Certified B Corporation

SHIFT INSIGHT

Evaluation of the revised condition of registration on student outcomes (B3): Case study thematic report

Report to the Office for Students by Shift Learning

WWW.SHIFT-LEARNING.CO.UK

SHIFT LEARNING

Contents

WWW.SHIFT-INSIGHT.CO.UK

Executive summary

Introduction and method

The Office for Students (OfS) regulates higher education provision in the interests of students, aiming to ensure that 'every student, whatever their background, has a fulfilling experience of higher education that enriches their lives and careers'[1]. A core part of this regulation is that higher education providers deliver positive outcomes for their students on higher education courses, so that students continue after the first year of their course (continuation), complete their qualification (completion) and progress to outcomes such as professional employment or further study after graduation (progression). The B3 condition of provider registration, which forms part of the regulatory framework, reflects this requirement. The OfS consulted on revising the B3 approach over three phases and, in July 2022, published the revised B3 condition[2].

As part of a wider programme of activity to evaluate the introduction of revised B3, the OfS commissioned Shift Learning to conduct interviews with 40 higher education providers between April and June 2023 (stage 1) – exploring the initial response of the sector to the introduction of the revised approach. The related research report can be found on the OfS website[3].

Following stage 1 of the evaluation, the OfS commissioned Shift Learning to conduct a small number of deep-dive case studies to explore and better understand any provider-level changes that revised B3 may have driven.

The aims of this evaluation were to:

- Identify **positive** changes – arising specifically as a result of revised B3, or broader drivers of change.
- Identify if any **negative or unintended** changes had arisen – specifically as a result of revised B3, or broader drivers of change.
- Test the **assumptions** of the OfS theory of change (see appendix 1).
- Identify **external factors** that could explain or have contributed to changes.
- Identify the **role of provider context** in influencing behavioural change and outcomes.
- Explore and identify whether **regulatory burden** impacted behavioural change.
- **Empirically trace** the process linking the introduction of revised B3 to behavioural change within providers.

A small sample of providers was identified to reflect a range of behavioural change areas emerging from interviews undertaken in stage 1. Individual case study reports on the resulting sample of four providers were developed for OfS internal use, but will not be published to protect provider anonymity. The findings presented in this report are overarching themes emerging from across these four case studies.

The sampled providers were largely focused on making quality and monitoring changes. At the time of data collection (spring 2024), any provider actions taken in response to revised B3 would have been in the context of their interpretation of the OfS guidance on implementation[4], the OfS-published assessment criteria[5] (though no OfS B3 assessment work was completed at the time), and OfS B3 dashboard data[6]. The sample did not include any providers who had been selected for a formal assessment of compliance

[1] www.officeforstudents.org.uk/careers/explore-the-office-for-students
[2] www.officeforstudents.org.uk/media/0dc38475-3730-4173-88e7-42989be88262/revised-condition-b3-student-outcomes.pdf
[3] www.officeforstudents.org.uk/media/3xqhy2nr/evaluation-of-the-revised-condition-of-registration-on-student-outcomes.pdf
[4] www.officeforstudents.org.uk/publications/regulatory-framework-for-higher-education-in-england/part-v-guidance-on-the-general-ongoing-conditions-of-registration/condition-b3-student-outcomes
[5] www.officeforstudents.org.uk/media/7737/statement-on-prioritised-categories-for-2022-and-2023-assessment-cycles.pdf
[6] www.officeforstudents.org.uk/data-and-analysis/student-outcomes-data-dashboard/data-dashboard

with the condition by OfS. It is important to note that across all four providers, changes were in the early stages of implementation and so findings here are presented as early indicators of outcomes.

Key findings

This report explores the evidence collected and considered at different points in the four providers' overall response to revised B3, including:

- Their awareness of revised B3
- The process of problem identification
- The process of choosing required action
- Implementation of changes
- Any intermediate outcomes from these changes.

Awareness of revised B3

Awareness and understanding of revised B3 was a pre-requisite for any subsequent behavioural changes to have been motivated by revised B3.

- Awareness of revised B3 was highest among quality and senior team roles in all four providers. These staff were responsible for staying updated about regulatory changes and their impacts, as well as engaging in discussions about the implications of not meeting B3 thresholds, often focusing on potential sanctions and reputational damage.
- In the two smaller providers, this knowledge was concentrated among one or two core team members, while the two larger providers had broader quality and strategy teams involved.
- Despite the initial time investment required, understanding B3 was seen as necessary and important for quality and senior teams, with little perception of it as an unnecessary burden.
- In all four providers, awareness of revised B3 varied outside of those with core regulatory responsibility, with many module leads and academics being unaware or only vaguely familiar with it.
- Efforts were made to embed revised B3 knowledge more broadly through staff briefings and informal discussions, aiming to increase academic staff buy-in and ownership of revised B3 data and action planning, while making staff aware of the regulatory context and potential implications of non-compliance.

Problem identification

Case studies explored how providers reviewed and used the published revised B3 data in particular, and whether this was used to inform decision-making about where action was required.

- Key decision-makers in all four providers prioritised internal data on student outcomes and staff insight over published B3 data when identifying areas of need. Internal data was seen as more up-to-date and allowed them to stay ahead – proactively identifying and addressing potential issues.
- All four providers reviewed B3 data dashboards, but found the publication lag made data less useful for proactive decision-making. Instead, they used the dashboards for verification and comparison with internal data and as an additional pulse check.
- There was existing emphasis on data-informed quality monitoring before the introduction of revised B3. Providers were already gearing internal data and reporting towards the regulatory framework introduced in 2018, making revised B3 a continuation of this rather than a significant shift in focus.
- While B3 data was not the primary driver for changes, in these four providers it did help with confirming and expediting priority areas for action, which seemed largely driven by its regulatory nature and potential implications of falling below threshold. Reviewing revised B3 data alongside data from other areas, such as the Teaching Excellence Framework (TEF) and National Student

Survey (NSS), was a crucial foundation for evidence-informed decision-making across these four providers.

Choice of action

Case studies reviewed how providers decided on actions required after identifying areas of need.

- In all four providers, decisions on what action to take were led by quality teams, with senior-level approval required. Wider staff consultations also fed into this process, to secure buy-in and gather feedback on current processes and improvements needed.
- A desire to improve and enhance performance and practice was stated as the primary motivation for actions in all four providers, not directly linked to revised B3.
- In the three case studies where providers were focused on quality and monitoring processes, there was evidence through interviews and supporting documents that decisions on required actions were ongoing before the introduction of revised B3. These decisions were generally influenced by contextual factors specific to each provider, as part of which the wider OfS regulatory framework, including original B3, may have sometimes consolidated the need to act.
- The OfS regulatory framework, including B3, was seen to have informed continual policy and practice reviews, even if not driving them.
- Although revised B3 was not the initial driver for changes, the potential regulatory implications of not meeting revised B3 thresholds appeared to have helped progress actions and ensure compliance in all four providers, in which revised B3 provided additional motivation for prioritising improvements.
- Provider actions related to revised B3 included the introduction of course- and module-level action plans incorporating B3 thresholds, with module leaders required to report performance against these to determine what action may be required. B3 data was also used to make a case for student success coaches in one provider and was judged as making a strong contribution to getting this initiative ‘over the line’.

Implementation

Case studies explored the actions that providers were implementing, which were slightly different and bespoke at each provider, and their context. They also highlighted any barriers that hindered the implementation of actions.

- While not directly linked to the introduction of revised B3, changes to improve student outcomes, across all four providers, were perceived as contributing to increased workload for academic and professional services staff, both now and in the future.
- Required actions associated with revised B3 were perceived by quality teams as necessary to enhance performance and practice. In this respect, revised B3 was not directly mentioned by any of the four providers as creating an additional burden.
- Implementing changes to processes required substantial internal time, resources and technical skills for data collection, processing and analysis. While this was seen as necessary work to ensure quality provision and ensure baseline regulatory requirements in quality are met, it did pose challenges for the providers.
- The two smaller providers explored in these case studies also faced additional challenges due to a lack of data-analysis staff, necessitating new hires to manage data tasks and increased workload for existing staff in quality-related roles. While new hires were not focused only on revised B3 requirements, it did fall within their remit.

Emerging outcomes

Case studies aimed to identify behavioural change to achieve specified outcomes and whether revised B3 had supported or driven these behaviours.

- All four providers planned to use internal data and reporting to track and monitor the success of changes. Despite the early stage of implementation, it was anticipated that any revised quality-monitoring processes would lead to improved student outcomes over time, due to the subsequent actions they would help to identify a need for.
- None of the providers planned to use B3 data dashboards to track success. This was due to the lag in publishing this data, with providers perceiving that it did not give a contemporary picture of performance.
- However, evidence from the three case studies in which providers were focused on quality monitoring indicated (through interview accounts and supporting documents) that revised B3 terminology, thresholds and/or definitions were being used in programme and module monitoring, as well as providing the scaffolding for action planning, prompting a more detailed programme and module-level reflection.
- The admissions-related case study revealed anecdotal evidence of a clearer, more transparent admissions process and reduced student dropout rates, although it was too early to officially track these outcomes using internal data.
- One quality and monitoring case study showed evidence of revised B3 data specifically playing a role in contributing to provider-wide strategic monitoring and development of institutional KPIs.
- Within the same provider, revised B3 data was also used to make a case and move forward with student success coaches and enhanced student learning resources.
- In one of the quality and monitoring case study providers, we also heard how revised B3 data had been used in their decision to enter into a partnership for collaborative provision with another provider, as part of their due diligence process. No significant concerns were flagged in a review of the provider's B3 data, but if there had been any concerns about the partner provider falling significantly below thresholds, this may have influenced their decision to enter the partnership.

Conclusion

Findings from this research highlight the positive work going on within higher education providers to continually enhance their performance and practice. While revised B3 did not emerge as a strong driver for change in any of the four providers, we did see evidence that – as part of the regulatory landscape – revised B3 was supporting and expediting the drive for continuous improvement.

We found several examples of B3 data prompting targeted course- and module-level action planning, in which B3 thresholds formed a central pillar. If performance was below B3 thresholds, or perceived as in danger of falling below them, course and module leaders were required to explain reasons for this performance and devise a specific and measurable action plan. We heard that action planning per se was not new, but B3 data offered a more nuanced and targeted approach. One provider had also incorporated B3 data within their broader institutional KPIs.

The regulatory aspect of B3 was mentioned several times as a useful tool for mobilising support and resource. Three of the four providers had specifically mentioned B3 in staff briefings conducted about proposed changes to quality and monitoring, as a means of gaining 'buy-in' from academics.

Use of thresholds within B3 (original and revised) seemed an important step for providers on their journey towards being more data-informed. Our discussions with key contacts suggested that B3 thresholds had motivated further focus on data-informed monitoring and evaluation. That is not to say providers were not using data to inform policy and practice pre-B3, but more that B3 had further enhanced this focus.

It was too early to observe specific and measurable outcomes regarding the impact of the changes that providers discussed. However, quality monitoring and evaluation changes did seem to provide a foundation for progressing enhanced student outcomes, with revised B3 contributing to how providers were approaching changes and enhancing practice.

Introduction

Background of the revised B3 condition

The Office for Students (OfS) regulates in the interests of students, aiming to ensure that 'every student, whatever their background, has a fulfilling experience of higher education that enriches their lives and careers'. A core part of this regulation is that providers deliver positive outcomes for their students on higher education courses, so that students continue after the first year of their course (continuation), complete their qualifications (completion) and progress to outcomes such as professional employment or further study after graduation (progression). Forming part of the OfS's regulatory framework, the B3 condition for provider registration (which covers student outcomes) reflects this requirement.

Providers had been following the February 2018 version of the regulatory framework[7], including the original version of the B3 condition, before the consultation on the revised B3 condition and subsequent consequential amendments to the regulatory framework[8]. This is an important contextual factor, as changes to process and practice found as part of this evaluation may have been influenced by earlier work set in motion to meet the 2018 regulatory framework.

In the original B3 condition, a range of student outcome measures were used to evidence the performance of a provider over time, mainly focused on continuation, degree outcomes and progression for first degrees were also considered. Completion was a new measure introduced in the original B3 condition and was included for context rather than for use as the basis on which to make compliance judgments.

The OfS consulted on revising the B3 approach over three phases and, in July 2022, published the revised condition B3[9] as part of the regulatory framework. Revised B3 applied to all registered providers from 3 October 2022. The revised B3 regulatory expectation for positive outcomes starts with a provider's outcome data for three measures: continuation, completion and progression into managerial or professional employment, further study or other positive outcomes. The three indicators of continuation, completion and progression are further categorised by mode and level of study. For each indicator, the OfS has published a numerical threshold setting the minimum performance a provider should be delivering. The OfS will consider a provider to be 'delivering positive outcomes' if it is performing at or above each of the numerical thresholds in relation to the indicators and split indicators, or if the OfS assesses the provider's context to justify the performance below threshold. The B3 compliance assessment process is set out in Regulatory Advice 20[10].

Stage 1 revised B3 research

As part of a wider programme of activity to evaluate the introduction of revised B3, the OfS commissioned Shift Learning to conduct interviews with higher education providers between April and June 2023 (stage 1) to explore the initial response of the sector to the introduction of revised B3.

Stage 1 interviews were completed with quality contacts from 40 higher education providers, sampled to reflect variation in financial typology and data for the B3 indicators. Participants were all in roles with responsibility for B3 compliance. These stage 1 interviews specifically explored: providers' understanding of the condition; how they approached self-evaluation of their compliance with B3; and the extent to which B3 (or other factors) had contributed to institutional changes across a range of areas, including quality and

[7] www.officeforstudents.org.uk/publications/regulatory-framework-for-higher-education-in-england/
[8] www.officeforstudents.org.uk/media/7540/b3-consequential-amendments-to-rf.pdf
[9] www.officeforstudents.org.uk/media/0dc38475-3730-4173-88e7-42989be88262/revised-condition-b3-student-outcomes.pdf
[10] www.officeforstudents.org.uk/publications/regulatory-advice-20-regulating-student-outcomes/#:~:text=This%20regulatory%20advice%20sets%20out,under%20condition%20of%20registration%20B3

monitoring processes, student support and admissions policies. The full report and findings from this research can be accessed on the OfS website[11].

[11] www.officeforstudents.org.uk/media/3xqhy2nr/evaluation-of-the-revised-condition-of-registration-on-student-outcomes.pdf

Evaluation aims

Understanding the context of how providers might respond to B3

Following the stage 1 research, in autumn 2023 the OfS commissioned Shift Learning to conduct a deep-dive exploration of any provider-level changes B3 may have driven, through a small set of provider case studies. Individual case study reports have been developed but not published, to protect provider anonymity. This report presents findings from across all four – highlighting emerging themes and the local conditions that may have resulted in similarities and differences regarding providers' responses to revised B3.

In discussion with the OfS, and in light of findings from stage 1 research, specific areas were highlighted for further investigation via the case studies. These were:

- Changes to quality assurance, monitoring and enhancement processes
- Changes to course or portfolio offer
- Impact on student support and pastoral care
- Impact on admissions
- Changes to partnership activity.

Stage 1 interviews suggested that revised B3 may play a part in explaining changes in each of these areas – however, alternative explanations may have been possible and were fully explored. Providers' contextual differences and their reactions to regulatory burden were also considered across all case studies.

Evaluation aims

The aims of this evaluation were to:

- Identify **positive** changes – arising specifically as a result of revised B3, or broader drivers of change
- Identify **negative or unintended** changes – arising specifically as a result of revised B3, or broader drivers of change
- Test the **assumptions** of the OfS theory of change (see appendix 1)
- Identify **external factors** that could explain or contribute to changes
- Identify the **role of provider context** in influencing behaviour change and outcomes
- Explore and identify whether **regulatory burden** impacted behavioural change
- **Empirically trace** the process linking the introduction of revised B3 to behavioural change within providers.

Evaluation approach and methods

Theory-based impact evaluation case studies

We utilised a theory-based case study approach for this evaluation. This involved creating a theory of what the impact of revised B3 was likely to be, then systematically looking for evidence to prove or disprove the theory, while accounting for other influencing factors.

While stage 1 research aimed to understand the wider awareness and context for the impact of revised B3, the case studies aimed to understand how and why providers had changed behaviours, and the extent to which these may be attributed to the introduction of revised B3 or other drivers.

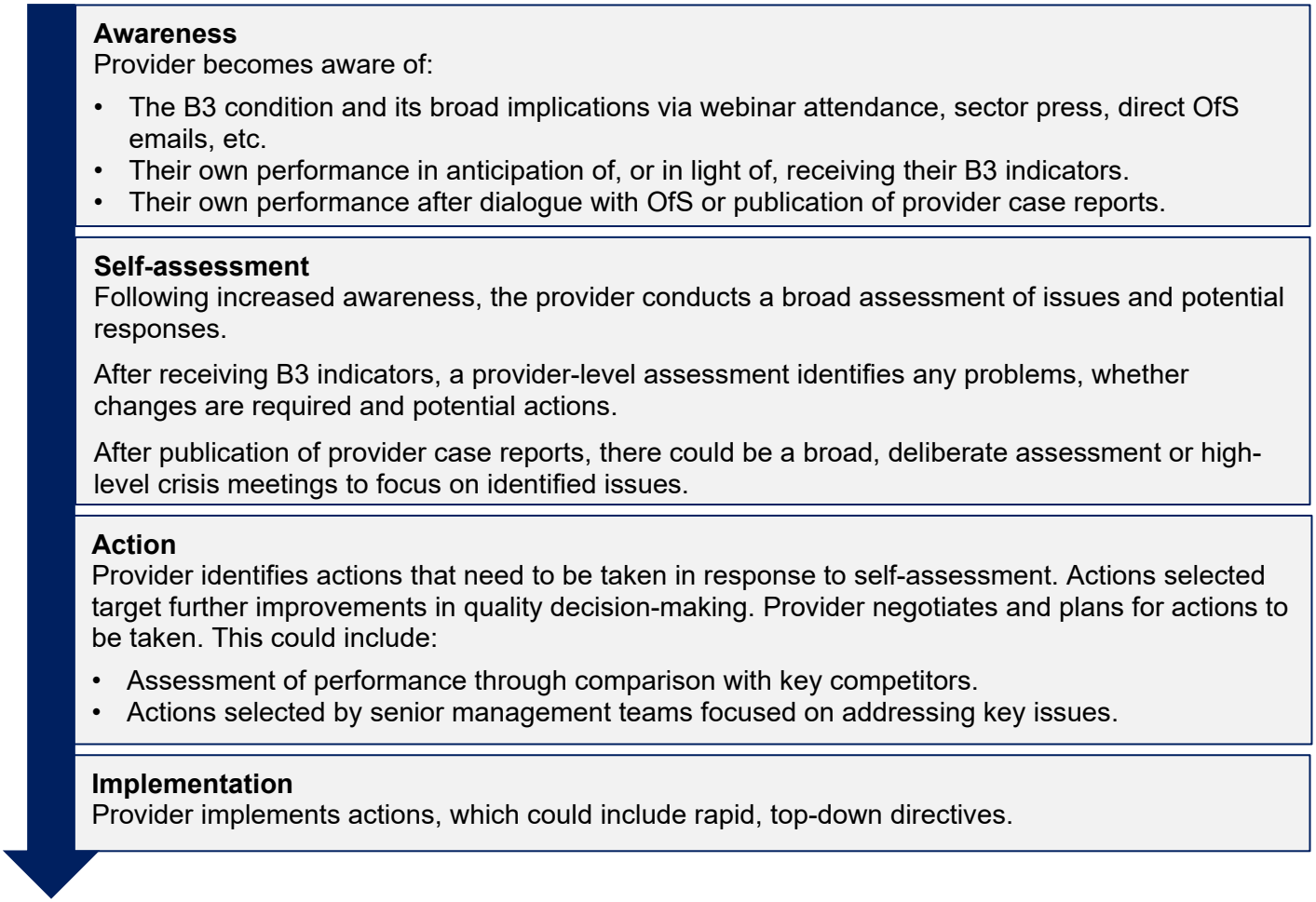
The OfS theory of change (see appendix 1) was used as a starting point to frame the case study evaluation. We were looking to explore the mechanisms of change that occurred in each provider, to link the inputs and activities to subsequent intermediate outcomes (i.e. behavioural change areas) and how this could be traced through to emerging direct or indirect outcomes.

We aimed to explore the impact of revised B3 across four distinct phases, enabling a deep dive at each phase. These phases were:

1. **Awareness:** Provider becomes aware of revised B3.
2. **Self-assessment:** Provider reviews their performance against revised B3 requirements and data dashboards.
3. **Action:** Provider identifies any action that needs to be taken.
4. **Implementation:** Provider implements actions and changes.

Figure 1 details some of the activities we expected to see across these phases of change and which we probed for further information and evidence.

Figure 1: B3 case study phases of change activities examples



Within the phases, we wanted to determine the motivations for any change we observed. We anticipated the research revealing various drivers, but by drawing on the OfS theory of change (appendix 1), we identified these four key drivers of change as a starting point:

- Desire to improve performance
- Fear of sanctions
- Fear of reduction in recruitment
- Fear of reputational damage.

We created an evaluation matrix for each case study (see an example in appendix 2) that documented anticipated motivations for change, combined these with the four defined phases and added specific questions we wanted to ask during each phase.

Sampling

Providers were purposively selected based on six behavioural change areas in which changes had occurred or were planned, based on what had been mentioned in stage 1 research. These were: 1. Quality assurance processes, 2. Partnership activity, 3. Courses or portfolio offer, 4. Student support and pastoral care, 5. Support for at-risk groups, 6. Admissions.

We know that provider context – particularly size – has a strong impact on the level of challenge around regulatory burden. This has been a trend in much of the research Shift Learning has conducted for the OfS

and is highlighted in our provider engagement report[12]. While it was not feasible to reflect the whole sector within a limited number of case studies, we did therefore aim to represent various provider financial typologies[13], regions and B3 threshold groupings and used this as secondary sampling criteria (see appendix 3 for sample breakdown). We did not include providers who fell outside these typologies – those in the 'unclassified' category – in case it was too easy to identify them. The sample did not include any providers who had been selected for a formal assessment of compliance with the condition by the OfS.

Based on the above, we started the evaluation fieldwork with a sample of six providers. However, two of those that agreed to participate stopped responding during fieldwork, and so those case studies were unable to proceed. These providers were both identified as having made some level of change to their student support offer. Early scoping calls with these providers did not enable us to gather enough depth on their awareness, process or action in response to revised B3 and so they have not been reported here.

Table 1 outlines the behavioural change area of each of the four providers that did progress to a case study, the changes they told us they had planned when interviewed in stage 1, the methods we used for data collection and which financial typology and B3 grouping they fell into.

Table 1: Overview of method for each case study

Case study	Behavioural change area	Changes planned	Methods	Financial typology and B3 grouping[14]
1	Quality and monitoring	Realigning their continuous monitoring of modules to revised B3 regulation.	Interviews with: • 2 senior members of the quality team • 1 quality team lead (interviewed twice) • 1 senior academic • 1 head of school • 1 strategy lead • 1 employability lead. Document review of policy and planning documents, internal reports, emails, briefing documents, action plans and summary documents on proposed changes to the quality and monitoring system. A short survey to programme and module leaders and professional services staff.	Qualifying income (QI) over £200m. Threshold group 1 (few students affected by indicators below threshold).
2	Quality and monitoring	Realigning their annual quality reporting to revised B3 and piloting a new approach to module monitoring	Interviews with • 1 director of quality (interviewed twice) • 1 director of planning • 1 associate dean of student outcomes – also a module leader (interviewed twice) • 2 module leaders	Specialist: other. Threshold group 3 (a large proportion of students affected by indicators below threshold).

[12] www.officeforstudents.org.uk/publications/report-on-provider-engagement-with-the-office-for-students
[13] www.officeforstudents.org.uk/publications/provider-typologies-2022
[14] See appendix 4 for information about provider grouping.

Case study	Behavioural change area	Changes planned	Methods	Financial typology and B3 grouping[14]
		across 10 different modules.	• 1 ed-tech portfolio manager • 1 data analyst (student experience/outcomes). Document review, examining committee meeting agendas and minutes, teaching and learning strategies, academic regulations, module-evaluation data and reports, module-enhancement plan templates, a funding application for a data analyst role and email communications.	
3	Admissions	Creating a new admissions policy with formalised person specifications and interview criteria.	Interviews with: • 1 head of higher education (interviewed twice) • 4 programme leaders. Document review, examining the 2021/22 admissions policy, 2023/24 admissions policy, 2023/24 person specifications and interview questions from four courses, a PowerPoint presentation used to inform staff of the new policy and a new higher education reporting dashboard (in development).	Majority level 4/5. Threshold group 2 (some students affected by indicators below threshold).
4	Quality and monitoring	Reworking their annual programme quality monitoring and reporting to include B3 metrics of continuation, completion and progression, alongside the key further education metrics.	Interviews with: • 1 director of academic standards (interviewed twice). • 1 group quality director. • 1 higher education quality director. Document review, examining programme-review templates, programme self-review documents and papers to the divisional board.	Majority level 4/5. Threshold group 2 (some students affected by indicators below threshold).

Process-tracing approach

To comprehensively meet the aims of the evaluation, we adopted a process-tracing approach. Process tracing (within monitoring and evaluation) aims to establish whether something influenced a specified change or sequence of changes. Process tracing can help show not only whether a change occurred, but

how and why[15]. This was especially valuable for evaluating whether observed changes across the selected providers were influenced or caused by the introduction of revised B3 or were unrelated.

Data collection and analysis

Data collection focused on the knowledge, decisions and activities involved across the four phases of awareness, self-assessment, action and implementation.

Initial key informant interviews

As a starting point for each case study, an interview was conducted with the person who took part in stage 1 interviews and had an overview of B3-related activity across the provider. This person was able to confirm current practice and any progress made between the stage 1 interview and this research.

Interviews with wider provider roles

Interviews were a core method in each case study. These were held with roles identified during the key informant interview and included those with an overview of activity in the provider and those involved in decision-making and sign-off of new activities. These roles included staff in registry, strategy and planning roles, quality and monitoring leads, and faculty or department leaders. We did not limit interviews to leadership roles, as the process-tracing method we utilised led us to both academic and professional services staff with involvement in the areas under exploration.

Document review

We reviewed a broad range of documentary evidence to confirm actions mentioned in interviews or to elicit further information.

Documents were reviewed across all case studies and included annual quality reports, committee meeting minutes in which B3 was an agenda item, module-evaluation documents and relevant policy documents (e.g. around admissions).

Faculty surveys

A short survey was disseminated to programme and module leaders and professional services' staff in the first case study. Despite reminders, this had a low response rate, but allowed us to capture their views on the changes made to the quality and monitoring system and why they thought those changes were made, as well as corroborating some of the statements made by key decision-makers in interviews.

Data analysis

We developed a workbook for each of the case studies that mirrored the evaluation matrix (see appendix 2). This acted as a single point for data collection and analysis, including links to evidence and process-tracing tests, while documenting what further evidence was required and interviewers' initial thoughts and reflections. Following each data-collection episode (e.g. interview, document review, survey), our case study team met to apply any new evidence to the workbook, calibrate findings, and record initial and ongoing findings across all four case studies. We also applied process-tracing tests using a strength of evidence table (see appendix 4), which enabled us to review data and assess how evidence either strengthened or weakened our existing findings, including assessing contradictory accounts.

[15] Punton, M and Welle, K (2015). Straws-in-the-wind, Hoops and Smoking Guns: What can process-tracing offer to impact evaluation? IDS CDI Practice Paper, No. 10, 2015.

Limitations and mitigations

Process-tracing method

The process-tracing method facilitated a rigorous and transparent approach to assessing contribution to change. However, process tracing can sometimes be viewed as a 'top-down' deductive method. Researchers are entering the field to prove or disprove whether something is true and, in doing so, there is a risk of finding what they expect to find (confirmation bias). To guard against this, if evidence was not found to support the hypothesis, we fully discussed the potential reasons and also any alternative explanations for how the hypothesis could have been disproved. Once the alternative explanations had been discussed by the research team, this sometimes necessitated follow-up interviews with providers. These follow-up interviews aimed to further probe and test our line of reasoning around the alternative explanations and whether we needed further evidence to strengthen our assessment of the hypothesis being confirmed or disconfirmed.

We further mitigated against confirmation bias through regular interviewer meetings, calibration and reflection, as well as thorough briefing and training of interviewers ahead of entering the field. These regular meetings enabled the research team to question each other about the conclusions they had individually reached from the evidence collected. These meetings enabled each researcher to reflect on their thoughts about the evidence collected as part of process tracing and their summation about what it was telling them of the role of revised B3 within it. Assumptions were challenged and, if necessary, data was revisited to test the neutrality of opinion. As part of this process, we compared specific documentary evidence with what interviewees had told us about this subject area, exploring any similarities and differences. Where interview and documentary evidence did not quite align, we discussed why this may be, and highlighted areas where further evidence may be required to substantiate and corroborate the evidence found. Providers were then contacted for follow-up interviews to collect any further evidence if necessary.

Case studies

We are mindful that the evaluation involved just four providers. Although these providers were carefully selected, they are not representative of all providers. We focused on process and decision-making and the strength of evidence exhibited as part of the process-tracing method (appendix 4). In this respect, we can say that the evaluation was methodologically robust, but caution is still required when making any sector-wide generalisations.

There was also potential bias in the sample of providers that agreed to take part in this research and a potential positive bias in the information shared with us by providers. There was a risk that, due to providers knowing the research was commissioned by the sector regulator, they may have been selective in whether they chose to take part and subsequently in the information they chose to share with us. We mitigated against this through strong and transparent communication about Shift Learning being an independent third party conducting the research and putting in place strong protocols for providers' anonymity and confidentiality in the research, in order to elicit a more honest account of provider behaviour and motivations. OfS staff have not at any point known the identity of the participating providers.

Case studies explored the influence on provider behaviour of revised B3 specifically. It is important to note that the original version of B3 was already being followed as part of the regulatory framework published in February 2018, before the consultation on revised B3 and subsequent consequential amendments to the regulatory framework. This is a key contextual factor, as changes to process and practice found as part of this evaluation may have been influenced by earlier work set in motion to meet the 2018 regulatory framework. In interviews and document reviews, we probed into the influence of wider regulatory conditions, including the original B3, and used the timeline of events to help establish to what extent revised B3, as opposed to original B3, was seen to be driving change.

It is important to take into account the timing of the research when considering the findings. At the time of data collection (spring 2024), any provider actions taken in response to B3 would have been in the context of their interpretation of the OfS guidance on implementation[16], the OfS-published assessment criteria[17] (though no OfS B3 assessment work was completed at the time), and OfS B3 dashboard data[18]. It is important to note that across all four providers, changes were in the early stages of implementation and so findings here are presented as early indicators of outcomes.

Generalisability

When discussing generalisability, it is important to note the difference in purpose and end goal between case studies and quantitative studies. While statistical findings are mainly generalised to populations, case studies develop a rich and in-depth investigation of phenomena to generalise to circumstances and situations, with the help of in-depth analytic investigation. The method adopted for this evaluation takes this further, with the use of process tracing to assist in the critical exploration of what we were told within interviews and sense checking of verbal statements via documentary evidence wherever possible.

For this research, we explored motivations for change across four case studies, investigating behaviours, processes, practices and context. We looked for similarities and, where found, these contributed towards what is termed analytic generalisation[19] and have been included in this report.

It could be argued that four case studies were an insufficient volume for analytic generalisation. However, across these four case studies, we interviewed numerous people and examined a volume of documentary evidence. This was all situated within a theory-based evaluation that was underpinned with a theory of change and based on discussion with participants from stage 1. Cases were purposively sampled to exhibit variation in provider typology and the context of providers has been fully recognised and made explicit. This has contributed to ensuring a robust in-depth investigation that can offer some cautious generalisations across case studies. That said, we would not want to say findings are a reflection of what may be found across the higher education sector and any generalisations are only valid at the particular time of the study.

[16] www.officeforstudents.org.uk/publications/regulatory-framework-for-higher-education-in-england/part-v-guidance-on-the-general-ongoing-conditions-of-registration/condition-b3-student-outcomes
[17] www.officeforstudents.org.uk/media/7737/statement-on-prioritised-categories-for-2022-and-2023-assessment-cycles.pdf
[18] www.officeforstudents.org.uk/data-and-analysis/student-outcomes-data-dashboard/data-dashboard
[19] Yin, R. K. (2012). *Case Study Research: Design and Methods* (p.18). London: Sage.

Findings by phase

Introduction to findings

Individual case study reports have not been published to protect provider anonymity, but this report presents emerging themes and highlights similarities and differences across the four providers regarding their response to revised B3. The following sections explore the evidence collected and considered at different points in providers' overall response to revised B3, including:

- Their awareness of revised B3
- The process of problem identification
- The process of choosing required action
- Implementation of changes
- Any intermediate outcomes from these changes.

Awareness and understanding of revised B3

Awareness and understanding of revised B3 was a pre-requisite for any subsequent behavioural change to have been motivated by the revised condition. This section explores the levels of awareness of revised B3 among participants at the four case study providers.

There was awareness of revised B3 among those in key decision-making roles in each of the four providers. While this revised B3 awareness did not necessarily motivate changes in provider behaviour, it is an important pre-condition to understanding the extent to which revised B3 influenced any changes.

Across the four providers, awareness of revised B3 tended to be highest among quality and senior team roles, for whom OfS regulation formed a core part of their duties. These individuals were also key decision-makers for the area of change being explored as part of this research in their provider (i.e. the specific change to quality and monitoring processes or admissions). They were required to be cognisant of the regulatory environment and how changes to it could impact their provider. In the two smaller, majority level 4/5 providers, this involved only one or two core team members, whereas in the other two providers, this responsibility was spread across larger quality and strategy teams.

When asked about their understanding of revised B3, key individuals involved in decision-making for the area of change being explored were highly aware of how it applied to their provider and the potential impact of not achieving revised B3 thresholds. All spoke of building this knowledge through a combination of reading OfS documentation about revised B3, attending the OfS webinars, being involved in responding to the B3 consultation and reviewing the B3 data dashboards. While it took time to understand and familiarise themselves with the regulatory framework and subsequent revisions to B3, they ultimately saw it as necessary work and an important part of their job role. As such, we did not hear much direct discussion of any unnecessary burden on these roles because of the introduction of revised B3.

As part of building their own awareness, staff in quality and strategy roles across the four providers were engaging in discussions (internally and with other providers) around the implications of not meeting B3 thresholds. These largely focused on potential sanctions and any reputational damage that may come from the publication of the data if it were below threshold and benchmarks. Interestingly, across interviews, the terms 'benchmark' and 'threshold' were often used interchangeably, generally being used as a way of identifying and referring to poor data that was 'under benchmark' or 'under threshold'. Understanding of the differences between these two terms was variable across all four case studies, with some level of ambiguity in use of the terms seen across all four[20].

[20] See Regulatory advice 20: Regulating student outcomes for further detail on B3 thresholds and benchmarks: www.officeforstudents.org.uk/publications/regulatory-advice-20-regulating-student-outcomes

Senior staff from three of the four providers recalled conducting some form of risk assessment when the revised B3 dashboards were published, which involved looking at their internal data and the B3 data dashboard to assess whether they were likely to be selected for OfS investigation and how at risk they were. This risk assessment resulted in a judgement regarding whether they felt OfS investigations and sanctions were likely. Interviewees within these three providers told us that they did not ultimately judge themselves to be at risk based on their B3 data dashboard and their own internal data. As such, a fear of sanctions due to revised B3 was not seen to be a motivator for subsequent changes in behaviour at their provider. It is important to note that sanctions can still form a key reason for compliance without necessarily being a key driver for change in practice.

In three of the four case studies, those in quality and senior roles were particularly keen to increase academic staff awareness of revised B3. This was achieved as part of staff briefings to introduce planned changes at the provider. An underlying motivation for this was to give module staff more ownership of their B3 data and associated action planning. We found that in these providers, revised B3 was being referenced when communicating the rationale and need for change, especially regarding quality-monitoring process and practice changes, as a way of helping to gain buy-in from wider staff. Referencing revised B3 as a driver for change was thought to gain buy-in, as academic staff would be aware of any potential consequences of not meeting B3 thresholds or provider-devised key performance indicators (KPIs) relating to revised B3. In one of the providers, proposed quality and monitoring system changes were seen to require more academic staff time than previous systems, with a fear that academics may not have seen this as directly related to their job roles. Referencing revised B3 in this instance was seen as a solid rationale for academic staff buy-in, given its regulatory role and importance. However, one provider purposely chose not to reference revised B3 when communicating quality and monitoring changes to wider staff. Instead, they positioned changes in terms of the positive impact they would have on students, as this was felt to better resonate with staff than a potential perceived 'threat' of regulatory action.

While B3 was not always mentioned when introducing associated changes, those in quality and senior roles across all four providers mentioned that attempts had been made to embed knowledge of revised B3 more broadly in their provider. This involved a combination of staff briefings and informal discussions about the publication of the B3 dashboards. However, beyond interviewees for whom regulation was a core responsibility, awareness of revised B3 varied. It was common for module leads and academics to be unaware of revised B3 or know of it only by name. Some recalled hearing it mentioned in staff briefings, for example, but could not describe what it was or why it was important.

Problem identification

Case studies explored how the four providers reviewed and used the published B3 data, and if this was used to inform decision-making about where action was required.

Key decision-makers in all four providers told us they used internal data and staff insight as starting points to guide decisions about where action was needed, rather than starting with B3 data. Continuation, completion and progression were being monitored in some way via internal data at all four providers. This included looking at students' academic achievement and progression, attendance and engagement with online learning environments, as well as looking at NSS and Graduate Outcomes Survey results.

While there was clear evidence that the providers had reviewed the B3 data dashboards once available, those in quality teams at all four providers thought there was too much lag in this data for it to be useful in driving decision-making and strategic action, which is why they chose to monitor the same measures using their internal data.

Ultimately, all four providers stated that their internal datasets and reporting had alerted them to any areas of poor performance before they received their B3 data dashboards, meaning these had not prompted them to look at any new areas where action may be required. Providers wanted to

ensure they were ‘ahead of the game’ by using internal data to help them identify underperforming areas that may cause issues in the future. Further, we found that in all three of the quality and monitoring case studies, whilst action had not yet been taken, conversations about required areas of change had begun before the introduction of revised B3.

Reviews of B3 data were another pulse-checking exercise, rather than providing a core motivation to act. Quality leads at all four providers knew the implications of falling below B3 thresholds and wanted to check that their internal data aligned with the OfS’s data when the dashboards were released. They also used the B3 data dashboards to help them assess how they performed against areas they felt the OfS may prioritise for exploration. In these four providers, this was not ultimately a core driver for change, as they did not see themselves at risk of OfS investigation based on their data. Even though, in some cases, they could see metrics that fell below threshold, they were satisfied that there was either contextual information that would be taken into account to explain this, or that the specific area was unlikely to be a priority area of focus for OfS investigations.

In one provider that had made changes to their quality and monitoring processes, there was evidence that the split metrics, published as part of the revised B3 dashboards, had also been used within decision-making. In particular, they mentioned utilising the time series metric to help interrogate their data further, although they did not report any changes being initiated as a direct result of this analysis.

Another provider that had made changes to quality-monitoring processes described how they had also interrogated their split metrics further. However, they had chosen to use internal reporting tools for this, rather than published B3 data. This was a large provider with a dedicated data analysis team and well-established reporting tools and so they felt they had a greater ability to analyse the internal data, such as being able to cross-tabulate and look at two split indicators concurrently.

Comparing B3 data with other providers also helped to provide a pulse-check of provision. Two providers highlighted that they were comparing their B3 data with other similar providers to offer a fuller picture of their competitor landscape. In these cases, a comparison of data against B3 thresholds was added to an existing competitor analysis process, rather than a new process being created because of revised B3’s introduction. B3 data was used alongside other insights to see how they compared to, and could learn from, other providers. One other provider described how they had chosen not to compare their B3 data with other providers, due to the specific context of their provision and student demographic.

We saw evidence of other aspects of regulation feeding into providers’ assessments of which areas of operation may need action, forming part of ongoing quality monitoring and reporting. Participants highlighted that as part of the regulatory framework introduced in 2018, there were existing B conditions, including B3, and an increased focus on monitoring student outcomes before revised B3 came into force. All four providers were aware of an increasing focus on data in the higher education sector and were planning to recruit more roles to help with data-informed quality monitoring, reporting and evaluation. Internal data and reporting were therefore already geared towards the regulatory framework and, as such, the release of B3 data was not seen to drastically change this direction of travel.

The providers from all three quality and monitoring case studies also spoke of other elements that had influenced their assessment of areas that may require attention. Specifically, other regulatory elements and results of other sector data. Revised B3 was part of a larger suite of regulatory tools, including the other B conditions, the A conditions of registration for access and participation, and the TEF, that the providers were using to help identify areas in need of update to improve practice, but with a mind to this needing to be compliant with sector regulation. Results from other sector data such as the NSS and the Graduate Outcomes Survey were also being fed into their internal reporting to help in identifying areas where intervention may be needed, helping to build a fuller picture of provision.

The key decision-maker within the admissions-related case study also mentioned compliance with sector regulation and best practice reference points having fed into their review of areas of need within their higher education policies – specifically naming the Competition and Markets Authority, and the Quality Assurance

Agency (QAA) Quality Code for Higher Education, as well as compliance with the B2 condition of registration.

While B3 data was not found to be an initial motivation for change in any of the case studies, it was often described by interviewees as having confirmed and expedited priority areas when looking at where action was required. For example, in the admissions case study, a review of internal data showed overall completion rates were lower than they would have liked on certain courses and the head of higher education voiced concerns about the implications this could have for their revised B3 data and the potential risk of regulatory action or reputational damage. This fear of sanctions or reputational damage as a result of revised B3 therefore made it all the more important for the provider to look at what was causing low completion rates, and any action needed to improve this. This sat alongside their wider, longstanding motivations related to improving student outcomes, not linked to revised B3.

As such, no single tool was felt to have directly led to specific changes being made, but a combination was perceived to build a comprehensive picture of provision by helping identify areas where they fell below sector expectations and where there may be areas requiring improvement.

Choice of action

Case studies reviewed what providers did to decide on the action required after identifying areas of need.

Decisions regarding what action was required were led by those in data and quality teams across all four providers, with final sign-off required from senior-level colleagues. In each provider, some degree of staff consultation was conducted before any action was decided on. It was common for programme leaders to be invited to give feedback on current processes, areas of best practice and others they felt required improvement. This form of consultation was also thought to help secure staff buy-in to subsequent changes, especially if these were likely to result in more time required of them.

The most cited motivation driving decisions around actions was a desire to enhance student outcomes by improving performance and practice. Across all four case studies, there was evidence to suggest that action was being chosen due to its perceived impact on students, helping to enhance quality at their providers. For example, the key decision-maker within the admissions case study stated that the need to enact changes within their admissions process was largely motivated by a desire to ensure a fairer and more transparent process that would best serve students, with old practices deemed to fall short. Similarly, providers in all three quality and monitoring case studies evidenced meetings across several years (before revised B3) in which processes were regularly reviewed in light of the student outcomes. In this sense, the process of deciding what action was required was not found to be driven by the introduction of revised B3, as a focus on enhancement was previously evident.

While enhancement was stated as a strong motivation behind actions, the wider regulatory framework, including revised B3, did appear to feed into decisions. The providers in all three quality and monitoring case studies stated that policy and practice were continually reviewed to identify ongoing areas of need and action that could improve student outcomes. As part of these ongoing reviews, they were looking at both the regulatory framework and best practice guidelines from mission groups and other organisations. Providers specifically mentioned the OfS Conditions of Registration, QAA's UK Quality Code and revised requirements brought in by the Institute for Apprenticeships and Technical Education (IfATE). No providers gave evidence of how these wider aspects of regulation had directly fed into action being decided on, but we heard discussion of how these aspects were considered as a whole to help track their performance against sector standards. Actions that could support performance in any of these areas, if needed, would likely be viewed favourably.

Although not the initial driver for choice of action, revised B3 was seen to play a part in helping providers to enhance performance and student outcomes. In particular, in the three quality and monitoring case studies, programme monitoring had been geared towards the language of revised B3

(continuation, completion and progression) to provide a regulatory-aligned view of student outcomes and associated action plans for areas falling short of specified targets. We saw evidence of B3 benchmarks and thresholds being used in internal course monitoring and reporting, while informing several KPIs as a way of monitoring practice.

B3 data was also used to make a case for student success coaches within one of the quality and monitoring case studies and was judged as strongly contributing to getting this initiative 'over the line'. We also heard and saw evidence of how B3 thresholds were now being used as part of course- and module-level quality monitoring and enhancement to establish where action may be required. This had resulted in planned changes to assessments and module credits, although these changes had yet to occur.

Across the case studies, it appeared that revised B3 may have acted to expedite change due to the regulatory implications. Providers were continually reviewing processes and assessing what action may be needed to ensure a positive student experience, but also, in turn, actions that may result in positive metrics, including those relating to B3 measures of continuation, completion and progression. We saw strong evidence from all four providers of their awareness of the implications should they not achieve B3 thresholds. Although none voiced immediate concern or felt at risk of sanctions or investigation, they were sharply focused on performance being at or above the threshold for key revised B3 metrics, which in some part was to avoid investigation and ensure compliance. Fear of sanctions was not found to be the initial motivation or driver for their choice of action, but the impact on their B3 data was considered.

This was particularly evident in the admissions case study. While B3 was not officially documented as a reason for deciding to review the admissions process, it was considered by interviewees to have provided further reason for prioritising this area, given the impact poor admissions processes could have on publicly available B3 data and the implications for the provider's reputation. Again, we saw that revised B3 may not have instigated this change, but had contributed to a sharper focus and catalyst for change due to the regulatory implications of not taking action. A general theme of compliance was found to be a strong motivation for change in the admissions case study. The key decision-maker discussed a need for their higher education policies to stand up to scrutiny and comply with various sector regulations and reference points, such as Competition and Markets Authority advice on consumer protection law[21], UK Quality Code for Higher Education[22] (2018, revised May 2023) and other B conditions of registration.

Implementation and barriers to implementation

In this section, we highlight how action was implemented and any barriers that hindered progress.

Reviewing and updating quality and monitoring processes and practice was described as a substantial task by interviewees in the three providers that had made changes in this area. Although not viewing themselves as having particularly pressing compliance issues, they had revised existing systems to more closely mirror the terminology and data required for revised B3. Across these three cases, quality and monitoring processes were described as continually being reviewed to ensure they were fit for purpose, with a regulatory compliance motivation implied within this. The introduction of revised B3 was seen as an opportune time to implement changes to enhance and expedite teaching, learning and student support practice to enhance student outcomes. Participants from these providers stated that it required internal time and resources to do this well. Reviewing and updating quality and monitoring processes was seen as an integral part of this, but often demanded higher level technical skills than had previously been required, due to collecting, processing and analysing multiple, often large, datasets and data points.

[21] assets.publishing.service.gov.uk/media/6475b2f95f7bb7000c7fa14a/Consumer_law_advice_for_higher_education_providers_.pdf
[22] www.qaa.ac.uk/docs/qaa/quality-code/revised-uk-quality-code-for-higher-education.pdf

We also heard of additional time requirements for academic staff due to enhanced programme monitoring. While there was not a direct connection made to the introduction of revised B3, there was an indirect link to an increase in academic staff workload due to changes brought about in response to revised B3.

We did hear of differences in additional workload required and associated challenges experienced in providers, based on their size. In the large provider (QI over £200m), those in quality team roles did not refer to an increased focus on data as being burdensome, as they had dedicated teams and resources for such tasks. However, a smaller specialist provider described how they lacked internal resources for data analysis and had recently needed to hire someone to focus specifically on data. While there was acceptance of this as a necessary additional resource for those in quality roles, it does suggest potential additional challenges and resourcing requirements for smaller providers where such tasks cannot be absorbed by existing staff.

In the admissions case study, challenges related to being a small, majority level 4/5 provider were reported, with only one role having overall responsibility for higher education quality monitoring and reporting. Implementation of the new admissions system also fell to this individual. Additionally, both the majority level 4/5 providers experienced further challenges – having to increase staff awareness of higher education regulation and terminology, especially as this differed substantially from further education terminology, which was more embedded due to further education provision being their main activity.

Emerging outcomes

Case studies looked to identify any early findings of how provider behavioural changes had contributed to specified outcomes that they aimed to achieve, and whether revised B3 could be seen to have supported these outcomes. It is important to note that, across all four providers, changes were in the early stages of implementation and so findings here are presented as early indicators of outcomes.

All four providers intended to use internal data and reporting measures to track and monitor success. Across all four providers, the changes were in the early stages of implementation, so hadn't yet impacted key indicators such as improved continuation and completion rates. They all intended to use internal data to track success, as this would allow them to respond in an agile way to any issues or poor performance, and it was seen to be more current. None of the four providers we spoke with intended to rely on the B3 data dashboards for monitoring, largely due to the lag in data being published.

While we did not observe the specific impact of changes on student outcomes, providers anticipated that revised quality-monitoring processes would set them on the road to longer-term enhanced student outcomes. For the three quality monitoring case studies, the intention was that changes would enable staff to better identify any subsequent actions that could be taken to enhance practice and student outcomes. Due to the stage of implementation of these changes, it was still too early for any subsequent action to have been taken and any results of this seen. However, there was evidence that changes had led to programme and module-level reflection, resulting in subsequent action plans for required intervention. For example, in one case study, we saw evidence of changes to student support as a result of identified poor continuation data, which was elicited via their new quality monitoring system. This in turn had led to plans to introduce student success coaches as a direct result of these new action planning measures. Another provider told us that revised quality and monitoring processes had highlighted a need to improve the quality of student learning resources to make them more engaging and they were investigating how best to do this.

We also heard how, in one provider, revised module monitoring processes and reporting against B3 thresholds had identified issues with some practical and written assessments, as well as module credits. Anticipated changes included reviewing the weighting and timing of assessments, by splitting large modules with multiple assessment points into smaller credit units with their own assessment. It was felt this would help students to achieve and progress through smaller units, rather than failing a larger credit

module with multiple assessments, which may impact their degree classification and ability to progress to the next academic year.

We were told by individuals in each of these three providers that revising quality monitoring and reporting processes had, to some degree, supported their identification of areas for intervention and more directed action planning. While B3 was not the only consideration in action planning, it was often fundamental in driving a focus on data and galvanising progress, due to underlying concern about regulatory consequences if thresholds and targets were not met.

In the admissions case study, we heard anecdotal evidence of the new admissions policy and practice providing a clearer, more transparent process for applications, while programme leaders anecdotally reported fewer students were dropping out as a result. As with the quality and monitoring cases, this was yet to be officially tracked using internal data. This provider's assessment of the success of the new process appeared somewhat hindered by their context as a smaller, majority-level 4/5 provider with less mature processes for tracking higher education data and activities.

In one of the quality and monitoring case studies, we found strong evidence that B3 data was used to make a case for student success coaches and helped to get this initiative 'over the line'. At this provider, the idea of student success coaches had surfaced before revised B3, but implementation had been slow due to Covid, the pivot to online teaching and learning, and other activities taking precedence. With continuation, completion and progression now a part of the regulatory framework via revised B3, this enabled a strong rationale to be created for student success coaches as part of initiatives to improve and enhance student outcomes.

In one of the quality and monitoring case studies, we found strong evidence that B3 data specifically played a role in contributing to provider-wide strategic monitoring and development of institutional KPIs. We discovered that revised B3 and the introduction of threshold measures had contributed to the specific development of new institutional KPIs, such as a new employability KPI stating it should track 1% above OfS progression thresholds by 2030.

We observed that B3 thresholds had been introduced as a measure of performance for quality monitoring purposes at course and module level within a couple of the quality and monitoring case studies. We discovered that revised B3 thresholds were being used for quality monitoring purposes and to highlight areas where targeted improvement and enhancement may be required. OfS B3 definitions were also used to ensure internal reporting aligned with B3.

In one of the quality and monitoring case study providers, we also heard how the provider had used revised B3 data in their decision to enter into a partnership for collaborative provision with another provider. It seems the decision to progress the partnership was ultimately influenced by the nature of the project itself – however, as part of due diligence, they did interrogate the external provider's B3 data. As they did not find any significant concerns in this data (which may have risked them having regulatory compliance issues) it was deemed low risk to enter into the partnership. The key contact mentioned that if there had been concerns about the provider falling significantly below thresholds, this may have changed the outcome.

Conclusions

Evidence collected from the four case study providers highlighted that improvement and enhancement of student performance, support and outcomes were a strong driver for changes in practice. While revised B3 was not viewed as a strong driver for change in any of the providers, we did see evidence that – due to being part of the regulatory landscape – revised B3 was supporting and expediting the drive for continuous improvement.

The regulatory aspect of B3 was mentioned several times as a useful tool for mobilising support and resource. Three of the four case study providers had specifically mentioned B3 in staff briefings conducted about proposed changes to quality and monitoring, as a means of gaining 'buy-in' from academics, partly due to its regulatory nature and the provider's requirement to comply. That said, one provider purposively chose not to mention B3 when making changes to quality and monitoring, perceiving that this would 'turn academics off' and instead emphasised instead the need to enhance student outcomes.

We observed several examples of B3 data prompting targeted course- and module-level action planning, in which B3 thresholds formed a central pillar. If performance was below B3 thresholds, or perceived as in danger of falling below them, course and module leaders were required to explain reasons for this performance and devise a specific and measurable action plan. We heard that action planning per se was not new, but B3 data had offered a more nuanced and targeted approach. One provider had also incorporated B3 data within their broader institutional KPIs.

Use of thresholds in both the original and revised B3 seemed an important step for providers on their journey towards being more data-informed. Our discussions with key contacts suggested that B3 thresholds had motivated further focus on data-informed monitoring and evaluation. This does not mean that providers were not using data to inform policy and practice pre-B3, but rather that B3 had further enhanced this focus. As part of this, we heard provider reflections about their need for enhanced data collection, processing and analytical capabilities (especially smaller providers). Reporting against B3 thresholds had been part of a suite of evidence that had underpinned and created a strong rationale for the recruitment of new posts focused on data processing and analysis. The journey to a more data-informed approach was ongoing in all of the case study providers.

It was too early to observe specific and measurable outcomes regarding the impact of the changes that providers discussed. However, quality monitoring and evaluation changes did seem to provide a foundation for progressing enhanced student outcomes, with revised B3 contributing to how providers were approaching changes and honing practice.

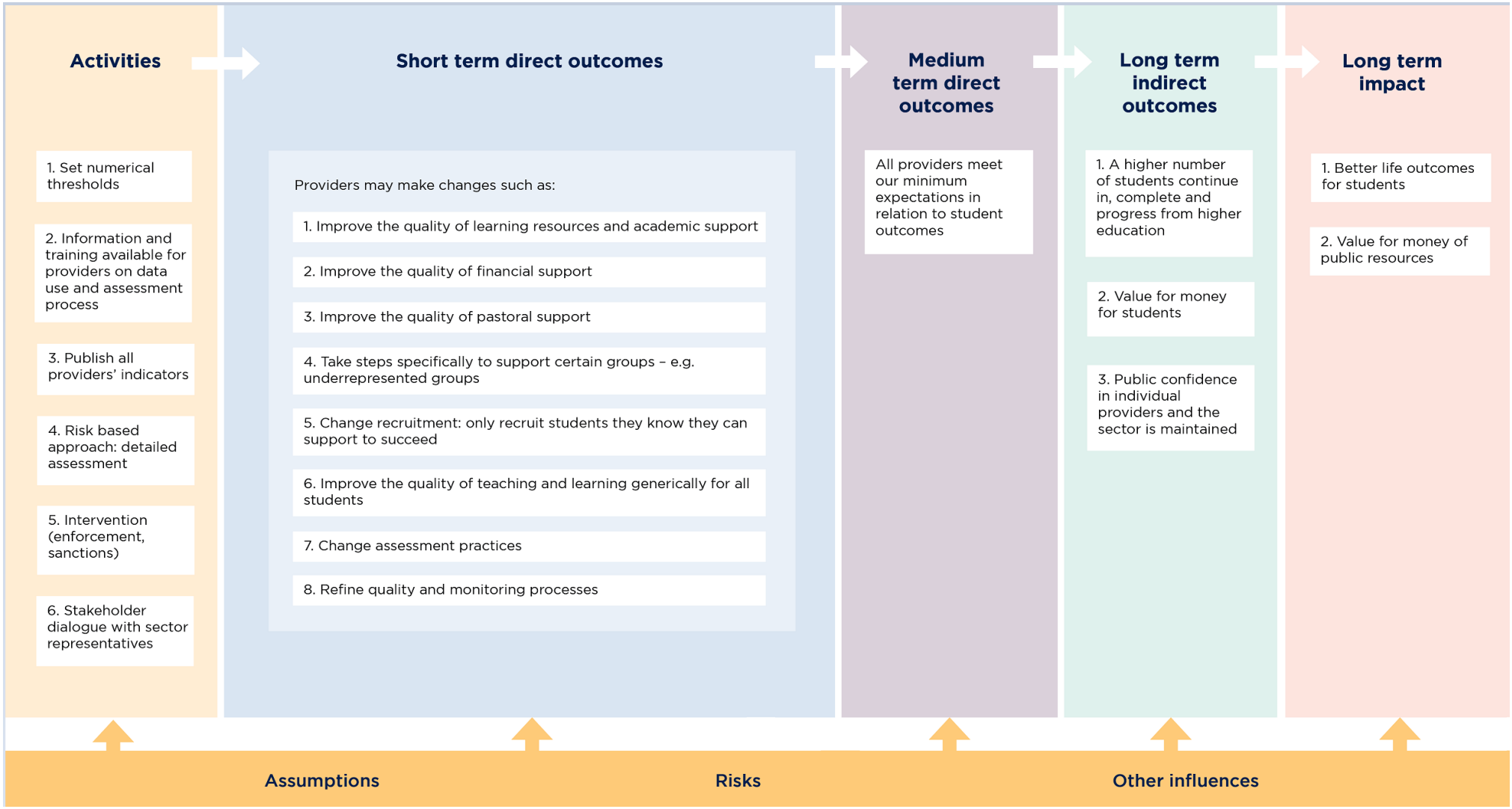
Findings from this research highlight positive work ongoing within higher education to continually enhance performance and practice, and the important role played by revised B3 in supporting these efforts.

Appendices

Appendix 1: Simplified B3 theory of change

The following is an OfS document (devised in March 2023) that is subject to further development. It helped to frame the case study evaluation research.

Figure 2: Simplified B3 theory of change



Appendix 2: Example case study evaluation matrix

We created hypotheses about what providers were doing in response to B3 and the primary motivations driving action. Alongside this, we devised an evaluation matrix tailored to each case study that used the evaluation questions as a starting point to explore these hypotheses. The evaluation matrix was used as the basis for analysis workbooks, team conversations, question guides and interviewer briefings, as well as to scaffold process-tracing (PT) tests. An example of the hypotheses and the evaluation matrix used for the quality and monitoring case studies is given below.

Example: Changes to quality and monitoring processes

Overarching hypothesis

- H1: Provider has made changes to **quality and monitoring processes** due to revised B3.

Linked hypotheses

- LHyp (sanctions): This was the result of a desire to avoid OfS sanctions.
- LHyp (demand): This was the result of anticipated changes to student demand.
- LHyp (reputation): This was the result of a fear of reputational damage.
- LHyp (enhancementB3): This was the result of a desire to improve and enhance performance and practice (linked to introduction of B3).

Alternative hypotheses

- AH1 (other OfS tools): Changes to quality and monitoring processes were due to other OfS regulatory tools (such as TEF, the A conditions for access and participation, and other B conditions).
- AH2 (enhancement): Providers have changed thinking and/or behaviour due to a desire to improve and enhance performance and practice (not linked to introduction of B3).

Null hypothesis

- NH1 (no change): No changes to quality processes occurred.

If evidence for AH1 and AH2 were not found, we would explore if changes had occurred due to other unanticipated influences and factors (AH3), but would only do this after exhausting all other alternative hypotheses. Any findings linked to AH3 would likely move into their own specific and separate hypotheses.

- AH3 (unanticipated): Changes to quality and monitoring processes were due to other unanticipated influences and factors.

Table 2: Example evaluation matrix

Evaluation questions linked to impact phase	Information required	Source of evidence	Example of evidence (related to PT test)[23] Note these are only potential EXAMPLES
Phase 1: Awareness – Provider becomes aware of data on own performance on B3 conditions			
1A. Is the provider aware of B3 – how knowledgeable are they of this condition of registration and its requirements?	Level of knowledge of B3 and requirements, as well as how these apply to their institution.	Interviews and follow-up correspondence.	Provider offers a full outline of B3, the demands on providers and how the institution is responding and ensuring compliance. (Hoop test, H1 and linked hypotheses). *Rule out all but alternative hypotheses if…*

[23] See appendix 4 for detail on PT tests

Evaluation questions linked to impact phase	Information required	Source of evidence	Example of evidence (related to PT test)[23] Note these are only potential EXAMPLES
	What was their understanding of the potential impact of not achieving B3 benchmarks?		• Provider is unaware of B3 • Provider has heard of B3 but does not demonstrate knowledge of implications for practice.
1B. What contributed to this understanding?	What communication, training or other contact with the OfS or other bodies has been received? What was the impact on understanding? What internal communication was prompted as a result of such external communication?	Interviews and follow-up correspondence.	Key informant (KI) received urgent email from vice-chancellor with dashboard, indicating that urgent action had to be taken if they were not to face sanctions (Straw in the wind, LHyp [sanctions]).
1C. What was the timing of this knowledge?	Time at which revised B3 was understood by key stakeholders.	Interviews and follow-up correspondence. Documents, e.g. meeting minutes, reports.	Provider shows that the understanding of B3 occurred prior to the last time at which changes to quality and monitoring were considered (Hoop test, H1 and linked hypotheses).
Phase 2: Problem assessment – Internal assessment by provider that identifies problems and potential actions in response			
2A. How did the provider review and use the published data to inform decision-making about where action may be required?	Whether B3 data was discussed by decision-makers with a view to determining if action was required.	Interviews and follow-up correspondence Documents, e.g. meeting minutes, reports.	Meeting minutes show revised B3 data discussed in a senior team meeting (Straw in the wind, H1 and linked hypotheses). Emergency meetings held (Straw in the wind, LHyp [sanctions]).
2B. To what extent did providers compare their data with other providers?	Whether B3 data was discussed in a comparative context by decision-makers.	As per 2A above	Provider accounts and documents show no reference to comparative data (Hoop test, disconfirmative, linked hypotheses as appropriate). Provider accounts and documents show reference to comparative data (Straw in the wind, linked hypotheses as appropriate).
2C. Did comparing data to other providers act as a motivating factor for change?	Whether B3 data was discussed in a comparative context by decision-makers with a view to determining if action was required.	As per 2A above	Provider accounts and documents show reference to comparative data in the context of a discussion of whether to respond in any way (Smoking gun, linked hypotheses as appropriate). Provider accounts and documents show reference to comparative data in the context of how likely they are to be the subject of OfS assessment and possible sanction (Smoking gun, LHyp [sanction]). Provider accounts and documents show reference to comparative data in the context of how this may impact their wider reputation (Smoking gun, LHyp [reputation]). Provider accounts and documents show reference to comparative data in the context

Evaluation questions linked to impact phase	Information required	Source of evidence	Example of evidence (related to PT test)[23] Note these are only potential EXAMPLES
			of how this may impact student demand (Smoking gun, LHyp [demand]).
2D. Did any of the other OfS regulatory tools contribute changes (other B conditions, TEF, A conditions for access and participation)?	Whether other OfS regulatory tools were discussed by decision-makers with a view to determining if action was required.	As per 2A above	Provider documentation shows reference to other OfS regulatory tools (Straw in the wind, AH1 [other OfS tools]).
Phase 3: Choice of action - Internal negotiations within provider on actions to be taken in response to identified problems			
3A. What source of information did providers consult in making decisions about which actions to take?	Whether revised B3 conditions, OfS communications or data dashboards were used in decision-making around what action to take regarding quality and monitoring processes. Whether B3 prompted consultation of broader sources of information and tools to inform decision-making around what action to take regarding quality and monitoring processes. How these things were used.	Interviews and follow-up correspondence. Documents, e.g. meeting minutes, reports.	Provider accounts and documents show reference to comparative data in the context of a discussion about which of some competing actions to take (Smoking gun, linked hypotheses as appropriate). Provider accounts and documents show reference to information about students taking notice of B3 data in the context of a discussion about which of some competing actions to take (Smoking gun, LHyp [demand]).
3B. Who did providers speak with before making decisions about which actions to change within strategy/process/practice?	Were those with knowledge of or responsibility for revised B3 conditions consulted as part of decision-making?	As per 3A above	Provider accounts and documents show consultation with student recruitment about potential impact on student numbers in the context of a discussion of which competing actions to take. (Smoking gun, LHyp [demand]).
3C. What were the key motivations for, and considerations in, prioritising and making these decisions?	Were revised B3 conditions, inputs or implications mentioned when evaluating alternative courses of action?	As per 3A above	Evidence that discussions involved fear of OfS sanctions (Smoking gun, LHyp [sanctions]). Evidence that discussions about alternative actions involved fear of anticipated changes to student demand caused by B3 (Smoking gun, LHyp [demand]).
3D. Where did the decision-making lie?	Who was involved in decisions around changing quality and monitoring processes?	As per 3A above	Evidence that broader faculty and professional services teams (specify) were involved in discussions of changes to quality and monitoring processes. (Straw in the wind, LHyp [demand]).
3E. Were there any unanticipated influences and factors which led to decisions to make changes to strategy/process or practice?	What other unexpected factors and influences were there? How did the economic situation affect provider behaviour?	As per 3A above	Specific reference to other factors. (Straw in the wind, AH2 [other factors]).

Evaluation questions linked to impact phase	Information required	Source of evidence	Example of evidence (related to PT test)[23] Note these are only potential EXAMPLES
3F. What (other) strategic responses has the provider CONSIDERED in terms of strategic direction?	Outside quality and monitoring, what changes has the provider considered? Are any of these relevant to B3? e.g. changes to its admissions approach, course portfolio/offer, student support (financial, academic, pastoral, tailored support for at-risk groups). Changes considered to utilisation and reporting of data, to enhance continuation, completion and progression metrics (including changes to condonement, degree algorithms).	As per 3A above	Not linked to PT hypotheses, but covered within research tools and reported in the individual case studies and thematic report.
Phase 4: Implementation – Provider implements actions in attempt to respond to identified problems			
4A. What changes has the provider made to its quality-monitoring processes?	Changes made to quality and monitoring processes. How data was utilised to inform these changes. Role of the data in determining the success of these activities.	Interviews and follow-up correspondence. Documents, e.g. meeting minutes, reports.	Provider outlines changes to quality and monitoring, and specifically references this was because of B3 – documentation confirms this. Evidence of changes supplied, e.g. pre-B3 and post-B3 academic regulations (noting change), committee meeting minutes signalling action, academic staff confirm that changes occurred within specified timescale.
4B. What are the emerging final outcomes within these providers (positive, negative, intended, unintended)?	Is there evidence of changes to continuation, differential degree outcomes, professional employment or postgraduate study, or to any forward indicators of this, e.g. module outcomes, employment intentions, attendance?	As per 4A above Also any aggregated data on these areas, perhaps particularly at module level.	Not linked to PT hypotheses, but covered within research tools and reported in the individual case studies and thematic report.
4C. What (other) strategic responses has the provider TAKEN in terms of strategic direction?	Outside quality and monitoring, what changes has the provider taken? e.g. changes to its admissions approach, course portfolio/offer, student support (financial, academic, pastoral, tailored support for at-risk groups). What did this look like?	As per 4A above	Not linked to process-tracing hypotheses, but covered within research tools and reported in the individual case studies and thematic report.

Evaluation questions linked to impact phase	Information required	Source of evidence	Example of evidence (related to PT test)[23] Note these are only potential EXAMPLES
Other evaluation questions not linked to phase: 1. Regulatory burden			
What burden have providers experienced (i.e. any additional work or resources required in order to manage compliance in relation to condition B3)?	Estimates of hours spent. Details of new hires or resourcing.	Interviews, evidence of new hiring, change to job descriptions.	Not linked to process-tracing hypotheses – could potentially become one if seen to be a barrier to intention.
What is the perceived burden of the implementation of revised B3 for providers?	Provider views on perceived burden.	Interviews.	Not linked to process-tracing hypotheses, but covered within research tools and reported in the individual case studies and thematic report.
What suggestions do providers have for how burden could have been reduced in the past, and how burden could be reduced in future?	Provider suggestions.	Interviews.	Not linked to process-tracing hypotheses, but covered within research tools and reported in the individual case studies and thematic report.
Other evaluation questions not linked to phase: 2. Context			
Do intermediate outcomes vary for providers in different contexts?	Information on provider context, e.g. student types, demographics, institution size, specialism, nearness to B3 thresholds.	OfS data, contextual interviews.	Not linked to process-tracing hypotheses, but covered within research tools and reported in the individual case studies and thematic report.
What explanations are there for the differences in intermediate outcomes?	Why providers feel their particular approach has so far been effective/ineffective.	Interviews.	Not linked to process-tracing hypotheses but covered within research tools and reported in the individual case studies and thematic report.
What were the barriers to considering or implementing changes?	Barriers providers felt hindered any progress. For example, knowledge of requirements, inadequate resources, prioritization of other areas.	Interviews.	Not linked to process-tracing hypotheses but covered within research tools and reported in the individual case studies and thematic report.

Appendix 3: Sample representation

Details of the representation of different OfS provider financial typologies[24], regions and B3 threshold groupings are given below.

Table 3: Financial typologies of providers in the case studies

Financial typology	Representation
QI over £200m	1
QI £100m-£200m	0
QI less than £100m or unknown	0
Majority level 4/5	2
Specialist: creative	0
Specialist: other	1

Table 4: Regions of providers in the case studies

Region	Representation
North West	2
Yorkshire and the Humber	1
London	1

Table 5: 2023 B3 threshold groupings of providers in the case studies

2023 B3 threshold groupings	Representation
Group 1	1
Group 2	2
Group 3	1

Participants were sampled so as to obtain representative views from different groups of providers in terms of how their data compared with the B3 thresholds (Table 5). This was done to understand whether a provider's position against the thresholds affects the degree and type of actions taken. The B3 threshold sampling groups use information on the proportion of students at a provider who are in a student group whose indicator rate is below the threshold set out for the regulation of condition B3 in any of the four years for which the indicators are calculated. The indicators are continuation, completion and progression. The student group is defined by the mode and level of study. For example, 'full-time first-degree students' is a student group.

The identification of indicators below threshold was based on those below the numerical threshold with greater than 95 per cent statistical confidence. These calculations used data from the OfS's Student Outcomes data dashboard[25] published in September 2022 (note that this has since been KPId).

The number of students in each group where the indicator value is below the threshold are added together and divided by the total number of students at the provider. More details, including definitions of the three

[24] www.officeforstudents.org.uk/publications/provider-typologies-2022
[25] www.officeforstudents.org.uk/data-and-analysis/student-outcomes-data-dashboard/data-dashboard

indicators, can be found under the ‘How is this calculated?’ section of the OfS's Key Performance Measure 1[26] webpage, which uses the same methodology.

- Group 1: any provider that has a lower proportion of students affected by indicators below thresholds than the sector average (as defined by the OfS's Key Performance Measure 1) for continuation, completion AND progression.
- Group 2: providers that have at least one outcome with a higher proportion of students affected by indicators below the threshold than the sector average, but that do not have any proportions of students below threshold greater than 30 per cent.
- Group 3: any provider that has at least one indicator where the proportion of students affected by indicators below the threshold is greater than 30 per cent.

[26] www.officeforstudents.org.uk/about/how-we-are-run/key-performance-measures/kpm-1-extent-of-poor-student-outcomes

Appendix 4: Process tracing and strength of evidence

The main purpose of process tracing (within monitoring and evaluation) is to help establish how and why a change occurred, and the motivations driving change. This was especially valuable for evaluating if observed changes across the four providers were influenced or caused by the introduction of revised B3, or were unrelated.

Before process tracing started, hypotheses about what providers may be doing in response to revised B3 were devised. As we had sampled based on an overarching hypothesis (knowing that changes had occurred or were planned), we focused on what motivated these changes – for example, was it a fear of sanctions, to ensure more students completed their studies, or to enhance practice more generally?

For each individual case study and related hypotheses being tested we devised an evaluation matrix. This matrix outlined the 'what, how and why' of the evaluation – detailing the areas to be explored, information required, scope and method, limitations and process-tracing tests.

The evaluation matrix formed the basis of a workbook where all data from interviews and document evidence reviews was recorded. We also subjected all evidence to process tracing and strength of evidence tests as part of an ongoing and iterative process.

Case studies aimed to corroborate findings from interviews with other sources and ideally different types of sources (e.g. internal documents). We used process-tracing tests (Table 6) along with a strength of evidence table (Table 7), which allowed us to triangulate data and assess how collected evidence either strengthened or weakened our existing findings, including assessing contradictory accounts and the trustworthiness of sources. In brief, the process tracing tests assess evidence as follows (further technical detail is set out in Table 6):

- **Straw in the wind**: there is neither necessary nor sufficient evidence to confirm the hypothesis. It is the weakest type of evidence, as it only suggests the hypothesis is plausible.
- **Hoop:** the evidence is necessary for the hypothesis, but it's not sufficient to confirm it. It is slightly stronger evidence, as it suggests that the hypothesis is relevant, but it's not enough to confirm it is the cause of the change being investigated.
- **Smoking gun:** there is sufficient evidence to confirm the hypothesis, but not necessary to confirm it. Evidence is strong as the evidence confirms the hypothesis, but it does not discount other causes of the change.
- **Doubly decisive:** there is both necessary and sufficient evidence to confirm the hypothesis. It is the strongest evidence and confirms the hypothesis as the cause of the change being investigated.

Table 6: Process-tracing tests for causal inference

Straw in the wind	Hoop	Smoking gun	Doubly decisive
Passing: Affirms relevance of hypothesis, but does not confirm it.	**Passing:** Affirms relevance of hypothesis, but does not confirm it.	**Passing:** Confirms hypothesis.	**Passing:** Confirms hypothesis and eliminates others.
Failing: Hypothesis is not eliminated, but is slightly weakened.	**Failing:** Eliminates hypothesis.	**Failing:** Hypothesis is not eliminated, but is somewhat weakened.	**Failing:** Eliminates hypothesis.
Implications for rival hypotheses **Passing:** *slightly* weakens them. **Failing:** *slightly* strengthens them.	**Implications for rival hypotheses** **Passing:** *somewhat* weakens them. **Failing:** *somewhat* strengthens them.	**Implications for rival hypotheses** **Passing:** *substantially* weakens them. **Failing:** *somewhat* strengthens them.	**Implications for rival hypotheses** **Passing:** *eliminates* them. **Failing:** *substantially* strengthens them.

Source: Collier, D. (2011) Understanding Process-tracing. *PS: Political Science & Politics 44 (4):823-830*

In addition, evidence had to be assessed for whether we could trust the source[27]. The evaluation corroborated findings from interviews with other sources. If evidence was not found, this could be for a number of reasons – the evaluation did not have full access to empirical data (e.g. minutes of meetings confirming discussion or action), or that we received different and contradictory accounts of the same event/action. Where possible, we specifically noted the reasons behind the strength of evidence offered. In this respect, strength of evidence is weakened or strengthened by the degree to which empirical accounts and access to supporting information and data exhibited a level of trustworthiness.

To help guide us in deciding both the theoretical value of evidence and empirical trustworthiness, we designed a strength of evidence table (Table 7). These are examples and, as with all combinations of theoretical value and empirical trustworthiness, a judgement call was required, alongside triangulation of data.

Table 7: Examples of strength of evidence

Strength of evidence	Empirical trustworthiness	Theoretical value of evidence
Weak	One uncorroborated account (stakeholder interview).	Weak theoretical value because if the activity took place or the motivation for action was confirmed, then evidence should be found.
Medium	Multiple sources, enabling some corroboration, but detail sparse. Some trustworthy internal documents (e.g. meeting minutes, strategy documents, operational plans) but this is likely to be just one or two.	Medium theoretical value because if the activity took place or the motivation for action was confirmed there would be more evidence found.
Strong	Multiple sources, including trustworthy internal documents (if found). Information supplied by multiple trustworthy sources.	Strong theoretical value because there are few plausible alternative explanations for finding the evidence.

[27] Beach, D. & Pedersen, R.B (2019) *Process Tracing Methods: Foundations and Guidelines* (*2nd ed.*), pp155-158. University of Michigan Press, Michigan