our phase three consultation our proposals the general approach we would be likely to take to intervention when a provider was judged to be delivering below the numerical threshold for a particular indicator. Our intervention is likely to be escalatory, normally including the use of specific ongoing conditions requiring improvement, and therefore for registered providers there will be an appropriate period where a provider is likely to be given the opportunity to improve. However, where a provider does not improve over time, or there is a significant breach of condition B3 when we

⁹ See 'Setting numerical thresholds for condition B3', available at www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/student-outcomes/.

first intervene, we may take the view that current and future students would be better served by other providers that have demonstrated their ability to deliver outcomes that meet our minimum requirements.

53. We have considered whether it remains appropriate to continue with our proposals at this time in light of the pandemic. We take the view that the imposition of minimum requirements should continue to apply regardless of the circumstances in which a provider is operating. We note that the data currently available for continuation shows that there does not appear to have been an immediate impact on this indicator during the pandemic. We do not yet know the longer-term impact on the sector or on student outcome indicators; for example, the effect of the pandemic on responses to our Graduate Outcomes survey has not yet been tested for graduates from the 2020-21 academic year. However, we think it would be wrong to set lower minimum requirements as a consequence of the pandemic, which will only have a temporary impact on student outcomes. We expect providers to do more to support future students who may have had lost learning in their secondary education as a result of the pandemic.
54. We have nonetheless taken a number of steps to allow us to understand and, where appropriate make allowances, for the immediate impact of the pandemic:
 - a. We have set out clearly that the evidenced effects of the pandemic on a provider's student outcomes would be one of the contextual factors we would consider in any decision on whether a provider has complied with condition B3.
 - b. We have disaggregated our data sets to show each of the last four years. This will allow us to see whether performance in the most recent years is markedly different from previous performance.

Proposed indicators – general points made about all indicators

55. The phase one consultation proposed to define our requirements by setting numerical baselines for acceptable performance for indicators relating to continuation, completion and progression to managerial and professional employment or higher-level study.
56. A large number of comments related to the use of indicators in general.
57. Several respondents suggested that continuation, completion and progression are unconnected to the quality of the course and should not be included in a definition of high quality. This argument was made most strongly for progression and completion, and less so for continuation. Respondents argued that there are factors outside the control of the provider that affect student outcomes, and that the use of these indicators could have a negative effect on widening participation by discouraging providers from recruiting students from underrepresented backgrounds, or via non-traditional routes, such as mature students entering higher education via Accreditation of Prior Experiential Learning. Respondents took the view that evidence shows that students from underrepresented groups tend to have worse outcomes compared with students from groups who are more widely represented. Because of this, some respondents suggested that the use of these indicators favoured more research-intensive providers with higher entry tariffs.
58. Some respondents suggested that continuation, completion and progression measures only partially capture student achievement and outcomes and do not consider the wider social

benefits derived from higher education. Other respondents argued that there are often positive reasons for students leaving higher education.

59. Respondents also commented that the use of continuation and completion indicators in particular would disincentivise the development of flexible and modular learning and offers of credit transfer models of delivery, and that this was in contradiction to the government's policy aims. Some respondents asked how a continuation indicator would work for courses that are less than or one year of length. It was also suggested that, the more flexible study opportunities and modes a provider offers, the harder it is to measure completion for that provider.
60. Respondents suggested that flexible provision including distance learning, micro-credentials and modular study modes should be treated separately in the student outcome indicators, as it is not easy to compare with the outcomes from more traditional provision. Respondents also suggested that courses with intermediate awards (such as HNC-HND and integrated masters') and courses with integrated foundation years were particularly vulnerable to misrepresentation in student outcome data. Respondents suggested that many students on these types of courses change routes once they have the information and experience to understand what course or subject is most appropriate for them and their needs. Students moving to other pathways or courses at this stage should not be seen as failure, but rather that foundation courses are working as intended.
61. Students studying via short courses or modular study have often completed their studies before the reporting point, and students who opt to take extended study breaks risk being overlooked when compiling outcomes data. There should be an option to record such students separately, so that a provider's flexible approach does not compromise its student outcomes data.
62. It was suggested that clear and unambiguous definitions of continuation, completion and progression need to be set, so that it is clear what data is being measured and there is no risk of potential double or triple counting during aggregation.

Our response

63. We consider that providers should be held accountable for the outcomes of their students, as the contribution that they make to their ability to complete their higher education courses and succeed afterwards is considerable. Given the substantial financial investment in higher education made by both individual students and the taxpayer, there are strong value for money arguments for considering the outcomes from that investment.
64. Our proposals in relation to condition B3 sit alongside our broader approach to regulating quality and standards.¹⁰ Our approach to regulating quality and standards in general, condition B3 specifically, and assessing excellence through the TEF, recognises the importance of student outcomes as an integral aspect of high quality. While we recognise that other aspects of quality and standards are important, we disagree with respondents who argued that outcomes are unconnected with the quality of a course. We can see through various analyses,

¹⁰ See www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/student-outcomes/.

including our Exploring Student Outcomes analysis and individual provider benchmarking, the impact that providers can have on outcomes for similar student populations.

65. However, as we set out in Proposal 2 of our consultation, we acknowledge that it would not be appropriate to use the proposed outcome measures in a way that required all providers to achieve extremely high absolute performance as a minimum. For example, there will always be some students whose life experiences mean that not completing a qualification is the right choice for them. We also recognise that some students may leave higher education for positive reasons, and that students' motivations for study extend beyond gaining relevant employment. However, we do not accept that these factors are so extensive as to make it impossible to establish a minimum requirement, including numerical thresholds.
66. The outcome indicators we propose to use for condition B3 only partially capture the social benefits derived from higher education. As set out in our phase three consultation, one of the tests for the indicators we will use is the availability of data. We will always be open to considering use of other robust data sources and evidence that enable us to assess the wider benefits of higher education as context of a provider's performance.
67. We have set out in the phase three consultation document, in Proposal 3, our position on the use of existing student outcome indicators in relation to modular provision. In general, and as set out in the phase one consultation, we expect to consult further on appropriate measures of success for this kind of course.
68. We agree that it is important to have a clear and unambiguous definition of the indicators we will use. These are set out in a separate data consultation.¹¹ This also sets out how interim awards and students transferring from one provider to another would be counted within the indicators. In both of these circumstances our proposal is that this would be counted as a positive outcome.

Proposed indicators – continuation

69. The phase one consultation proposed that student continuation would be an indicator used to assess student outcomes. We consider continuation to be a relevant measure because continuation rates provide a strong indication of whether a student has been appropriately recruited onto a suitable course that matches their abilities and aspirations, and whether they then receive the support they need. Continuation rates identify at an early stage whether a student's investment of money and time has, for whatever reason, not led to a successful outcome. The continuation rates indicator focuses on an early point in the student lifecycle, currently identifying where students leave their studies within the first year for full-time students, or within the first two years for part-time students.
70. There was broad support for the use of continuation as a measure. Respondents agreed that it is a legitimate and reasonable measure of quality, and has been used within the sector as a measure for many years. Respondents suggested that it can be measured reasonably and

¹¹ See www.officeforstudents.org.uk/publications/student-outcomes-and-teaching-excellence-consultations/outcome-and-experience-data/.

effectively and can be substantially influenced by a provider. Respondents also suggested that information about continuation is important to facilitate student choice.

71. A very small number of respondents did not agree with the use of continuation as a measure. This was because they suggested it was not a measure of the quality of the course and might undermine initiatives to promote widening participation, particularly where personalised flexible-return approaches could be developed. Respondents suggested that providers might limit recruitment of students with characteristics that put them at a higher risk of non-continuation, for example mature students.
72. A larger group of respondents suggested that because student characteristics have an effect on continuation rates, these should be taken into account when setting the baseline, assessing providers against it or both. Respondents suggested that considering student characteristics when using a continuation indicator would mitigate any negative impact on widening participation.
73. A small number of respondents commented that non-continuation may be in the best interests of some students because of their personal circumstances and using a continuation indicator might encourage providers not to support this choice. This might include supporting decisions to transfer to another course within a provider. Another respondent suggested that there can be positive reasons for non-continuation, such as students pursuing opportunities that arise during their studies as a result of the professional networks to which industry specialist providers can connect their students.
74. Some respondents commented on what they saw as potential consequences of assessing continuation rates. For example, some respondents suggested that use of a continuation indicator might make providers conservative about introducing courses that might have high non-continuation rates, despite employer demand. Examples of the courses this might affect were computer sciences, business, mass communication, engineering and technology. Other respondents suggested that providers might be reluctant to change the content of courses with high continuation rates, and stated that refreshing course content is crucial to prepare students to enter the world of work and to meet labour market needs.
75. Other comments were that while the use of continuation rates was valid, this should be as a 'soft indicator' to identify potential regulatory concern, rather than as a regulatory requirement in its own right.

Our response

76. We note that, while there was disagreement with the proposed approach to using absolute performance, most respondents agreed with the use of continuation in the assessment of quality. In relation to the comments from a small number of respondents that it was not a measure of quality, we see student outcomes, including continuation, as an integral part of an assessment of quality. We agree that the use of continuation rates on their own would not be an effective measure of the quality of a course, but all higher education courses at registered providers are also subject to other conditions of registration that cover requirements for the academic experience and resources and support for students. We also agree that continuation indicators should be used as a soft indicator that might trigger investigation into other quality matters.

77. However, using continuation as a soft indicator would not, in our view, be adequate to protect students and the taxpayer from weak outcomes. Our view is that student outcomes will be affected by the quality of the course, but that they should also be assessed in their own right, both as a distinct aspect of quality and because of value for money and consumer protection considerations.
78. We have considered the responses that suggested student characteristics should be taken into account when assessing continuation rates in order to mitigate any possible risk to widening participation. Our view is that providers should be aiming to deliver similarly positive outcomes for students from all backgrounds, but we recognise that student characteristics have historically affected performance. We have therefore set out in our phase three consultation how we propose to address this. As set out in paragraph 73, we acknowledge that some non-continuation may be in the best interests of students or for positive reasons, but think that this is not likely to be the case for the majority of students who do not continue on their course. We do not think this prevents the use of such an indicator to assess performance.
79. In relation to comments that providers might be disincentivised from delivering courses that are in high demand from employers but might have high non-continuation rates, we consider that demand from employers and students should incentivise providers to continue to deliver those courses and to work to improve the continuation rates of students. However, we recognise that the subject may be a relevant matter of context, and we have therefore set out in our phase three consultation how we have considered this issue further.
80. Respondents also suggested that providers may be unwilling to change the content of their courses if these courses have high continuation rates. The other conditions of registration for quality and standards will provide mechanisms to ensure that overall course content and academic experience remain high quality. For example, our phase two consultation set out proposals for condition B1 which, if implemented, would require courses to be 'up-to-date'.

Proposed indicators – completion

81. The consultation proposed that we would continue to assess student outcomes by constructing a student completion indicator. We consider that this indicator is one of the most relevant measures of student outcomes available, and tells us whether a provider is recruiting students able to succeed through to the end of its courses. For students studying for higher education qualifications in the UK who are taught or registered by an English higher education provider on courses funded by the OfS, this indicator can be reliably constructed from data that is already available.
82. Respondents broadly recognised that completion was a legitimate indicator of student success and, like the continuation indicator, a measure that has been in use within providers for a long time and an outcome over which they have some control. Respondents noted that acceptable rates of completion are unlikely to be achieved without good-quality resources and academic support and this is likely to mean the provider is also delivering high quality provision.

83. Several respondents asked for further information on the detail of the construction of the indicator, for example whether student transfers would be treated positively and whether data would be split by degree classification.
84. Some respondents suggested that projections of completion (the method used to construct the current completion indicator) should not be used as a basis for setting baselines or making regulatory interventions. However, other respondents recognised that using cohort tracking methodology to conduct an actual completion rate would result in data being heavily lagged and this was also not ideal for regulatory purposes.
85. Several respondents suggested that the use of a completion indicator might have unintended consequences; for example, that providers might concentrate on achieving completion rather than maintaining academic standards and this could result in grade inflation.
86. Some respondents made specific suggestions about the construction of the indicator. These included that:
 - a. Completion with a different or lower award should not be considered a negative outcome within the indicator, as that might be what best meets a student's needs and circumstances.
 - b. Student transfers should be counted positively. The receiving provider should not be penalised for any poor completion rates arising from this transfer if the majority of the credits were delivered by the previous provider.

Our response

87. We note that there was broad agreement regarding the use of a completion indicator as part of an assessment of student outcomes. Our data consultation on the definition and construction of the indicators proposes two possible approaches to measuring completion. One is a projected completion rate, while the other would be an actual completion rate based on tracking cohorts of students. Respondents to the phase one consultation recognised that there were pros and cons of using either approach, and we are now seeking views on which approach we should use. In our phase three consultation we have proposed numerical thresholds based on both methods.
88. Where it is unclear from the available data whether a particular type of outcome should be viewed as positive within the construction of an indicator, we propose to treat it as either positive or neutral, rather than negatively. For example, we propose to treat the outcomes of those students who have transferred to another course or provider as neutral, and also recognise that completion of a lower award than originally intended could still be interpreted as a positive outcome. We take the view that this, in combination with the way we propose to set numerical thresholds and the use of other information about an individual provider's context in reaching regulatory judgements, will ensure that the use of completion as an indicator of quality remains appropriate.
89. In relation to comments that the use of a completion indicator may lead to grade inflation if providers are reluctant to fail students to achieve good completion rates, we consider this unlikely. as failure to uphold academic standards in this way could lead to regulatory intervention in relation to other conditions of registration, and to degree awarding powers for

providers that hold such powers. These other regulatory tools therefore create strong incentives for providers to maintain standards as they improve completion rates.

Proposed indicators – progression

90. The consultation proposed that we would continue to assess student outcomes by constructing an indicator for progression to higher-level study or professional and managerial employment.
91. We consider that a progression indicator is appropriate to use for regulatory purposes, because this will tell us whether a provider's students have successful outcomes beyond graduation. Although individual students will define their success beyond graduation in relation to their own goals and motivations, it is important to ensure that graduates are achieving outcomes consistent with the higher education qualification they have completed. Low rates of progression into employment and higher-level study destinations commensurate with the qualification they have completed may suggest that a course has not equipped students with knowledge and skills appropriate to their intended learning aims, or did not effectively support them to transition into the workplace.
92. The consultation also proposed that we would use the Graduate Outcomes (GO) survey to construct the progression indicator rather than the Destinations of Leavers of Higher Education survey. This would mean that we would only have progression indicators for the cohorts of qualifiers in the 2017-18 academic year onwards. We proposed that employment that does not have degree study as a prerequisite would not be counted as a positive outcome for the purposes of the indicator.
93. Only a small number of respondents suggested that progression to higher-level study or professional and managerial employment was an appropriate measure of student success or quality. A larger number commented that it was not, because the purpose of higher education is broader than ensuring students achieve professional or managerial employment. Some respondents suggested that using a progression indicator also fails to recognise the responsibility of students to achieve successful outcomes. Others suggested that progression was a suitable measure of value for money in terms of the return to the Treasury from graduate income, but this was not the same as an indicator of quality.
94. Several respondents suggested there were wide-ranging factors, outside the control of the provider, that affect progression rates; for example, the region in which the provider and its students are located, and whether it is a deprived area or an area with high skills needs. Respondents also suggested that graduate employment outcomes are significantly influenced by social factors, structural inequalities and employer bias. A specific example was given in relation to land-based subjects, where most students were said to be female and outcomes that could be affected by sex discrimination in the labour market could mistakenly be attributed to providers. Respondents stated that because of these social issues graduate employment rates tend to be higher for students from more advantaged backgrounds, regardless of the effective performance of the institution.
95. A number of respondents also considered that providers were not able to sufficiently influence progression because of the role played by student choice, for instance where students chose employment that is not at managerial or professional level based on their personal interests or

values. Personal circumstances may also be a factor: some students may not immediately take up employment, for example students with severe disabilities, some with caring or other commitments, or students who are studying for reasons of personal health or wellbeing.

96. Respondents suggested that, if a progression indicator was used, performance against a benchmark value should be considered, because they considered performance to be heavily dependent on factors relating to sex, ethnicity, socioeconomic group, location of graduate and entry tariff. Respondents also suggested that, if progression data was not benchmarked, it would be likely to penalise providers that focus on widening participation and those located in relatively economically disadvantaged regions. It was also suggested that only considering absolute performance would have a bigger impact on providers that deliver certain subjects.

Construction of the indicator to use 'professional and managerial' employment

97. Some respondents commented that certain vocational or work-based courses do not lead to employment classed as 'professional and managerial', such as early years education (nursery), veterinary nursing, teaching assistants courses, but that these lead to relevant and valuable employment outcomes. There was a suggestion that the indicator should therefore measure progression to 'employment' or 'suitable employment'.
98. Other respondents commented that graduates from the arts or humanities might enter roles of social importance that might not be classified as 'professional and managerial', such as entry level roles in the cultural and heritage sectors, or working for charities or community organisations. Other respondents said that in the land-based sector, graduates usually enter their chosen sector at a level below that generally recognised as graduate employment, but progress swiftly thereafter. A progression indicator that only measures professional and managerial employment may therefore not capture successful outcomes for these graduates.
99. Other respondents pointed out that the pilot subject-level TEF saw many Bachelor of Laws (LLB) graduates taking a subsequent vocational law course who were considered as non-progressing, even though this is one of the best outcomes for LLB students.
100. A number of responses suggested that freelance or self-employment should be recognised as a successful outcome. This was said to be because graduates draw heavily on the skills developed during study. Some respondents noted that self-employment and portfolio careers are increasingly the preferred career path for many graduates in disciplines such as the creative and performing arts. It was suggested that a progression indicator that did not count self-employment as a positive outcome might disincentivise providers from encouraging entrepreneurship and start-ups.
101. Other respondents suggested that the indicator should count voluntary activity, and activity that contributes to civic and social wellbeing, as a positive outcome and a valid progression from degree level study. This is because respondents said there is an increasing proportion of graduates entering the world of social enterprise and the third sector following graduation, and it is important that the voluntary sector also benefits from skilled and qualified graduates.

Scope of the Standard Occupational Classification definitions

102. A number of respondents noted what they saw as issues with the use of the Standard Occupational Classification (SOC) definitions to determine whether a role is classified as graduate employment. They suggested that SOC codes include too many similar roles where

some are considered graduate jobs and others not, and that they are problematic for specialist providers, where courses to a career such as 'artist blacksmith' have no SOC equivalent. One respondent noted that SOC must be kept up to date to ensure that providers operating within innovative and dynamic sectors are being properly and fairly regulated.

Progression from Level 4 and 5 courses

103. A number of respondents highlighted issues in relation to progression outcomes for students completing courses at Levels 4 or 5. They suggested that students studying at these levels might not have any intention to progress to a Level 6 qualification, and may not aspire to managerial or professional employment for a number of years following graduation.

Alternative suggestions

104. Some respondents suggested that a progression indicator should take into account how competitive each employment sector is and the likelihood of gaining a job in the relevant profession, which would provide context for the distance travelled and value added by a provider.

105. Another suggestion was that a progression indicator should refer to the proportion of graduates retained locally, to encourage greater collaboration between providers and local employers in line with the economic strategy.

Issues with the Graduate Outcomes survey

106. One respondent suggested that the current process to measure progression through a survey does not capture the outcomes for all leavers (the response rate for GO is around 50 per cent). Therefore, respondents suggested, it is not representative of all student outcomes and might be particularly problematic at subject level, where progression numbers under this measure might be very low.

Use of LEO data as a supplementary indicator

107. The consultation asked respondents for views on using the Longitudinal Education Outcomes (LEO) dataset to provide further indicators in relation to graduate outcomes.

108. Some respondents saw the benefit of using LEO, due to its reliability as a dataset; the use of contextual background information such as the region, age and ethnicity of the graduates surveyed; and the time period covered, allowing capture of a wide range of students, including those who take longer to progress to professional careers.

109. Others considered that LEO provides information that could be used as supplementary data alongside the Graduate Outcomes survey, and that it provides data that helps prospective students make informed study choices and providers plan their provision, as they can see the progression of their alumni over a longer period.

110. However, respondents overwhelmingly did not support using the LEO dataset. This was because it was considered to be relatively new and excludes some groups, such as international and postgraduate students and those who gain employment or become resident outside the UK.

111. It was also considered that a potential time lag of up to 10 years between graduation and reporting meant that the dataset would not be an accurate reflection of the current quality of

the higher education at a specific provider, and would not be useful in assessing the progression outcomes for newly registered providers.

112. Additionally, respondents did not consider that salary data provides an indicator of the quality of a course or of a provider, the value to society of a particular career or the personal benefit gained by the student either from their studies or from their choice of employment.
113. It was also noted that the use of LEO data could affect providers' behaviour, for instance by reducing recruitment onto nursing or other courses leading to less highly paid employment, or providers encouraging graduates to move out of lower-income regions in search of more highly paid employment.
114. In addition, it was considered that its use could disadvantage providers and courses that recruit a higher proportion of students from underrepresented groups because of their unequal access to some career opportunities, or could disadvantage providers and courses recruiting a high proportion of female students because of the gender pay gap in some employment fields.

Our response

115. We have considered the points made by respondents that progression is an inappropriate indicator of quality for the reasons described by respondents in paragraphs 94 to 96. We accept that students' motivation for entering higher education may not be solely about achieving a managerial or professional employment outcome, but we take the view that considering the extent to which a provider is preparing students to be able to take up graduate level employment or further study is in the interests of both students and taxpayers. For this reason, we also do not think it is appropriate to construct this indicator to consider any employment as a positive outcome. The detail of the proposed construction of progression indicators is set out in our indicators consultation.
116. We agree with respondents who suggested that progression is relevant as a measure of value for money. We set out in our phase one consultation our view that an assessment of student outcomes is an integral part of our consideration of quality, but that it also relates to value for money for both the student and the taxpayer. A low level of progression to employment that has a degree as a prerequisite may suggest that students are not developing relevant knowledge and skills.
117. In relation to the construction of the indicator, we recognise that some courses delivered by some specific, specialist providers result in employment that is directly relevant to the vocational courses studied and is highly sought after. This employment is sometimes not classed as managerial or professional on the basis of SOC coding – for example construction, artistic blacksmiths and nannies – and so would not count positively within a progression indicator. All the examples provided by respondents have previously been taken into account as relevant context as we have assessed progression outcomes for individual providers, and we would continue to do this. This means that in practice, even though progression from these courses to a relevant occupation would not count positively in the construction of the indicator, we would be likely to consider it positively within our assessment of that provider's context. As we will be considering performance at a subject level this will make it easier to take account of this circumstance.

118. Other examples that were provided by respondents, such as teaching assistants and veterinary nurses have been reclassified in relevant codes in SOC2020, and would therefore count as positive outcomes. Our consultation on the construction of the indicators provides details about the construction of the progression indicator. It also explains that we propose to count positively within the indicator graduates who are self-employed and undertake voluntary work that maps to SOC groups 1 to 3.
119. In relation to progression to employment from Level 4 and 5 qualifications, we agree that students might intend to study at these levels without any intention to progress to a Level 6 qualification, and that their progression is less likely to be directly into a managerial or professional employment. This is apparent in the current progression rates from Levels 4 and 5. We are proposing to address this concern by treating 'other undergraduate' (which includes Level 4 and 5 courses) as a separate level of study from 'first degree'. This means that a separate numerical threshold would be set for courses at these levels. The value of that numerical threshold would be influenced by the general sector trends for progression and would therefore be likely to be set lower than the numerical threshold for first degree. We are also proposing that split indicators will be constructed to allow us to separately identify level 4 and level 5 courses. This will enable us to see whether there are a particular trends within the 'other undergraduate' grouping.
120. In relation to comments that SOC codes need to be kept up to date, SOCs are produced by the Office for National Statistics and are normally updated every 10 years. It is possible that new courses and occupations will fall outside the classifications, but it will always be possible for us to take context into account to overcome any issues in relation to the mapping of courses and relevant occupations to SOC codes.
121. Beyond this, we recognise that the application of any standard classification will not be able to fully reflect the nuances of individual employment circumstances in all instances. However, we consider that the classification produced by the ONS is a robust and nationally recognised method of classifying occupations, which supports a transparent and well understood definition.
122. We think that there is validity in some of the points made about the factors that may influence progression rates and may not always be in the control of a provider. However, our view is that providers do and should have a considerable influence on the outcomes of their students, and the existence of factors beyond their control does not prevent us being able to use progression as a measure, particular if performance across the sector as a whole may be affected by similar external factors. This is because we propose, when we set numerical thresholds for the assessment of progression, to take account of performance across the sector, and this will account for the effect of many of these factors. Beyond this, the consideration of context for an individual provider further mitigates the concerns raised about the factors that might influence progression rates. We have set out in our phase three consultation how we propose to set numerical thresholds and how our assessment will take account of context.
123. In relation to comments about the use of the GO survey, and that the response rates mean it is not representative of student outcomes, we recognise that this is survey data. However, we consider that the GO survey is the best source of information on what graduates are doing. We consider that it is appropriate to use it to construct indicators for use in assessing

condition B3 alongside consideration of contextual factors. Recent research commissioned by the Higher Education Statistics Agency has shown that the responses are, in general, sufficiently representative to use without weighting and we have proposed an approach to response rate thresholds and data suppression in the technical consultation that will allow us to take into account localised instances where the data may not be sufficiently robust to use for regulatory purposes.

124. When the GO survey was developed, the census point was set later than that of the Destination of Leavers from Higher Education survey, at 15 months after a qualification is awarded rather than six months. This recognised that it was more likely to be more meaningful to survey students 15 months after graduation, when they could be expected to have taken up employment or study opportunities, and addresses the point made by some respondents that graduates may start their careers in a non-managerial position but progress swiftly after this. It was recognised that this extended period may nevertheless increase the likelihood of lower response rates, and this balance was tested through consultation with providers and others.
125. We note that respondents were not in favour of using graduate salary-based datasets such as LEO to construct indicators. Our interpretation is that this is in part because respondents thought we were proposing to set baselines for graduate salary-based indicators. Our phase one consultation had suggested that we might consider using LEO data to construct supplementary indicators, which might help with the interpretation of data from GO. We had not proposed to set baselines for any supplementary indicators. We are still of the view that graduate salary-based indicators – whether calculated from LEO or GO data – may provide useful context to interpret providers' performance in relation to progression to graduate employment, but also consider that graduate salaries data would not enable us to measure progression into graduate employment as effectively.
126. Our proposed progression measure is more appropriate for use, as it enables us to better meet our policy objective of ensuring progression to graduate employment rather than to highly paid employment. However, we recognise that there is a strong link between graduate employment and earnings. We proposed in our phase one consultation that LEO data could inform our general monitoring activity. We also expect that LEO data could be part of the additional data we review when considering contextual factors for an individual provider when we assess compliance with condition B3 (see proposal 5 on implementation in the phase three consultation document).

Proposed indicator – degree outcomes

127. The phase one consultation proposed that we would no longer consider the percentage point gap between the proportion of leavers from Level 6+ undergraduate degrees who were awarded first or upper second degree classifications and those who were awarded classified degrees. We proposed that this should no longer be a measure in relation to the baseline regulation of student outcomes for providers that have an approved APP, because the indicator relates to the gaps in degree classifications awarded to students with different characteristics, rather than considering an absolute measure of the outcomes a provider delivers for its students. We proposed that we would continue to consider this indicator for any provider without an approved APP.

128. A very small number of comments about this proposal pointed out that many further education colleges do not have an APP and may be disadvantaged by this proposal, and that further clarity is needed on how providers without an APP will be affected by this approach.

Our response

129. The few respondents who commented were concerned that the proposed approach would disadvantage providers without an APP. Our proposal was intended to remove duplication of regulation for providers that have an APP, but we have further considered whether it is necessary to continue to include this indicator. Given the comments on complexity of the number of indicators that would be generated for condition B3, we have decided that this indicator is not necessary, given that it does not set numerical thresholds for absolute performance and instead looks at gaps in performance. We have therefore decided that condition B3 will not include an indicator for gaps in degree outcomes.

Mode and level of study

130. The phase one consultation proposed that indicators would be constructed separately, both for full-time and part-time modes and for different levels within each mode of study.

131. The phase one consultation also invited views on whether any additional indicators should be considered, for example: the separation of first degree courses that include an integrated foundation year from other first degrees; top-up courses; sandwich year provision; apprenticeships; and accelerated degrees or distance learning.

132. Respondents were generally in agreement with the levels for the indicators proposed in the consultation, because they are well established and already routinely engaged with by providers. Respondents said it was appropriate to allow different baselines for different levels of study to be considered. These have been carried forward in our phase three consultation.

Apprenticeships

133. A significant number of respondents said that apprenticeships should be considered as an additional indicator because these courses have distinctive characteristics and the intentions and experiences of students are often different. Respondents took the view that student outcomes are likely to be better if learning is linked to employment and the programme is designed to provide a direct route into graduate-level employment.

134. Respondents also suggested that, as study is delivered in partnership with an employer, there are factors that may affect outcomes that are not relevant for other types of courses, such as the study being paid for by the employer and changes within an employer organisation or a student's employment status affecting their ability to continue.

Integrated foundation year

135. Respondents were generally in favour of separating courses with an integrated foundation year into an additional indicator. Respondents were of a view that, because courses with an integrated foundation year have a potentially different student population, typically holding non-standard entry qualifications, there would be value in looking at the performance of these courses separately from first degrees. Respondents said there is data to suggest that students on an integrated foundation year have historically been less likely to continue, and that not

separating out this provision may incentivise providers to remove it. It is viewed as an essential route in some subject areas, such as engineering, and was seen to provide an important role in widening participation.

Pathway courses

136. One respondent considered that other pathway courses offering preparation for higher education courses at Levels 3 to 7 should be considered and assessed separately, rather than compared with courses that have different intended outcomes, to ensure assessment of outcomes is adequate.

Courses with a sandwich year

137. A small number of respondents said that sandwich year courses should be considered as an additional indicator, because the student experience is very different for this type of course and the sandwich element is likely to impact on students' employability and therefore outcomes. The additional level of detail would provide useful information to providers considering changes to current delivery mechanisms.

Distance learning

138. A small number of respondents said that distance learning should be considered as an additional indicator. This was because distance learning students may be more likely to be studying for personal interest or academic stimulation, rather than to achieve particular employment outcomes. Separating this data would provide information about the different outcomes for students compared with their counterparts on campus, and would be useful information to providers considering changes to current delivery mechanisms.

139. One respondent suggested that blended learning is likely to grow after the pandemic and this may also be another indicator that should be considered.

Accelerated degrees

140. A small number of respondents said that accelerated degrees should be considered as an additional indicator. This was seen to be potentially useful information to providers considering changes to delivery mechanisms, and would ensure parity of treatment for those offering alternative provision.

Top-up degrees

141. A small number of respondents said that top-up degrees should be considered as an additional indicator because they have distinctive characteristics. Having separate data would enable further analysis and actions to be focused on these distinctive groups and fully capture the different routes into particular programmes, such as nursing. Respondents suggested that top-up degrees generally recruit students who have a genuine desire to continue their studies, and therefore continuation and completion rates are likely to be higher.

Separate indicators for Level 4 and 5 courses

142. A small number of respondents suggested that 'other undergraduate' should be split further into Level 4 and Level 5 indicators. This was because 'other undergraduate' draws together a broad range of courses and qualifications that may not be comparable. Combining the levels may also deter innovation in sub-degree qualifications for fear of a negative impact on measurable indicators. Respondents also suggested that splitting the indicators by level would

make the data easier to understand, and that differentiation between the levels is important for progression indicators. This is because students who complete a Level 4 course expect a different outcome from those completing a Level 5 course, and many have no intention of progressing to further levels of study.

Other suggestions

143. A small number of respondents suggested other indicators for mode and level of study, which are outlined below with reasons where they were given:
- a. Modular, flexible or lifelong learning, because this type of provision is increasingly used to meet upskilling needs or to allow individuals to dip in and out of learning experiences in line with their personal needs and desires. Not having a separate indicator may discourage providers from delivering this, for fear of impact on overall measures of outcomes.
 - b. Professional postgraduate courses, such as short courses with teaching included and courses to prepare students for professional examinations.
 - c. Extended degrees.
 - d. Any mode of study that gives students experience of employment while studying, beyond apprenticeships and sandwich years, as this would provide valuable insight into any benefit of this experience.
 - e. A specific separate indicator for integrated masters' degrees.
144. There was general support for including separate indicators for different types of postgraduate courses as set out in the proposal, as this was seen to strengthen the focus on these courses. Some respondents suggested that as postgraduate research funding is the responsibility of UK Research and Innovation, this might lead to duplication of monitoring. Other respondents suggested that a further separation could be made for postgraduate indicators, for example between Master of Arts and Master of Fine Arts.

General points

145. Other points that were made by respondents in connection to the proposed indicators for mode and level of study included the following:
- a. A single baseline for each level of study may not achieve transparency and public confidence, given the number of factors for which it must account.
 - b. Having a different baseline for part-time students may not be in line with setting minimum expected outcomes for all students.
 - c. The differentiation between 'first degree' and 'undergraduate with postgraduate elements' may not have much meaning to students or applicants, as it may just be seen as progression. Therefore, there may be less value to having these separate indicators.
 - d. A small number of respondents felt strongly that no further indicators for mode and level should be considered in addition to those set out in the proposal, because the existing levels were sufficient. It was also suggested that students registered on courses with

other types of delivery, such as apprenticeships and accelerated degrees, are recruited with the appropriate qualifications to undertake the course at its entry point, so should have the same expectations regarding outcomes.

Our response

146. We note the broad agreement with constructing indicators separately for full- and part-time students and for different levels of study and, for the reasons set out in the phase one consultation, are continuing to propose that indicators will be constructed and assessed on this basis.
147. We agree that apprenticeship provision is sufficiently different from other types of courses, and recognise the fact that apprentices have employment contracts and 80 per cent of their training is delivered on the job by their employer. Our phase three and data consultations therefore propose that we will treat apprenticeships as a distinct 'mode' of study, separated by whether the apprenticeship is undergraduate or postgraduate in level.
148. In relation to integrated foundation years, we recognise that there is evidence that such courses have historically produced lower continuation rates and that this may be a reflection of the entry qualifications of students studying on these courses. Because these courses allow entry for (often mature) students without the formal qualifications necessary for entry to higher education, they can be an important mechanism for widening participation. However, we are also mindful of the point also made by some respondents that students register on courses with an integrated foundation year in order to progress to a first degree, and therefore sign up to the same learning outcome as any other student registering on a first degree course. Students on these courses also pay the same fee for the foundation year as for the first degree, and these courses therefore represent a more expensive route into higher education than many alternatives (for example access courses).
149. We have considered comments that suggest courses with an integrated foundation year should be considered separately from first degree courses because of the characteristics of the students who are generally recruited to these courses. We are mindful that there was strong agreement in responses to the phase one consultation that we should not set separate baselines based on student characteristics. We are therefore proposing in our phase three consultation that full-time first degree courses with an integrated foundation year should be separated in the data as a split indicator relating to different course type, but not a separate level of study with a separate numerical threshold. This is because the learning aim of the course onto which the student registers is the same as a first degree. However, we will separate out foundation years as a split indicator, which will allow us to see whether there are differences in a provider's performance between different course types without providing an incentive to accept weaker performance. Further detail is set out in our phase three consultation. It also sets out proposals to use a benchmark value as part of the process, to check the influence that factors such as student characteristics and entry qualifications may be having on a provider's performance.
150. In relation to creating separate indicators for courses delivered at Levels 4 and 5, rather than combining them in an 'other undergraduate' category, we are also proposing in the phase three consultation that we will show these as split indicators reporting on separate course types, but they will not have separate numerical thresholds. This is because constructing

separate indicators for each level would result in large numbers of sparsely populated indicators.

151. We have considered the suggestions made by respondents about other types of courses that they thought should be separated out within the indicators. We set out in our phase one consultation that any further splits of data would be subject to data availability and reliability. For this reason we are not able, at this time, to separate out accelerated degrees because these courses were only identified distinctly in data reporting in 2019-20 and only for those providers returning data to the Higher Education Statistics Agency (HESA) rather than through the Individualised Learner Record. We are also not clear that structural features of distance learning will persist in a way that might warrant separation in our measures as concepts of distance and blended learning become more blurred. We will consider whether there is merit to separation in the future, and note that this may be an aspect of our consideration of provider context when making assessments.
152. Other suggestions, such as pathway courses, top-up degrees, courses that offer employment opportunities, professional postgraduate courses, short courses with teaching included, and courses to prepare students for professional examinations, are not possible to distinguish in the data because they are returned by level of course. In relation to pathway courses, we propose that the biggest impact will be for registered providers that almost exclusively deliver these courses and that we will therefore be able to take the context of the provider into account when interpreting performance in the data.
153. Separating out Masters of Fine Art from MAs would rely on using course titles reported in data as free text, so data and therefore the correct identification of courses would be imperfect and potentially unreliable. We are therefore not looking to split these courses out, but looking at subject indicators is likely to have the same effect.
154. Some of the other suggestions are already covered by the levels set out in our phase one consultation; for example, extended degrees and integrated masters' are included in the 'undergraduate with postgraduate components' split.
155. We have considered the suggestion about splitting courses at Levels 4 and 5 rather than combining them in a single 'other undergraduate' indicator. We are therefore proposing in our phase three consultation that we do not separate out these levels with separate numerical thresholds. However, we will separate out these levels as a split indicator, which will allow us to see whether there are differences in a provider's performance between different course types and without providing an incentive to accept weaker performance. Further detail is set out in our phase three consultation.
156. In relation to separating out modular, flexible or lifelong learning, we agree that this would be helpful because of the structural features of these courses and the likelihood of different definitions of success compared with more traditional courses, but it is currently not possible to identify such courses in the data. As set out in the phase one consultation, this is an area we intend to develop as the government policy for the Lifelong Loan Entitlement develops, and in line with Data Futures. In the meantime, if a provider has a significant amount of modular or flexible courses that affect the performance in its indicators this can be taken into account as context, but this would need to be supported by evidence that these outcomes

were positive for students signed up to appropriate learning aims, and in line with their expectations.

157. Further detail about how each indicator will be constructed is included in our data consultation.

Student characteristics

158. The consultation proposed that the continuation, completion and progression indicators would be split to show the performance of different student demographic groups within each level and mode of study. This would continue to show performance for a given demographic group in aggregate, over a time series of up to five years for which data exists for the provider, but would not provide information for that group separately for each year of the time series. To date, these 'split indicators' have shown the performance within each indicator, broken down as follows:

- age
- Participation of Local Areas (POLAR) quintile
- English Indices of Multiple Deprivation (IMD) quintile
- ethnicity
- disability
- sex
- domicile.

159. We did not propose to split each subject or partnership indicator by demographic group, because our view was this would create a substantial number of data points that would be too small to form the basis of meaningful judgements.

160. We welcomed views on whether there were further characteristics, such as other protected characteristics listed in the Equality Act 2010, which we should consider including in the list of 'split indicators', noting that the construction of any additional splits based on student characteristics would be dependent on data availability, quality and reliability.

161. A very small number of respondents disagreed with the use of split indicators based on student characteristics, stating that if the policy aim is that all students, regardless of background, were expected to achieve the same numerical baseline it should be unnecessary to look at demographic data. It was also suggested by some respondents that the performance of demographic groups should be monitored as part of APPs rather than as part of an assessment of quality.

162. A number of respondents commented that the demographic splits as proposed could lead to very small numbers of students and therefore unreportable outcomes. The issues relating to the use of data based on small population sizes and their statistical reliability were noted by respondents.

163. However, the majority of respondents agreed with the proposed split indicators because these were the categories for which data was routinely gathered and are in alignment with split indicators use in APPs and previously for TEF. Most respondents suggested that gathering data for further indicators would create additional burden, and this was seen as undesirable.

164. A small number of respondents provided specific comments on some of the proposed split indicators.

Age

165. Only a few comments were made in relation to use of age as a split indicator. One suggestion was that a small number of age bands should be used, rather than under and over 21, as this would recognise providers whose students are typically older than those who traditionally study undergraduate courses.

POLAR and IMD

166. Some respondents were critical of the use of POLAR as a characteristic, because it was said to not capture broad enough measures of disadvantage. The most widespread criticism was of its use in relation to providers that are in, or recruit a large number of students from, London. Similar remarks were made by respondents in relation to major cities and densely populated areas generally. In these circumstances, the higher density of population in London, the relative absence of low-participation areas in London and the closer proximity of people of lower and higher wealth make this characteristic unreliable as a measure. Respondents noted that these weaknesses could lead to misleading or unmeaningful insights, misleading assumptions about student characteristics and a lack of statistical robustness.

167. Respondents also noted that POLAR data has less or no relevance to postgraduate and international students.

168. Some respondents suggested that the use of POLAR was unnecessary if the OfS was also looking at IMD data, as this would provide a more useful indication of students' social disadvantage compared to POLAR. Others suggested that tracking underrepresentation by area (TUNDRA) and UCAS multiple equality data could be used in addition to POLAR, to reduce the issues with the POLAR data.

169. A couple of respondents requested that lay definitions be included in the presentation of analytical work in relation to POLAR and IMD, because they viewed these measures as not as well understood as other split indicators.

Ethnicity

170. Respondents who commented on this split indicator agreed with its use, but expressed a preference for it to be made more granular. A couple of respondents suggested that a split between minority ethnic and white students was of limited analytical use, and expressed a preference for a mid-level granular split (such as Asian, black, mixed, other and white) where this was possible. However, another respondent specifically commented that these more granular splits would still be very broad, and would risk concealing a wide variation of performance between different groups who would fall under one category, for example between Indian, Chinese and Bangladeshi students in the Asian split. A couple of respondents also suggested that the OfS should consider a split indicator for students from Gypsy, Roma or Traveller communities.

Disability

171. Only a few comments were made in relation to the use of disability as a split indicator. As with age and ethnicity, there was a suggestion that it could be further broken down for example to show the type of disability, such as mental health.

Sex

172. Some respondents advocated that sex should be replaced with a split indicator for gender, and others suggested that a gender or gender identity measure could complement the sex split indicator. Respondents asserted that this would be appropriate as sex no longer followed a binary definition. Some of these respondents cautioned against a binary male-female split for either sex or gender split indicators and, for example, advocated the inclusion of neutral or non-binary options. One respondent noted that Data Futures would need to be compatible with any changes made in this regard.

Suggestions for additional or alternative split indicators

173. As noted in paragraph 163, the majority of respondents did not favour additional split indicators. However, there were a small number of suggestions for additional indicators.

174. Some respondents expressed concern that the proposed split indicators covered some but not all protected characteristics under the Equality Act 2010. Religion and sexual orientation were specifically cited by respondents in this regard. In relation to religion, respondents noted that it was compulsory for higher education providers to return data on this to HESA, so appropriate data should already be available.

175. In relation to sexual orientation respondents recognised that there was a lack of data and that little was known about the experience of students who are lesbian, gay, bisexual, transgender or of other minority sexualities or sexual identities (LGBT+). It was suggested that collecting such data and including it as a split indicator would increase providers' knowledge of the diversity of their students.

176. Respondents commented that when sufficient data was available for the omitted protected characteristics, they should be included; otherwise, it would not be possible to identify poor outcomes that needed to be addressed among certain subsets of students.

177. Respondents suggested several other characteristics that the OfS should consider adopting as further split indicators. These included students with caring responsibilities; care leavers; family estrangement; those from military or armed forces families; and first in family with higher education experience. It was suggested that these categories of students often have distinct support needs and including these as split indicators would enhance and support access and participation plans.

178. Eligibility for free school meals and English as an additional language were also suggestions, with respondents commenting that these would complement existing split indicators such as IMD and POLAR.

179. Other characteristics suggested for split indicators included asylum seekers, prior attainment, and commuter students.

Intersectional disadvantage

180. A few respondents suggested that split indicators should be intersectional and based on multiple, rather than single, characteristics. It was suggested that data was already available for many providers that would facilitate an intersectional analysis. However, several other respondents acknowledged difficulties in creating intersectional split indicators, due to the small student numbers, the significant underrepresentation of some groups, a potential lack of statistical robustness and the risk of individual students being identifiable from the data.

Our response

181. We have considered the responses that argued that we should not seek to split indicators by student characteristic if our policy aim is to ensure that students from all backgrounds achieve the same outcome, and that this comparative data is something that is considered and regulated through APPs. We have also considered responses that argue that our approach makes it easier for high-tariff providers to meet baselines. However, looking at performance of student characteristics this way will allow us to identify whether a high-tariff provider is delivering positive outcomes for students from all backgrounds. While APPs currently require a provider to address gaps in performance, for example improving the continuation rates between different student groups, they do not set out a minimum level of performance that should be achieved for all groups of students. That mechanism is achieved through condition B3. In relation to comments about the issues with POLAR data, we will not be updating this classification in future and are therefore not proposing to continue with its use as a split indicator. Details of the proposals and reasoning are set out in our data consultation in Proposal 2.

182. We have considered the comments in relation to a more granular breakdown of data in relation to age, ethnicity and disability. We are proposing to disaggregate both age and ethnicity, but not disability because this would result in student populations within each split which would be too small to allow a meaningful assessment. A full explanation is provided in our data consultation.

183. We have decided not to include other protected characteristics because we do not currently have sufficient data for all characteristics for all providers to be able to do this. We will continue to use sex within our splits, because this is a protected characteristic and creating data based on this is therefore important for fulfilling our public sector equality duties.

184. In relation to caring responsibilities, care leavers, estrangement, military families and first in family with higher education experience, we agree that understanding the performance of these groups is beneficial, but creating split indicators would be likely to result in sparsely populated data points. We are therefore not proposing to use these categories to construct additional split indicators

185. Our indicators consultation proposes other splits for the data, including free school meals, associations between characteristics of students (ABCS) and GO quintiles, that we are able to construct from the data available to us and are relevant to understanding the performance of providers in relation to different groups of students.

Small cohort sizes

186. We proposed that indicators would be constructed for any size of student population but we recognised that data can become unreliable where the population is too small. We therefore

also proposed that regardless of the population size, our presentation of the indicators would describe the statistical uncertainty associated with each indicator value. We set out in our phase one consultation that the size of student population that should be included in an assessment of the indicators, and the methods used to communicate statistical uncertainty, would be tested through further consultation.

187. A large number of respondents suggested that small providers with small student cohorts face additional difficulties when compiling meaningful and robust student outcome data, as data for the TEF and National Student Survey (NSS) already demonstrates. Respondents pointed out that data for small providers is often subject to fluctuations as the actions of one or two students has a disproportionate impact on the provider's outcomes data, which they considered would not reflect the quality of the course.
188. Some respondents suggested it is not possible to make any meaningful assessment of the quality of a provider's courses by judging the outcomes data of a provider with thousands of students against that of a provider with a few hundred. Respondents suggested that baselines should be scaled to account for the size of the student cohort.
189. Respondents also suggested that small providers and those with small student cohorts often find they must work harder processing outcomes data to account for fluctuations and variances, and this puts a disproportionate regulatory burden on the providers, which larger providers may not face. Many providers have limited staff resource available in general, and this is not necessarily a sign of poor-quality courses.
190. Outcomes data for small providers tends to be aggregated to aid data analysis and can hide both very good and very poor performance, leading to an unclear picture of the provider's true performance. This risks unnecessary regulatory intervention and ensures that small providers will always be outliers in any system using absolute baselines.
191. It was noted that introducing indicators at a more granular level for each mode, level and subject of study may result in small numbers of students, and therefore less comparable and credible data.

Our response

192. As set out in the phase one consultation, we recognise that there are issues with statistical reliability when considering smaller student populations. Respondents did not make suggestions about how we might approach the consideration of small population sizes, although one option would be to exclude from our assessment student outcomes for any provider or student cohort under a particular size. Our approach to reporting indicators and split indicators when the population size is small enough to represent a risk of data disclosure about individuals (in breach of the General Data Protection Regulation) is set out in our data consultation and in the proposed text of the new condition B3.
193. However, as a principle it is important that we are still able to regulate providers where the student population is small, because it is in the interests of those students that we do so. We have therefore set out proposals in our phase three consultation that explain how we will use a measure of statistical confidence when considering a provider's performance in an indicator. This will allow us to identify, even in small cohorts, when we should intervene to address providers' performance that is below a numerical threshold and where context does not

adequately account for that performance. We consider that this approach also means that numerical thresholds do not need to be weighted or scaled according to student population size.

194. We recognise that smaller providers have more limited staff resource available, and will provide the relevant datasets to providers on an annual basis.

Indicators for subjects and identifying pockets of performance

195. The consultation proposed that the OfS will assess student outcomes at a more granular level, by considering the performance of a provider at subject level. We proposed to show performance for each subject within each mode and level of study, to allow us to identify pockets of poor performance.

196. The consultation proposed that indicators should be established for subjects at level 2 in the Common Aggregation Hierarchy (CAH) for each of the continuation, completion and progression indicators. We did not propose to set a different numerical baseline for each subject within each mode or level of study – the performance in each subject would be assessed against the baseline established for each mode and level of study.

197. There was some support for establishing indicators at a subject level, with respondents commenting that it was important to have consistency in performance between subjects, and that subject-level assessment would help to identify poor performance that would otherwise be hidden by being part of larger provision. Respondents who supported this proposal recognised that this would generate a substantial amount of information, and that the data would potentially require increased effort to interpret, but felt that this was a better option than not considering all factors affecting performance. Some respondents were broadly in favour of the proposal, as long as contextual information was factored into any assessment.

198. A large number of respondents suggested it was inappropriate to set the same baseline of performance for all subjects, because of the large variation of performance between subjects, with this view expressed most strongly in relation to the progression indicator. Examples given for differences in subject performance were medicine and dentistry compared with social sciences subjects. Respondents suggested that a single baseline for all subjects, set at a high level, might see some subjects across almost all providers fail to meet this level. Respondents also suggested that establishing the same numerical baseline for each subject might result in providers closing courses that were not performing well, rather than taking sufficient steps to improve provision, giving students less choice. Respondents therefore suggested that individual subject level baselines would be more appropriate.

199. While a significant number of respondents suggested there should be subject-level assessment and baselines, other respondents were concerned about the increase in regulatory burden this could create, particularly for small providers including further education colleges, because of the additional complexity in having more data and more baselines. Respondents pointed to the Pearce Review of the TEF, and to the burden for providers and the regulator of subject-level assessment. It was suggested that consideration of subject-level performance would lead to significant staff time being used to analyse metrics to spot areas of risk to providers, adding to regulatory burden. The publication of such data would also require

significant staff time and energy to interpret, contextualise and explain, so that the public has an accurate and fair view of providers.

200. Some respondents suggested that assessment of student outcomes should focus on assessment at provider level in the first instance, deferring subject-level considerations to a later date, to enable a more focused and manageable agenda for reform of regulation. Other respondents suggested that subject-level assessment could be removed for providers that were consistently above the baseline.
201. Other suggestions to mitigate the potential additional burden of looking at data at a more granular level were that the OfS should only look at pockets of performance for limited numbers of providers on a risk-based approach, or else define a zone just above the numerical baseline where the data might indicate that some provision could be below the baseline, and use this as the basis for further investigation and drilling down into data.
202. A number of respondents commented that constructing indicators at subject level would result in small sample sizes, and this would impact on the statistical robustness of the data and its ability to be interpreted. This was highlighted as a particular concern for small providers. It was suggested that there should be a minimum threshold for sample sizes for data validity, as some subject areas are relatively small, and there should be a threshold for inclusion according to the proportion of that provision in the provider's subject portfolio.
203. Some respondents suggested that subject-level granularity may mean that issues of genuine concern, relating to quality and standards at the institutional level, are overlooked because of the preoccupation with dense split subject-level data.
204. Some respondents also suggested that the consideration of interdisciplinary subjects would be difficult within subject-level indicators.
205. A few respondents commented that using the CAH might be difficult where this does not map neatly to a provider's own subject definitions. It was also noted that the change of reporting coding systems from the Joint Academic Coding System to the Higher Education Classification of Subjects may have consequences for the continuity of data from year to year.
206. It was suggested the OfS should also consider whether assessing at a more granular level represents value for money, and whether the benefits of detecting pockets of performance at subject level would make the added complexity and cost of the approach worthwhile. It was suggested that the OfS may wish to use the approach selectively where the size and type of courses would make it practical and informative.

Our response

207. We have considered comments suggesting it would be inappropriate to set the same numerical threshold for all subjects because of the variability in performance. We recognise that there may be differences in subjects, particularly for progression, but setting a single numerical threshold would take account of the performance across all subjects. Therefore, to use the example above, it would take account of medicine and dentistry and social sciences. We would not be performing a comparative assessment between subjects, but comparing all subjects against a threshold we would expect all subjects to be able to achieve. However, by breaking indicators down to show subject-level performance we would be better able to see

the context of the provider – for example, whether it only delivered social science courses – and can take this into account through our assessment. We are also mindful of the large number of responses that raised concerns about the complexity of the data and regulatory burden. We have looked carefully at the data as we have developed our detailed proposals and, on balance, our phase three consultation sets out proposals for the consideration of indicators at a subject level without setting individual subject numerical thresholds.

208. Our view remains that it is important for us to identify smaller pockets of weak performance within larger student populations. Our phase three consultation has set out proposals for how we would prioritise assessment each year. This exercise recognises that we do not have unlimited resources available to assess performance and will therefore prioritise accordingly, enabling us to adopt a risk-based approach and use our resources efficiently. Using a prioritisation exercise will also enable us to use an increasingly risk-based approach, the practical consequence of which is that providers that are consistently above the numerical thresholds are unlikely to fall within the prioritisation exercise.
209. In relation to comments that consideration of subject-level indicators might lead to the closure of some courses, we accept this is the case. As set out in paragraph 38, our intervention is generally likely to be escalatory, and providers will normally be given the opportunity to improve over an appropriate timescale. If they are unable or unwilling to do so, removing courses that do not or cannot meet minimum requirements will be a positive thing.
210. In relation to the small cohort sizes that subject indicators might produce, we have explained in the phase three consultation how we propose to address this and the proposed condition sets out our proposal for a minimum population size. We propose that very small populations will be excluded entirely from our assessment, with the practical consequence of excluding subject-level assessment for very small providers. Where population sizes are too small to consider the performance in subject-level indicators we will only consider performance at level of study.
211. In our phase three consultation we have set out how we intend to approach assessment so that we can consider provider-level performance as well as identifying pockets of performance.
212. We acknowledge that producing indicators for providers delivering interdisciplinary courses may not be straightforward, and that CAH may not align with the subject definitions used by providers. Our data consultation provides further detail on how interdisciplinary subjects will be treated, and seeks further views on the approach the OfS should take in relation to these subjects.
213. In relation to the points made about regulatory burden for staff in interpreting data and spotting areas of concern, our view is that providers should be monitoring student outcomes and identifying areas of risk as a normal part of their academic governance arrangements. Providers that already have appropriate academic governance arrangements should not, therefore, see additional burden. We will be providing an annual dataset to each registered provider, which is constructed from the data it provides to either HESA or the Education and Skills Funding Agency. This will show the provider its performance in relation to the numerical thresholds set by the OfS, as well as information about the statistical confidence associated

with its indicators. We anticipate that this will support providers of all sizes to understand which of their courses may be subject to regulatory action.

Partnerships

214. The consultation proposed that the OfS would construct indicators to be assessed in relation to condition B3 in a way that would reflect a 'universal' view of a provider's obligations for the quality and standards in all its courses. This would mean that the indicators we produced for a provider would include all the students it taught, and the students it registered and for whom it was the awarding body. Students registered at one provider and taught by another would contribute to the indicators for both registered providers.
215. The consultation recognised that there are a number of ways in which indicators could be constructed and the disaggregation of data presented, and that, while there is a benefit to identifying differences in performance at a more granular level, this could lead to a substantial volume of information that could be challenging to interpret.
216. Several respondents commented that it was appropriate to construct indicators to reflect all a provider's students because they recognised that both providers in sub-contractual arrangements have responsibility for the quality of courses, and that constructing indicators in this way would help the OfS to identify and intervene where courses delivered through partnerships did not meet minimum required quality.
217. Other respondents commented that constructing indicators in this way would increase regulatory burden on providers, if both partners had to make data returns in relation to the same students. Respondents suggested that this increased burden might be felt in particular by further education colleges. The burden of collecting data in relation to students studying at unregistered providers was also mentioned as a burden for validating bodies and one that might threaten the sustainability of some partnerships.
218. Some respondents did not agree with the inclusion of validated courses within indicators because they stated that the validating body's role is to provide oversight but the content and delivery of courses are the responsibility of the delivery provider.
219. A small number of respondents suggested that constructing and assessing data through a 'universal' view might disincentivise validating bodies from entering or continuing with partnership arrangements, because of the potential impact on their performance against numerical baselines of aggregating their data with that of their delivering partners. This could in turn restrict entry to the higher education market, with implications for access to higher education.
220. One respondent suggested that set numerical baselines may not be appropriate for some forms of partnership arrangement, particularly where students receive joint or double awards, because the measures may be inadequate or skewed, and that the OfS assessment of such courses may need more nuance.
221. One respondent also pointed out that interpreting student outcomes indicators that are based on lagged data may be complicated where partnerships have changed.

222. Several respondents requested further information about how the OfS may intervene in relation to courses delivered in partnership arrangements where these do not meet the minimum numerical baselines, and whether and how the OfS would engage with providers involved in the partnership provision that are not registered with the OfS.

Our response

223. In relation to comments about the increased burden in data returns, it is not the intention that providers in a sub-contractual arrangement would both return data relating to the same students. As is the case now, the provider that has the contractual responsibility for the student and that registers the student will be responsible for returning the data relating to that student.

224. We will require additional information about the partnerships that providers have, and about changes to these. This information is also required to enable us to fulfil our statutory duty to publish information on the Register, and by statutory order a list of providers that meet the definition of 'Listed Body' in the Education Reform Act 1988. We propose to do this by conducting a one-off data collection exercise to collect core information about partnerships arrangements, following which providers would, as now, be asked to update information when partnerships change. We expect to consult on our approach to the collection of this information and anticipate proposing that registered providers that validate courses at providers that are not registered with the OfS will be required to collect and return that data from their partners.

225. In relation to points that the validating body's role relates to 'oversight' of the partnership rather than students' outcomes, and that constructing and assessing data through a 'universal' view might disincentivise validating bodies from entering or continuing with partnership arrangements, we consider that providers that hold degree awarding powers are responsible for the secure exercise of those powers and should ensure that the courses they validate are of appropriate quality and standards.

226. Our interest in looking at the outcomes achieved through courses delivered by validated partners is to ensure that a provider is using its awarding powers appropriately, and that students on courses delivered through those partnerships are receiving positive outcomes and are protected from unacceptably weak outcomes on the same basis as other students. If the outcomes are below our regulatory requirements, and the delivery partners are also registered, then the requirement to improve outcomes would rest with both providers. The delivery provider might have a specific condition imposed requiring it to improve outcomes for the students it teaches, and the validating provider might have a specific condition imposed to ensure that outcomes delivered by its partner providers are improved. We consider that this is consistent with the oversight role described by respondents, and that both providers would have a role to play in ensuring performance improved.

227. We might also want to engage with the validating body about the operation of its degree awarding powers. Where the outcomes are delivered by providers that are not registered, we will not engage directly with those providers, and we would expect to engage with the validating provider. Further information about our approach to assessment, including the consideration of validated provision, is set out in our phase three consultation.

228. We consider that providers' ability to operate their awarding powers securely, and the protection of students, must take precedence over concerns about the impact on partnership arrangements. Partnerships should not operate where these result in poor quality, including outcomes, for students. However, we recognise that there may be a risk to high quality partnership arrangements, due to perceptions of increased burden for the validating body. We are therefore proposing in our phase three consultation that we will not prioritise courses delivered under partnership arrangements for regulatory investigation in the first year of the implementation of condition B3. Further details are set out in the phase three consultation.
229. In relation to joint awards, our view is that both providers would be responsible for the outcomes because there is a joint responsibility for the students.
230. We recognise that data is lagged and that changes in partnerships can occur frequently. However, we would not expect partnerships to change within the delivery to a particular cohort of students, and therefore the data and our assessment should be able to cope with changes in arrangements.

Indicators for transnational higher education courses

231. The consultation proposed that we would construct indicators for continuation and completion outcomes in broad terms, based on information returned via the aggregate offshore record from 2019-20 onwards. We proposed that these indicators should be constructed at an aggregate level and included in the assessment of condition B3. This would allow us to regulate quality for the relevant students, and to ensure that outcomes are above a minimum level of performance.
232. The use of indicators for transnational education (TNE) courses was generally welcomed and there was broad recognition that some form of TNE outcome indicator should be required. However, there were a significant number of comments about difficulties with data collection and verification, and as a result whether data could be considered sufficiently reliable. Respondents suggested that providers that deliver TNE courses could find themselves at a regulatory disadvantage if TNE data is used as an indicator for quality, given that there are significant and longstanding issues with TNE data collection and verification which have not yet been addressed.
233. Some respondents suggested that English providers may not have the influence or leverage to compel overseas partner institutions and employers to provide the necessary data to the required standard. This would particularly affect the construction of progression indicators in relation to TNE. Some respondents suggested that further information was required about data collection, and there were concerns about the potential increase in regulatory burden.
234. Several respondents commented that the context in which the TNE courses are delivered should be accounted for when considering these indicators, and specifically that it would be necessary the local labour market in the country when considering employment outcomes.
235. Some respondents suggested that it is not appropriate to compare UK-based student outcome data with TNE student outcome data, as the two will vary so significantly that no meaningful conclusions could be drawn from the results. Other respondents suggested that TNE data should not be aggregated with other validated provision, because this would risk masking poor performance or concerning behaviour within a partner institution or country.

236. Other respondents suggested that TNE student data should not be part of the quality indicators at all, because, while it is important that students studying for an English award at an overseas partner should receive the same quality of provision as English-based students, the culture and context the partner provider operates in may mean the student experiences is overall very different.

237. It was suggested by one respondent that a universal view that incorporated TNE courses overreached the OfS's authority and overlooked the fact that overseas partners delivering courses in partnership with an English provider are subject to their own national quality requirements. Intervention by the OfS could damage the reputation of UK higher education and undermine TNE partnership arrangements. Another point made was that numerical baselines that are appropriate for courses based in England may not be suitable for TNE because of the effect on outcomes of socioeconomic factors specific to the geographical context in which the courses are delivered.

Our response

238. We have considered comments about the challenges of data collection and verification. We maintain that constructing indicators in relation to transnational education is necessary to allow us to regulate quality for its students, and to ensure that positive outcomes are delivered for students, taking account the context in which they are delivered. However, we accept that there are still a number of unresolved issues with data collection, which mean that we do not think it appropriate for us to propose to set numerical thresholds and to assess performance in these indicators on the same timescales as for courses delivered in England. Once data is sufficiently reliable, we will look to construct indicators and set numerical thresholds according to the same principles and methods set out in the phase three consultation, and would expect to consult separately about this.

239. We have therefore amended our proposal and will undertake and publish sector-level analysis of data based on 2019-20 and 2020-21 data from the aggregate offshore record in spring 2022. We will use this analysis as regulatory intelligence to inform our regulation of other quality conditions in relation to TNE courses, to inform our understanding of student outcomes in TNE courses and to inform priorities for improving data collection about TNE students. We will initiate conversations with providers about the data to inform the construction and use of provider-level indicators.

240. For clarification, our original proposal suggested that we would construct indicators for students studying outside the UK for continuation and completion only, and not for progression. We recognise the different context of student outcomes delivered outside the UK, and agree that it needs to be considered separately. In due course, we expect to consult on the detail of how continuation and completion indicators for students on TNE courses will be constructed, and about how data collections may change to facilitate this.

Considering data about higher education courses that are not eligible for funding by the OfS as context

241. The consultation proposed that we would include information about the overall numbers of students studying on higher education courses that are not eligible for funding by the OfS (also known as non-prescribed courses) in the contextual data included in a provider's B3

workbook, so that our assessment is informed by the broad scale of this provision for any individual provider. We also proposed that in the longer term we would wish to integrate courses that are not eligible for funding by the OfS into the coverage of our student outcomes indicators.

242. A small number of respondents commented on the nature of these courses, which are often less than one year in duration, and that this needed to be factored into the OfS's plans, and asked for further information about how baselines might apply to higher education courses that are not eligible for funding by the OfS.

Our response

243. Our proposal was to consider the overall numbers of students studying on courses that are not eligible for funding by the OfS and to consider this as context of a provider's higher education offer. We were not at this stage suggesting that we would construct indicators or set numerical thresholds for courses that are not eligible to be funded by the OfS because the data was incomplete.

244. However, the return of data on these courses is now a mandatory element of both HESA and Individualised Learner Record returns, and we are therefore proposing in the phase three consultation that students aiming for these qualifications will be included in the construction of the indicators and will therefore also contribute to setting numerical thresholds. This means that, for example, an indicator for the continuation of part-time other undergraduate courses would include students aiming for any Level 4 or 5 qualification, regardless of whether it is eligible to be funded by the OfS or not. Modules and micro-credentials will not be captured in the construction of the indicators at this stage, and the measurement of outcomes for those types of courses will be subject to further consultation. Further details are set out in our phase three consultation.

Additional indicators

245. In the phase one consultation we welcomed views on whether we should consider any other quantitative measures of student outcomes, for example an indicator that projects rates of students progressing from entry to professional employment or further study (such as Proceed data), or measures that might provide supplementary or additional contextual information about a provider's performance.

246. There was limited support for additional measures to be used when setting requirements in relation to student outcomes, other than a measure of value added. Respondents suggested that a measure that captured the value added by the provider was particularly important if benchmarking was not being used.

247. Other suggestions were that the OfS should develop a value for money indicator based on cost and other factors. There was also a suggestion that we should adopt numerical proxies for the quality of undergraduate taught provision, as set out by Graham Gibbs in the Advance HE report 'Dimensions of quality',¹² or combine indicator values to measure performance of higher education systems in terms of inputs, outputs and outcomes rather than treating them

¹² Available at <https://www.advance-he.ac.uk/knowledge-hub/dimensions-quality>.

individually, as described in the 2019 Organisation for Economic Cooperation and Development report 'Benchmarking higher education systems performance'.¹³

248. Other respondents were strongly not in favour of indicators based on combinations of indicators, and in particular projections of outcomes such as the Proceed data, because as a projected rate they did not think it was appropriate for regulatory purposes.

Our response

249. Previous work carried out by the Higher Education Funding Council for England and the OfS identified useful information that may support the future development of a measure of value added or learning gain, but we consider that there is not currently an appropriate measure for use across the sector. This is also recognised in our proposals for the new TEF, where we propose that there would be a focus on the educational gains providers deliver for their students.

250. In relation to the other suggestions, there are practical considerations of data availability or applicability across the sector associated with some of the other suggested indicators. While some of these measures are potentially interesting and valuable, having considered their likely additional value, we have concluded that it would not be proportionate to impose additional data requirements on providers or direct OfS resources to develop additional indicators at this point.

How baselines would be set for each indicator

251. The consultation proposed that in setting numerical baselines for performance, we would be informed by:

- a. The range of performance in the sector in relation to each indicator.
- b. Our policy intention to set a high bar for quality and protect students from poor outcomes, including by setting higher more challenging minimum numerical baselines that apply to each indicator and all providers.
- c. Our view of what a student, parent or member of the public and taxpayer might think was 'high quality' in relation to successful outcomes (we refer to this as the public interest test).
- d. Statutory guidance from the Secretary of State issued under section 2(3) of the Higher Education and Research Act 2017.
- e. Equality considerations.
- f. The OfS's general duties in section 2 of the Higher Education and Research Act 2017.

252. We proposed that the numerical baselines should continue to be set by reference to the performance of the sector as a whole, and that we should use a range of data to inform this.

¹³ See <https://www.oecd.org/education/benchmarking-higher-education-system-performance-be5514d7-en.htm>.

The phase one consultation sought views on how the range of sector-level performance should be taken into account when setting numerical baselines, and gave as an example setting the numerical baseline at a level that would seek to improve the bottom 10 to 20 per cent of provider performance in the sector for a given indicator.

253. Several respondents made comments in relation to how baselines should be set. They agreed that sector-level performance should be considered to ensure that baselines are relevant and achievable, including, for example, evidence of achievement by groups of students and students dependent on mode of study. It was suggested that the policy aim of improving outcomes for all students should be balanced with the acknowledgement of the reduced likelihood of successful outcomes for some groups of students. Respondents suggested that baselines should not be set at a level that might disincentivise providers from admitting students from underrepresented groups.
254. Respondents sought clarification about how sector performance would inform baseline setting, for example whether this would make use of average performance in the sector, separate figures for each institution, and regional adjustments, and the weighting to be given in the process to contextual factors. Respondents also suggested that baselines should be set largely based on evidence and analysis, for example of economic factors, including in the context of impact on graduate employment prospects following the pandemic.
255. Some respondents suggested that an intersectional analysis of student characteristics in the consideration of the data and when setting baselines would provide a greater insight into weaker outcomes and facilitate improvement of quality in these areas.
256. Some respondents supported the idea of setting baselines that would target the lowest-performing providers, taking account of sector-level performance for a given indicator, and that a baseline set with reference to the bottom 10 to 20 per cent of providers would be in keeping with a risk-based, low-burden approach. However, the vast majority of respondents did not like this approach and suggested that it would contradict the use of absolute baselines. This is because setting a baseline based on comparative performance, e.g. the bottom 10 to 20 per cent, denies the possibility that all providers at a given time could be performing very well in absolute terms, or the opposite situation where there is widespread weakness across the whole sector.
257. Alternatives suggested were considering outliers in performance or using the data characteristics from the whole sample of providers to calculate standardised scores for each provider (T or Z scores, for example), to enable accurate comparison across providers.
258. Other respondents suggested that individual baselines could be set for each provider based on a number of contextual factors that impact that provider, effectively giving it a bespoke baseline and target. Others favoured an assessment against a benchmark value rather than an absolute baseline, and many respondents commented on what they understood to be the removal of benchmark values from the current system.
259. Respondents also suggested that the OfS should seek to understand the potential impact of the baselines considered, to avoid adverse consequences; for example, providers with high proportions of students recruited from underrepresented groups, or students with certain protected characteristics, being placed in the lower performance range.

260. It was suggested that the OfS should consider international comparisons and the position of UK higher education in an international context when considering baselines.
261. Several respondents commented on the proposals to set baselines in relation to outcomes, considering what minimum level of performance would seem appropriate to members of the public, or a 'public interest test'. A few respondents were in favour of this proposal, while suggesting that providers would need time to achieve this, as with access and participation goals, and seeking more information about how the baseline would be set considering this information.
262. However, others were not in favour of a public interest consideration because they felt that members of the public are likely to lack relevant or sufficient knowledge to inform a view about where regulatory performance indicators should be set. Others felt that such a test would give too much latitude to the OfS in setting baselines and making its assessments.
263. Some respondents requested additional information or clarification about how baselines would be set and reviewed, whether they would change over time and the process for doing this.
264. Respondents wanted clarification about whether a lower baseline will be set for students studying part-time, and the reasoning if this is proposed. It was suggested that establishing different baselines for part-time courses would run contrary to the proposal to apply the same minimum baselines to all courses, although it was acknowledged that some groups of students, including part-time and mature students, may have specific requirements beyond those set within baselines.

Our response

265. We have carefully considered the comments that setting baselines should take account of context, in particular the historical variation in outcomes between different groups of students, to ensure that providers are not disincentivised from recruiting students from underrepresented groups. This did not feature in our phase one proposals because we were concerned that setting baselines in this way might result in double counting the effect of the impact of student characteristics, because we proposed to consider context in our assessment of individual providers.
266. We have set out in our phase three consultation an explanation of how we propose to use data on sector performance to set numerical thresholds, and how we propose to use analysis to inform where these are set in order to ensure that appropriate consideration is given to student characteristics. As part of this we have also explained our consideration of the impact of setting numerical thresholds on different student groups and type of provider.
267. We have noted comments that targeting a certain percentage of providers would result in a comparative rather than absolute approach, and agree with comments that it should be possible for all providers to achieve an absolute minimum level of performance. We have therefore not included this approach in our phase three proposals. For the same reasons we are not proposing an approach that would identify and focus on outliers of performance.
268. In relation to the suggestion that we should create standardised scores, we are proposing that we consider a provider's performance against an individual benchmark as part of the assessment process. This is explained further in our phase three consultation.

269. In relation to the use of benchmarks as individual targets, we have not included this in our proposals. Our aim is to set minimum requirements that all providers should be expected to deliver for students from all backgrounds. Using individual provider benchmarks as targets could mean that if some benchmarks were at a low level because performance across the sector was weak, this performance would be tolerated. This system would not allow the OfS to impose a mechanism to that protected students from weak outcomes that were in line with wider sector performance.
270. We note that several respondents commented on their understanding that we were proposing to remove the use of benchmarks from the current system. The regulatory framework is clear (paragraph 350) that we do not use benchmarked performance in an assessment of student outcomes. This has never been part of our baseline regulation, and the phase one proposals did not represent a change from this established position.
271. We have considered whether it would be appropriate to make comparisons with international performance when setting baselines, and are not proposing this within our method. This is because each country has its own policy, regulatory aims and funding regimes, and, while we recognise that England has higher outcomes compared with some other nations, this does not justify choosing not to set regulatory requirements that apply in England and should not influence what those requirements are.
272. We have noted concerns that the suggested 'public interest' test is too broad and that the public is not sufficiently well informed to make a judgement. Our intention with this proposal was that we would challenge ourselves to take account of what a reasonable person might consider to be acceptable performance. It is also not the primary factor in determining where a baseline should be set, but a test of its appropriateness. In our phase three proposals we have amended this element to be clear that the purpose is primarily to test the student and taxpayer interest in relation to value for money. We consider this is an important factor, but as in phase one we propose that it is not determinative of where a baseline will be set.
273. The phase three consultation also sets out that we do not propose that baselines would automatically rise over time, but they should be reviewed on a periodic basis to align with the TEF assessment exercise.

Baselines for different student characteristics

274. The phase one consultation proposed that the performance of students from underrepresented groups and with protected characteristics should be assessed in relation to the same numerical baseline established for each indicator. This will ensure that all students are protected from low quality, including weak outcomes, by ensuring a minimum level of performance.
275. The majority of respondents who commented on this agreed that it would not be appropriate to set different baselines for different groups of students.
276. However, a small number argued that numerical baselines should be varied by student characteristics. Respondents cited the need to prevent perverse incentives that would discourage recruitment of students from certain groups, and the need to recognise the differential performance levels among different demographic groups.

Our response

277. As set out in the phase one consultation, we have proceeded with proposals to set a numerical threshold that applied to students from all backgrounds. In relation to preventing unintended disincentives for providers, our phase three proposals set out how we will consider the context of student characteristics in the setting of numerical thresholds and in the assessment of performance. We consider that these measures mean it is unnecessary to set numerical thresholds based on student characteristics.

Publication of data about providers' performance

278. In the phase one consultation, we proposed to improve transparency in relation to the assessment of student outcomes by publishing the indicators for individual providers, to show their performance in relation to the numerical baselines.

279. There was broad support for the proposal to improve transparency in relation to the indicators, as it was considered that this would maintain the confidence of the sector and ensure the indicators can be easily understood by all stakeholders. Some respondents thought that the indicators should be sufficiently granular that specific student groups and different types of provision can be reported separately, for instance higher apprenticeships, courses with an integrated foundation year and partnership provision. This would be more transparent and better enable the identification of pockets of poor performance.

280. Some respondents commented that the more indicators there were, the more difficult it would be to establish transparent baselines, and it would therefore be more difficult for students and other interested parties to interpret a provider's data.

281. A few respondents commented that the publication of data could mislead students, the public and stakeholders, presenting an inaccurate view of a provider's performance, unless it was presented with contextual information. It was suggested that this could particularly impact providers that recruit a large proportion of students from underrepresented groups, and may undermine the ability of students to make informed decisions about providers. A number of respondents considered that benchmarking rather than absolute data would provide more transparency and be more helpful for stakeholders in interpreting performance.

282. Several respondents argued that the publication of data about providers whose position based on absolute performance was below the baseline could damage the reputation of providers, particularly those that recruit large numbers of students from underrepresented groups. This could also damage the employment prospects of students who study at those providers if they are perceived as 'low value'. As with other aspects of the proposals, there was a suggestion that publication of the data could lead to providers changing their admissions requirements to avoid reputational risk.

283. A few respondents advocated complete transparency about how the baselines are set, and the assessment of each provider's performance including the process for considering context, so that they can be fully understood, along with advance availability of the data to enable opportunities for challenge and scrutiny. It was suggested that transparent criteria for baselines and indicators should be accompanied by effective support and guidance for providers, as this would prevent breaches. In addition, respondents suggested that the OfS

should keep the methodology used under ongoing review, and this should be done in a transparent manner.

Our response

284. We recognise that the number of potential indicators and data that we proposed to publish would result in a lot of information being available and that this might make interpretation difficult. However, we do not think this overrides the public interest in the transparency about provider performance and our judgements about this. We have set out in the phase three consultation further proposals regarding the publication of information, and that each provider will be provided with its own B3 workbook to enable a better understanding of what the data will look like. We propose that there will be some commentary available, which explains the construction of the data that we publish and will help with interpretation.

285. We consider that publication of data is an important regulatory tool and acts as an incentive for providers to understand and improve their performance if necessary. We note that respondents to the phase one consultation were strongly in favour of transparency in regulatory decision making, and suggested that this is a principle of the Regulators' Code. We think the principle of transparency of information for students and other stakeholders should normally override concerns about reputational damage.

286. We have proposed in the phase three consultation that we will consider a provider's performance against a benchmark value, to help us interpret the extent to which its context is driving performance, and proposals for how we will publish information to help people interpret data. Where there is no context or other factors that adequately explain a provider's performance, this is likely to lead to a decision that there is a breach of condition B3 and regulatory intervention, including requiring a provider to improve. Subject to our general policy on publication, which is currently subject to consultation, it is our expectation that any information about a breach and the reasons for that breach would be published, to provide transparency about our regulatory decision making.

287. We have set out proposals for setting numerical thresholds and the assessment of condition B3 in the phase three consultation, and will publish our final decision on the approach we will take following consideration of responses to it.

288. In relation to providers changing their admissions requirements, we acknowledge that this is a possibility. However, we think it is unlikely that providers would be able to maintain their current student numbers if they did this. Therefore, unless they are prepared to reduce their numbers, the alternative is that they must act to improve outcomes for the particular students they recruit. We consider that if providers are recruiting students who are not able to succeed and benefit from higher education, then a change in recruitment behaviour will be a positive thing.

Assessment of condition B3 and consideration of context

289. The consultation proposed that assessment of compliance with condition B3 would rely heavily on a provider's absolute performance shown in the proposed indicators, although the OfS would also consider the provider's operating context and other relevant factors to ensure proper interpretation of its absolute performance.

290. We proposed to set numerical baselines for student outcomes and to assess a provider's performance in relation to these. In addition, we said we would consider a provider's context to ensure we have properly understood its absolute performance. Paragraph 67 of Annex B of the phase one consultation listed four categories of contextual factors we proposed to take into account when assessing performance:
- a. Evidence that a provider no longer delivers, or no longer plans to deliver, courses that are included in the indicators, and the reasons for ceasing the delivery of such courses.
 - b. Any external factors outside the provider's control that might affect performance in relation to the numerical baselines; for example, in the immediate future, any impact of the coronavirus pandemic on a provider's performance.
 - c. Evidence of any actions a provider has already taken to improve its performance in relation to numerical baselines, and the effectiveness of those actions.
 - d. Evidence of a provider's plans to improve its future performance in relation to numerical baselines, and the credibility of those plans.
291. The consultation asked respondents for their views about these factors and any others we should take into account, and what weight we should place on them.
292. There was significant support for the inclusion of contextual factors when assessing performance shown in outcomes data, and respondents stated this was necessary to ensure its correct assessment and interpretation. Respondents suggested that the OfS should contextualise data when it generates it, rather than seeking additional contextual information from individual providers where courses fall below the baseline, because additional information submitted by providers may not be available to the public.
293. A large number of respondents suggested the use of benchmarking as the most effective and transparent way of taking context into account within any assessment. Respondents expressed concern about how context could be taken into account fairly and consistently without the use of benchmarking.
294. The most commonly mentioned contextual factor that respondents thought the OfS should consider was student characteristics, and their potential impact on performance against an indicator, including between different levels of award (because some awards attract larger numbers of students from underrepresented groups). It was suggested that acknowledging this does not equate 'baking in' inequalities to the sector.
295. It was suggested that ignoring evidence of the ways student demographics impact on student outcomes in setting and assessing against baselines risks reduced access to higher education by students from underrepresented groups, because providers focused on widening participation, whose students' outcomes may be lower, will be disproportionately affected and may reduce their focus or exit the higher education sector. Some respondents took the view that this may result in indirect discrimination.
296. Many respondents were of the view that in any assessment the OfS should use an intersectional approach to considering the student characteristics, which recognises the multiple characteristics that students have and the potential impact this has on their

performance. An example was mature students , with a respondent indicating that it was established that outcomes for these students have been heavily influenced by intersections of inequality such as race and class.

297. Only a small number of respondents commented in relation to the specific proposed contextual factors, but those who did disagreed that we should consider evidence that a provider no longer delivers, or no longer plans to deliver, courses that are included in the indicators, and the reasons for ceasing the delivery of such courses. The objections were that this was an infringement of institutional autonomy, as respondents suggested it sought to influence the management of a provider's provision and that providers sometimes decide to continue delivering courses with weaker student outcomes because of the development and learning opportunities that they offer to students who would otherwise not be engaged in higher education.

298. There was a substantial number of responses on a range of other factors that respondents considered were beyond a provider's control and should therefore be taken into account. The main factors cited were:

- a. International, national, regional and local economic and political events.
- b. Changes in government education, research and funding policies.
- c. Structural and social inequalities, notably in the labour market.
- d. Students' personal circumstances, including life events that may occur while they are studying or shortly after achieving their qualifications, which have an impact on their ability to complete their studies and to seek employment.

299. There was also a large number of comments about the impact of the pandemic and how this should be considered. Respondents suggested that the effects of the pandemic would be felt for many years and may result in reduced employment opportunities for graduates. Others suggested that providers have been affected differently by the pandemic, depending on where they are located, the type of provision and subjects taught and the students they expect to recruit. It was suggested that it would be difficult to quantify the impact of the pandemic consistently in order to assess providers' performance fairly. In particular, respondents suggested that those providers that recruit more students from underrepresented groups will have felt the disproportionate impact of digital poverty on their students' performance.

300. In addition to the contextual factors proposed in the consultation, respondents suggested that the OfS should take into account further contextual factors as follows:

- a. A provider's mission and strategy, which would include its access and participation plan, the role it plays in its locality, other higher education provision in the area and its history and tradition.
- b. A provider's funding, including the amount of premium funding, and the length of time the provider had been eligible for the funding.
- c. The provider's entry tariff, as this was stated to be a key driving factor in high quality student outcomes.

- d. Whether a provider is operating in its New degree awarding powers phase.
- e. The views of a provider's students and graduates as expressed through the NSS and Graduate Voice data, although respondents also suggested that the OfS should canvass student opinion more widely to obtain qualitative input.
- f. Efforts taken by providers in relation to improving performance against numerical baselines.
- g. Provider trends and previous performance.
- h. The sustainability of performance, specifically in relation to providers that may lack resource or capability to sustain acceptable performance even if able to demonstrate it at the point of assessment.
- i. A provider's reputation among its peers, and comparative performance.

301. Some respondents suggested that contextual factors need to be applied at a granular level, taking into account specific courses and student cohorts, rather than just at provider level.

302. A few respondents suggested that the proposed approach to setting and assessing against numerical baselines should be used to guide judgement, rather than the process becoming mechanistic or automated. Respondents suggested the OfS should consider the risk threshold that would trigger enforcement, to ensure this is proportionate, as it was suggested that enforcement should be in response to a trend in performance, shown across several indicators, rather than to isolated instances of performance below single baselines.

303. Several respondents also suggested that where the data flags potential risks or adverse changes, this should initiate dialogue between the OfS and providers, including, for example, to identify 'pockets of performance', rather than triggering decisions. While there was some reservation in responses about the burden of responding to the OfS with additional information, there was a clear wish among several respondents to be given opportunities to contextualise information from numerical data. Others suggested that a benchmark value for a provider could be used as a trigger for intervention, while setting a high absolute value that would exclude providers with an unambiguously 'good' performance.

304. Some respondents were of the view that the use of contextual information appeared to be a marginal part of the proposals, rather than integral to them, because of the limited detail on how these will be incorporated into assessment and what precise factors will be considered. More information was sought about the use of contextual information in the next phase of consultation, such as about how contextual information including student demographics will form part of assessment, the weighting of contextual information against other factors, and the precise factors to be included.

305. Other questions were whether assessment would only consider numerical data, how the OfS would assess the progression to graduate employment of international students, and the balance between the three student outcome indicators in the assessment process. For example, would performance below baseline in one area, such as continuation, lead to regulatory action if the other two were clearly above the baseline, or would overall performance take precedence?

306. Clarification was also sought about whether providers would be considered to be in breach of the condition if their performance fell below baselines for some indicators, such as subject indicators, but were satisfactory at the level of study or provider.
307. Some respondents also asked how the OfS would manage the possibility that low percentages would disguise the number of students affected at providers with large numbers of students or inflate the perceived change against percentages for providers with small number of students.
308. Another suggestion was that providers should be assessed in comparator groups. It was suggested that comparing a large multi-faculty university to a small further education college would not provide a helpful evaluation, and that providers tend to compete with similar providers in micro-markets, as opposed to there being one higher education market. One suggested comparator factor was entry tariff.
309. One respondent suggested that outcomes in relation to quality for condition B3 ought to be assessed by the Quality Assurance Agency for Higher Education rather than by the OfS, because condition B3 is concerned with quality and the Quality Assurance Agency is the designated quality body.

Our response

310. We have considered the comments about how context should be taken into account, in particular student characteristics and the suggestion that an intersectional approach is needed to understand the impact that combinations of characteristics have. We have set out in our phase three consultation how we will use our Exploring Student Outcomes analysis, which shows the effect of combined characteristics, in setting numerical thresholds. We are also proposing an approach that would use provider benchmark values to help us interpret the context of the performance of a provider. We think this will address points about how the impact of student characteristics can be considered in a consistent way, because all providers will have a benchmark value calculated based on a set of common factors.
311. In relation to comments that we should not take into account evidence that a provider has stopped or has planned to stop delivering poor performing courses, because this infringes on institutional autonomy, our experience of regulating tells us that this is something providers put forward as evidence in relation to their likely future performance. We understand that respondents were concerned this might incentivise provider behaviour in an undesirable way. However, we consider that it is relevant and we should take into account evidence that a provider has identified for itself that certain courses are not performing well and has as a result stopped delivering those courses. In these circumstances we are able to consider what performance looks like when these courses are removed. We therefore consider that this is relevant context and, as it is evidence of actions the provider has already taken, it is not an infringement on autonomy. We have set out in the phase three consultation that we would want to understand the timing and reasons for courses being discontinued, so that we could establish whether a provider was taking action only to avoid regulatory consequences. If this was the case then we would not be likely to give significant weight to the impact of the removal of courses in our assessment.

312. Respondents raised a large number of other factors that might be taken into account as context. We are proposing in the phase three consultation that providers will be able to provide any evidence that they consider relevant context as part of the assessment process for condition B3, and therefore all of the factors mentioned could be considered as part of our assessment. In presenting evidence of context, a provider should be able to demonstrate how that evidence has affected its performance such that it does not meet the numerical thresholds we have proposed. We are likely to put more weight on the evidence of factors that have not been accounted for already in the process of setting baselines and the assessment of performance, and that are genuinely outside the control of the provider. A provider's mission and strategy, the type of students it recruits, and its entry requirements are all things that we consider are within its control, and therefore are not likely to be given as much weight, although this will be dependent on the circumstances of the particular case. More detail and examples of context are provided in our phase three consultation.
313. In relation to comments about the differential impact of the pandemic on providers, we recognise that context is specific to a provider and the same factors will have a different influence on different providers. This is why we cannot provide a definitive answer on what weight we will give to particular factors – factors that are very relevant in one case may not have as much relevance in another. Providers are able to provide evidence of context and the particular impact it has had on performance, and this will be taken into account in our assessment.
314. Our phase three consultation has set out further details of our proposals for the assessment of condition B3, including the balance between indicators and the process for regulatory intervention. As set out in the phase one consultation, condition B3 will be assessed by the OfS itself, because we have significant expertise and capacity to produce the relevant indicators and to use them to analyse performance.

Assessment of initial condition B3 for providers seeking registration with the OfS

315. The consultation document suggested that initial condition B3 may not apply to a provider seeking registration without a track record of delivery of higher education. One respondent sought more information about how this would work in practice and how long the condition would not apply. Another sought further information about what would happen to a new provider that failed to meet baselines once sufficient data became available.

Our response

316. The proposal was that condition B3 would not apply as an initial condition in the circumstances where a provider had not previously delivered higher education and there are therefore no outcomes to assess. Condition B3 would still be applied as an ongoing condition, and the provider would be assessed against the condition as soon as sufficient data was available to draw reliable conclusions about its performance.
317. Our proposed condition B3 sets out that the initial condition will only apply to providers where recent data about higher education students has been submitted to either HESA or the Education and Skills Funding Agency. More detail of our proposed approach to the assessment of initial condition B3 is in Proposal 5 of the phase three consultation.

Definition of successful student outcomes

318. The phase one consultation proposed that the definition of quality and standards for successful outcomes included 'Students have the right skills from their course once in employment and employers are satisfied with the graduates they employ'.
319. Some respondents agreed that it was important to include employer perspectives within a definition of successful outcomes, and this was a valid part of any assessment of quality. However, a larger number suggested it was not appropriate to use employer satisfaction as a measure. It was suggested that there are a large number of variables outside of the provider's control that may impact on an employer's satisfaction, significant issues with measurement and barriers to collecting reliable data.
320. Respondents were not clear how a numerical baseline would be derived for the definition 'Students have the right skills from their course once in the employment and employers are satisfied with the graduates they employ'; whether there was an intention to survey employers to confirm that they are satisfied with the graduates they employ, and have the relevant skills and knowledge; and how this would relate back to individual providers and programmes of study, or take into account graduates who are self-employed.
321. Some respondents questioned whether the intention was to require a provider to be 'above the baseline,' and suggested that the definition should be revised to 'at or above the OfS numerical baseline' to confirm that, if performance is the same level as the baseline, the minimum expected requirement must have been met.

Our response

322. We consider that students developing the right skills for employment from the employer perspective is an important element of quality. It was not our intention to set a numerical baseline in relation to employer satisfaction with the graduates they employ, since as noted by respondents there would be difficulties with collecting reliable data. The development of appropriate skills was proposed as a part of condition B1, and we are not proposing to use this element in condition B3.

Coherence with our approach to access and participation

323. Some respondents made comments about the coherence of our proposed approach to the assessment of condition B3 with our approach to access and participation. Respondents commented that assessing the same outcomes for underrepresented groups in different ways as part of the process to approve an APP and for condition B3 would risk creating an inconsistent and confusing regulatory landscape. Additionally, respondents suggested it would be unhelpful to use different data, for instance the inclusion of international students for B3 but not for APPs.
324. It was also noted that awarding bodies would be held accountable for the performance of autonomous providers in partnership arrangements, often registered in their own right, but would not have the data for the same students counted for the purpose of their APP performance.

325. Responses also suggested that requiring providers to improve their outcomes in absolute terms was likely to cause providers to increase their entry requirements or reduce their use of contextualised admissions, in order to recruit students with a better chance of achieving positive outcomes, and this could adversely affect providers' performance against the targets in their APPs.

326. Similarly, it was suggested that if the benchmarking approach used for APPs were not adopted for assessing performance against condition B3, closing an access and participation gap could mean a provider dropping below a B3 baseline, thus breaching the condition.

Our response

327. In our document 'Consultation on regulating quality and standards in higher education: Analysis of responses', published in July 2021, we set out our view that we consider the approach to assessment of condition B3 to be fully aligned with our approach to access and participation delivered through access and participation plans and condition A1.¹⁴ APPs currently set out the actions being taken by a provider to reduce gaps between different student groups with reference to access, success and progression. Our proposals for condition B3 would provide a further safeguard, in the form of a minimum requirements for student outcomes that would apply to students from all backgrounds. Our view is that regulating minimum requirements for quality and standards is a safeguard for both quality and equality, because it ensures that all student groups (including those with protected characteristics) do not receive an education where the quality or standards are below a minimum acceptable level.

328. In relation to the comment about awarding bodies, APPs relate specifically to qualifying persons on qualifying courses and therefore do not allow for the inclusion of the outcomes of students at partner providers.¹⁵

329. We will be considering our approach to APPs, having due regard to the guidance from the Secretary of State, and will ensure that the approach aligns with condition B3.

330. We have set out in our phase three consultation proposals that we would look at benchmarked performance as a way of helping us to understand the context of a provider's performance.

Coherence with TEF

331. A few respondents made comments about the coherence of our proposed approach to regulating student outcomes with our approach to the TEF. The main themes related to whether the proposals were consistent with our operation of the TEF, and how easy it would be for external stakeholders to interpret providers' performance.

¹⁴ OfS 2021.23, available at www.officeforstudents.org.uk/publications/consultation-on-regulating-quality-and-standards-in-higher-education-analysis-of-responses/.

¹⁵ Further detail of the coverage of access and participation plans can be found in paragraphs 7 to 11 of 'Regulatory Notice 1: Access and participation guidance' (OfS 2021.38), available at www.officeforstudents.org.uk/publications/regulatory-notice-1-access-and-participation-plan-guidance/.

332. Responses suggested that for TEF to continue to use benchmarks and adjust for student diversity would be in contradiction to the proposed approach to condition B3. If TEF and B3 assessments used different approaches, this would imply that baselines are absolute whereas excellence is contextual.

333. However, it was considered positive that a system of weighting as used in the TEF (i.e. weighting for NSS metrics and for graduate outcomes) was not included in the proposals, as this could place undue emphasis on graduate outcomes at the expense of other metrics.

334. In addition, respondents suggested that the differences in the approaches to condition B3 and TEF means that stakeholders would need to ascertain whether a provider meets the baselines as well as whether it exceeds them to correctly interpret a provider's performance.

Our response

335. As stated in 'Consultation on regulating quality and standards in higher education: Analysis of responses', published in July 2021,¹⁶ our proposed approach is to regulate to secure a minimum baseline for student outcomes and to use TEF to incentivise excellence and innovation above this baseline and ensure continuous improvement. The consultation on TEF sets out proposals for how it will interact with our baseline requirements expressed in the quality conditions, such that a provider would not be eligible for participation in TEF unless it was judged to have also satisfied the baseline requirements for quality.

Regulatory burden

336. A large number of respondents thought the proposals would increase regulatory burden and also commented on the cumulative additional burden of the proposals in relation to access and participation and the TEF, suggesting there could be duplication between these regulatory activities.

337. Respondents suggested that there would be increased burden in relation to an increase in data returns to inform the proposed indicators, particularly for providers acting as validating bodies if they are required to collect and submit data for students registered with the validating partner and for TNE courses. Respondents suggested that any additional returns should align with other returns already made by providers to governmental and other regulatory, statutory, or professional bodies.

338. Other respondents suggested that the increased volume and complexity of indicators, particularly if levels of study were extended or indicators for subjects were introduced, would add regulatory burden due to the volume of data that would need to be analysed and understood by providers. It was suggested that this would be disproportionate in the context of a high-performing sector and what are likely to be a small number of pockets of poor performance.

339. It was suggested that the effect for small providers and further education colleges would be disproportionate because these providers generally have fewer staff and resources than larger

¹⁶ Available at www.officeforstudents.org.uk/publications/consultation-on-regulating-quality-and-standards-in-higher-education-analysis-of-responses/.

providers with the infrastructure in place to generate any additional data required and to monitor their own performance.

340. Respondents also commented that small providers and further education colleges would be more likely to have to provide additional contextual and other information to the OfS, because data would be less reliable based on smaller student numbers, and that any increases in administrative costs resulting from added regulatory burden could be passed on to students, or could reduce funds available to support students with specific needs or with protected characteristics.

Our response

341. We recognise that all regulation has the potential to impose some burden on regulated organisations; however, this burden needs to be balanced with the need for effective regulation. This being the case, we will consider whether imposing further requirements on providers is proportionate to the objective sought. We acknowledge that additional data returns in relation to TNE and validated provision will be required, and this will inevitably add some burden. However, we consider this necessary to ensure that we can regulate effectively on the quality of provision, including the outcomes delivered, and, in the case of partnership provision, to meet our statutory obligations in relation to the publication of information.

342. We have heard broad agreement from respondents that it is important to investigate pockets of poor performance within larger providers or cohorts of students. We agree with this, but the trade-off is more detailed data, which respondents also suggested leads to increased burden.

343. We will provide a B3 workbook to each provider each year, which should limit the need for providers to generate additional data for themselves. We would have expected providers to already be considering their performance against the indicators that we propose to use for condition B3 as part of their academic governance arrangements. Where this is happening there should not be a significant increase in burden. Where this is not already happening, we think it is appropriate that providers seek to interrogate data about their own performance.

344. We consider that our proposed use of a measure to consider statistical confidence in the data will mean that providers with smaller populations will not be unnecessarily burdened because of the size of their student population. Our regulatory interventions will focus on courses where there is a sufficient level of statistical evidence that performance is below a numerical threshold. In these circumstances, regulatory burden is necessary to protect students.