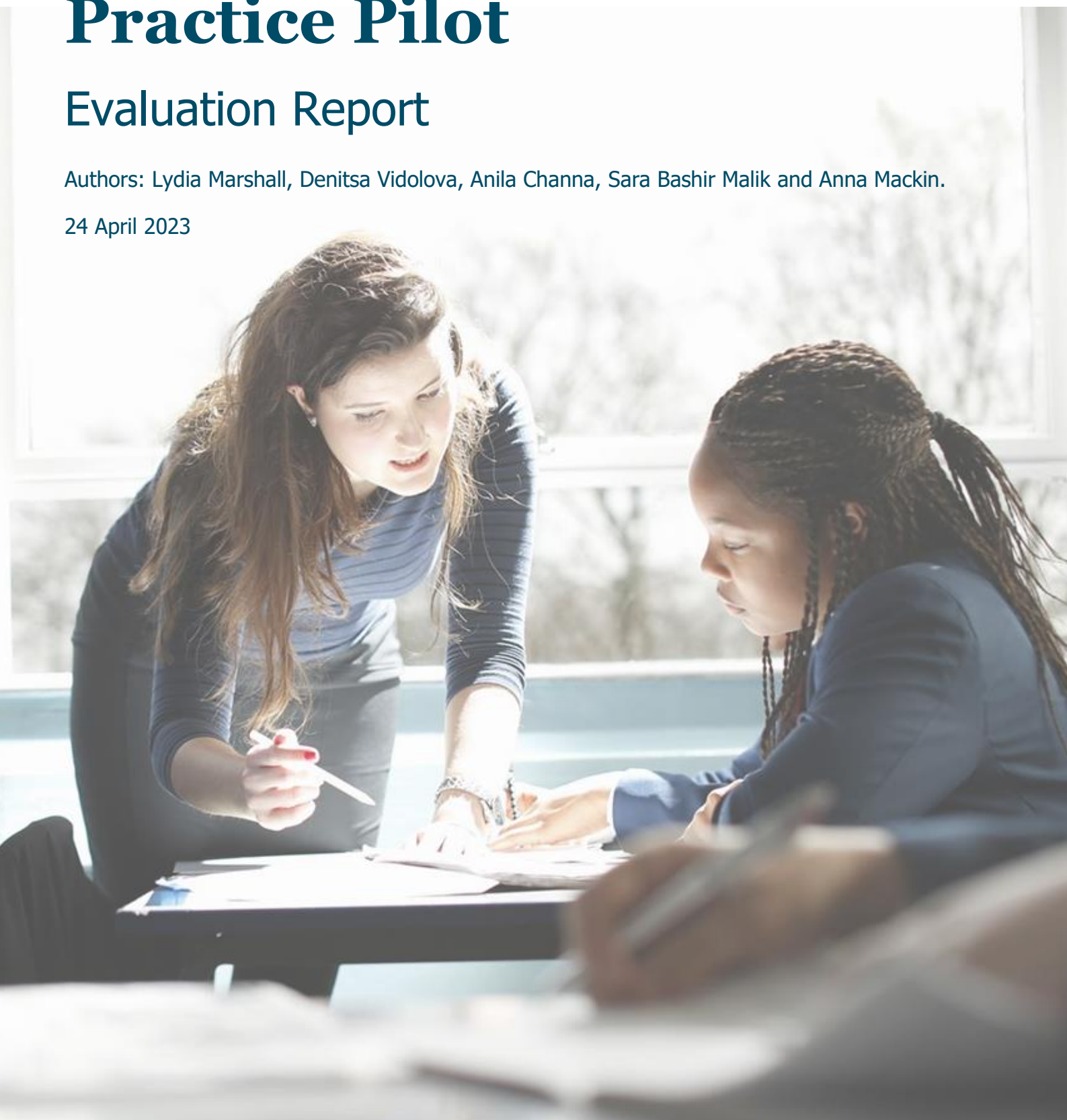


Intensive Training and Practice Pilot

Evaluation Report

Authors: Lydia Marshall, Denitsa Vidolova, Anila Channa, Sara Bashir Malik and Anna Mackin.

24 April 2023



About the National Institute of Teaching

The National Institute of Teaching boosts the quality of teaching and school leadership by carrying out research, applying the insights to its professional development programmes, and sharing findings with the sector.

Our programmes are designed by teachers and leaders who understand the joys and pressures of school life. We work in collaboration with schools across the country, ensuring that our findings benefit all schools, teachers and leaders.

About the evaluator

The project was independently evaluated by a team at Oxford MeasurEd:

- Lydia Marshall,
- Denitsa Vidolova,
- Anila Channa,
- Sara Bashir Malik and
- Anna Mackin.

The lead evaluator was Lydia Marshall.

Contact: lydia.marshall@oxfordmeasured.co.uk

Acknowledgements

We are grateful to the National Institute of Teaching for supporting this project. We are indebted to the providers, educators and trainees who generously gave their time and expertise to make this formative evaluation fruitful and rich. We would also like to thank the peer reviewers who provided useful and insightful feedback on the draft report.

Contents

Executive summary	1
1. Introduction	4
1.1 Policy context.....	4
1.2 The evidence in favour of ITAP.....	6
1.3 The pilot	7
1.4 The evaluation.....	7
1.5 Structure of this report.....	9
2. Methodology.....	10
2.1 Overview of design	10
2.2 Data collection.....	11
2.3 Analysis and reporting	11
3. Findings	13
3.1 Participants	13
3.2 Design and implementation	14
3.3 Evidence of promise.....	25
3.4 Feasibility.....	39
3.5 Scalability.....	45
4. Limitations.....	49
5. Conclusions	51
6. Considerations for the rollout of ITAP.....	53
6.1 Considerations for ITT providers.....	53
6.2 Considerations for organisations supporting ITT providers.....	54
6.3 Considerations for future evaluations of ITAP	55
References	56

List of tables

Table 1: ITAP guidelines.....	5
Table 2: Evaluation questions.....	8
Table 3: Sample achieved against target.....	13
Table 4: Provider wise pre- and post-pilot survey response rates	14
Table 5: Pilot delivery by the four providers	17
Table 6: Provider activities mapped against the five-element framework.....	18
Table 7: Model differences by provider	20
Table 8: Factors shaping six parameters of difference	22
Table 9: Model compliance against the DfE guidance	24
Table 10: Provider-specific differences to BAU	28
Table 11: Planning challenges and solutions.....	40
Table 12: Delivery challenges and solutions.....	42
Table 13: Solutions to participation challenges.....	45
Table 14: Barriers and solutions to scaling	46

List of figures

Figure 1: Evaluation overview	11
Figure 2: Design process	15
Figure 3: Five-element framework developed by NIoT and informed by Grossman (2018). 15	
Figure 4: Changes in trainees' knowledge about how and why teachers use questioning in the classroom	29
Figure 5: Changes in trainees' understanding about how to use questioning to establish what learners know	30
Figure 6: Change in trainees' understanding about how to use questioning to extend and challenge learners	30
Figure 7: Change in trainees' confidence to plan in advance how to use questioning in a lesson	31
Figure 8: Change in trainees' confidence to adapt and ask appropriate questions in response to what happens during a lesson	31
Figure 9: Post-pilot survey: trainees' enjoyment of ITAP pilot by provider	33
Figure 10: Post-pilot survey: trainees' ratings of the ITAP pilot compared to the rest of ITT by provider.....	34
Figure 11: Post-pilot survey: trainees' feedback on the relevance of ITAP pilot by provider	35
Figure 12: Post-pilot survey: trainees' feedback on the pace of ITAP pilot by provider.....	36
Figure 13: Post-pilot survey: trainees' feedback on amount of activities included in ITAP (all providers).....	38
Figure 14: Post-pilot survey: trainees' feedback on amount of activities included in ITAP (HF)	38
Figure 15: Post-pilot survey: trainees' feedback on amount of activities included in ITAP (LHU).....	39
Figure 16: Post-pilot survey: Proportion of trainees reporting challenges in participating in ITAP pilot.....	43

Executive summary



The pilot

Starting from the 2024/25 academic year, all initial teacher training (ITT) programmes in England must include an Intensive Training and Practice (ITAP) component.

The National Institute of Teaching (NioT) conceived and funded a pilot to try out different models of ITAP to inform future decision-making and delivery as the policy is rolled out. As part of this pilot, four established providers - Harris Federation, the Huddersfield Horizon SCITT, Liverpool Hope University and Star Teachers SCITT - designed and delivered models of ITAP to 318 postgraduate trainees and apprentices (referred to collectively as trainees in this report) during November and December 2022.

To support comparisons of different models and allow the evaluation to explore the implications of different design and

implementation choices, all four models focused on the topic of questioning.

The four models were developed using a participatory approach and with reference to a five-element framework developed by NioT and informed by the work of Pam Grossman. They all included a mixture of centre- and school-based activities to:

- **Introduce:** support trainees' learning about the theory of teaching and learning
- **Analyse:** support trainees to analyse expert teaching
- **Prepare:** provide opportunities for trainees to use approximations to practice and get feedback
- **Enact:** support trainees to apply their learning in the classroom
- **Assess:** monitor trainees' knowledge and skills

The evaluation

This report presents findings, conclusions and recommendations for ITT providers and organisations supporting them from a formative evaluation of this pilot. The evaluation was carried out by Oxford MeasurEd, who employed a mixed-methods approach to assess the design, promise, feasibility, and scalability of all four models. The evaluation aimed to (1) provide learning to inform high-quality design and delivery of ITAP, (2) contribute to the evidence base on ITAP and (3) offer lessons on how to best support trainee teachers during ITAP to optimise attraction into and retention in the profession.

The evaluation engaged over 250 respondents, including providers, educators, and trainees, through data collection methods, including qualitative observations, interviews, participatory workshops and quantitative surveys.

Quantitative findings should be treated as indicative only. The pilot was not designed to measure causal impact, and in the absence of a counterfactual, it is not possible to establish causality between the pilot and changes in intended outcomes. Survey response rates also varied notably between providers, and it is possible that the achieved sample may not be representative of all participants. Nevertheless, both the quantitative and qualitative research included a diversity of respondents and analysis triangulated findings across methods and participant groups to provide thematic findings and explanation

Key findings

Overall, there is reason to be optimistic about the promise of the four models and the promise of ITAP itself. All models were well-designed, demonstrated evidence of promise, appeared to be feasible, and can be considered scalable, providing a handful of barriers are proactively addressed.

- There is evidence of promise regarding acceptability and intended outcomes. Most trainees who responded to the post-pilot survey rated the models positively, finding them enjoyable and relevant. Indicative quantitative evidence also showed positive change in knowledge, understanding and confidence over time. The qualitative research found that all stakeholder groups perceived benefits for trainees to be linked to the singular focus on questioning and ITAP's marriage of theory and practice, including timely opportunities to apply learning practically. There were some concerns, however, about the models' fast pace and questions over whether all trainees benefited equally.
- The four models were well-designed and grounded in both professional expertise and evidence-based theory about teacher education. Experienced teacher educators from all four providers designed the models in a participatory manner with support from NIoT, Oxford MeasurEd (who delivered the formative evaluation) and Proxima Education (who designed "digital approximations" used in the pilot). Throughout the design process, providers referred to NIoT's five-element framework (see above), which included examples of activities that might be included under each element, and to other frameworks – for instance, to identify learning objectives. There was less evidence that the providers referred to the evidence of "what works" in training teachers when selecting specific activities for delivering their curricula, reflecting their reliance on their professional expertise as well as a limited evidence base on ITAP.
- All models were also feasible to deliver. Respondents reported planning-, delivery- and participation-related challenges that appeared to have been adequately resolved in the pilots. Three challenges to feasibility, however, may require greater focus in the future: potential overburdening of schools and mentors, added pressure on trainees, and ensuring

trainees are given sufficient support to understand how training content can be tailored to their teaching contexts.

- Scalability by repetition and by volume appears possible, albeit with additional planning and resources. Barriers such as upfront investments in design, maintaining quality as volumes increase, and the finite capacity of schools to support delivery will need concerted focus in order to be overcome. Using digital solutions as well as sharing learning and resources could facilitate successful scaling.
- Design decisions taken by providers appeared to result in trade-offs regarding promise on the one hand and challenges around feasibility and scalability on the other. For instance, models that took trainees away from their usual placement schools benefitted from a notable contrast to the “business as usual” for trainees and from providing opportunities for additional deliberate practice and exposure to expert practice. However, the decision to deliver ITAP in a different context was more resource intensive and may be challenging to implement with large cohorts or multiple times a year.

Considerations for ITT providers and organisations supporting them

Based on the learning from this pilot, we outline some key considerations for ITT providers as they prepare to design and deliver ITAP, and for organisations like NIoT who aim to support them. Section 6 of this report provides more details on these as well as considerations for future evaluations of ITAP.

Considerations for ITT providers	Considerations for organisations supporting providers
1. Plan ahead	1. Build a bank of resources to support providers
2. Proactively and carefully identify the topics that will be delivered using ITAP	2. Assist providers to share evidence and learning on implementation and effectiveness
3. Ensure ITAP provision explicitly explains the relevance of, and teaches trainees how to, adapt core content to different teaching contexts	3. Invest in building the evidence base
4. Communicate early, clearly and comprehensively to all stakeholders	4. Consult on and advocate for the right balance of guidance and flexibility from the Government
5. Avoid overburdening schools wherever possible through design choices and monitoring delivery	
6. Ensure that trainees are aware of available pastoral support	
7. Seek out and use resources available to support design thinking and delivery	

A photograph of a teacher with long blonde hair and a purple lanyard, smiling as she works with two students, a boy and a girl, at a desk. They are all looking at a book or paper on the desk. The background is a blurred classroom setting.

1. Introduction

The National Institute of Teaching (NfT) commissioned Oxford MeasurEd to conduct a formative evaluation of their pilot on intensive training and practice (ITAP) in the autumn term of 2022. In this report, we employ a learning-focused approach to present findings and recommendations from this formative evaluation to inform design and implementation of high-quality ITAP in the future.

1.1 Policy context

The impetus for the pilot evaluated in this report is an upcoming policy change related to teacher training. Starting from the 2024/25 academic year, all initial teacher training (ITT) programmes in England must include an ITAP component (DfE, 2022).

Teacher education in England has been undergoing major reform since the Carter review in 2015 (Carter, 2015). In addition to the introduction of the Early Career Framework (DfE, 2019a), these reforms have included, for instance, the rollout of the new ITT Core Content Framework (CCF) (DfE, 2019b), the push towards evidence-based curricula, and reforms to the qualifications for senior leaders, specialist teachers, and headteachers.

The introduction of ITAP is a continuation of these reforms. In July 2021, the ITT Market Review (henceforth “the Review”), which aimed to “enable the provision of consistently high-quality training, in line with the CCF, in an effective and efficient market” (ITT Market Review Expert Advisory Group, 2021: 4), recommended the introduction of a compulsory “intensive practice placement” in ITT across England. The Review argued that intensive practice would support better integration of theory and practice, consolidate trainees’ understanding of how teaching practice is

informed by research and evidence and consequently better prepare trainees to understand and apply learning material. The Government's response to the Review (DfE, 2022) accepted the recommendation to introduce intensive practice.

The Government's response to the Review defined the requirement in more detail and moved to describing it as "intensive training and practice" (DfE, 2021). Distinguished from school placements, the DfE state that a core objective of ITAP is to "provide trainees with an opportunity to intensify focus on foundational aspects of the curriculum, benefit from immediate feedback, and access appropriate expert support" (ibid.: 6).

The Review and the subsequent ITT Reform Accompanying Document (DfE, 2022) set out requirements and guidelines for the 2024/2025 academic year. The minimum time requirement for ITAP is four weeks or 20 days for postgraduate courses and six weeks or 30 days for undergraduate courses. However, the days spent on ITAP do not need to take place in one block and can be divided across the programme.

In addition to the time requirements, the Review and response set out a number of quality requirements, which are summarised in Table 1.

Table 1: ITAP guidelines

Principle	Quality requirements/guidelines	Source
1. Align to the CCF	It is expected that most of the selected ITAP topics will align with areas of the CCF. Nevertheless, providers will judge which areas will be appropriate for their training curriculum.	(DfE, 2022: 6)
2. Leverage own expertise and evidence	ITAP should be integrated into the provider's ITT curriculum that leverages that provider's expertise and draws on relevant research and evidence.	(DfE, 2021: 55)
3. Focus on pivotal topic	ITAP should cover three to five pivotal or foundational aspects of the planned curriculum.	(DfE, 2022: 66)
4. Deploy expert support	ITAP is expected to deploy experts with deep knowledge and expertise to lead and support delivery.	(DfE, 2021: 55)
5. Be distinct from school placement but include classroom exposure	ITAP can take place in a training centre or virtually but must also include substantial exposure to classroom practice that is distinct from the standard school placement.	(DfE, 2021: 55)
6. Allow observance of expert practice	ITAP should use a variety of methods, including for instance, observations and deconstructions to allow trainees to reflect on expert practice.	(DfE, 2021: 55)
7. Allow opportunity to apply theory	ITAP should allow trainees the opportunity to apply their learning, for example, through approximations or live practice.	(DfE, 2021: 55)
8. Provide feedback	ITAP should provide opportunities for immediate and targeted feedback.	(DfE, 2022: 6)

1.2 The evidence in favour of ITAP

There is a growing but currently modest body of theory and evidence on practice-based teacher education (PBTE)¹.

The ITT Market Review (ITT Market Review Expert Advisory Group, 2021) drew on a limited number of sources to evidence the importance of practice and the integration of theory and practice into ITT. In particular, the Review referenced the Deans for Impact's "Building Blocks" model (2017) and the "Practice with Purpose" report (2016) to argue for:

- the importance of a carefully sequenced training curricula with gradual increase in the complexity of goals, with knowledge and competencies that are pinpointed and sequenced
- the value of approximations and low-stakes practice

The Review also drew on Perry et al.'s (2019) review of the literature on ITT curricula and examples of the close integration of university and school experiences in ITT programmes in the Netherlands (Burn and Mutton, 2015).

Support for the ITAP concept can also be derived from the Carter Review (Carter, 2015). This 2015 review cited the work of Burn and Mutton (2013) and Sahlberg et al. (2014) to advocate for the application of models of ("research-informed") "clinical practice" into teacher training. As explained by Burn and Mutton, "research-informed clinical practice" can and has been used to refer to a host of different interventions but can be defined "by the intention to facilitate and deepen the interplay between the different kinds of knowledge that are generated and validated within the different contexts of school and university" (Burn and Mutton, 2013: 1). They use it to refer to an approach that seeks to integrate practical engagement in a school setting with research-informed knowledge in an intentional way to prepare trainee teachers for the classroom experience. Evidence from the United States (US) suggests that "clinical practice" can improve teaching practice, teacher retention and student outcomes, although here "clinical practice" is defined more broadly as "field experience" (NCATE, 2010).

In the last 15 years, there has been growing scholarship, mainly from the US, providing theoretical and case-study research on effectively integrating theory and practice in teacher education. A seminal piece of work in this area is put forward by Grossman (2018), who articulates a framework for responsive teaching that comprises:

1. **Representation:** making practice visible to trainees
2. **Decomposition:** breaking down a practice to its constituent elements
3. **Approximations:** providing practice opportunities for trainees that are proximal to the profession

There is also a small field of quantitative evidence on the impact of PBTE on teachers and their learners. According to Sims et al. (2021, cited in Sims, 2021), evidence in favour of the reform can also be found in recent research on teacher professional development (PD). A meta-analysis of 104 experimental trials showed that PD that includes intensive practice² has a more positive impact on student test scores than PD without intensive practice³ (ibid.). Cohen et al. (2020) provide

¹ Here we define PBTE as attempts to deliver teacher education in a way that prepares trainees for the practicalities of teaching and focuses on concrete elements of teaching practice, often through linking pedagogical theory and these practices throughout deliberate practice. Grounded in research at the University of Michigan School of Education, Teaching Works refer to "practicing practice" and define PBTE as training where "Novice teachers learn high-leverage practices through carefully sequenced classroom and field-based learning experiences." (Teaching Works, no date).

² Intensive practice here is understood as practice that includes the following elements: isolating a specific skill to work on, mentors modelling the skill to the teacher, the teacher rehearsing the skill, and the mentor giving feedback based on observing the practice. This is consistent with elements 3, 4, 6, 7 and 8 in Table 2.

³ Specifically, the meta-analysis found that "The average impact for 'PD including intensive practice' is .10 ($p=.03$); the average impact of 'All PD' is .05 ($p<.01$); and the average impact for 'PD without intensive practice' is .02 ($p=.18$)." (p. 7)

evidence for the value of deliberate and guided practice, finding that providing coaching between mixed-reality simulated practice sessions in teacher education courses led to more rapid development of skills and changes in attitudes.

More evidence is required to understand the impact of ITAP in ITT, and to understand the particular elements and approaches that hold best promise.

1.3 The pilot

NIoT initiated a pilot to try out a range of ITAP models and learn lessons to support effective design and delivery of ITAP in the sector. The pilot included four different models that all focused on questioning, which NIoT identified as a foundational aspect of ITT. The shared topic focus was intended to support comparisons of different models and allow the evaluation to explore the implications of different design and implementation choices.

Two models were developed and delivered by the NIoT and their partners Harris Federation (HF) and Star Teachers SCITT (STS). HF and STS are two large education trusts who deliver NIoT's teacher training programme, which generally leads to a Qualified Teacher Status (QTS) certificate, with an option of postgraduate certification in education (PGCE) awarded by Higher Education Institutes (HEIs).

The other two models were developed and delivered by external partners – the Huddersfield Horizon SCITT (HSS) and Liverpool Hope University (LHU). A deliberate decision was made to pilot various models to generate findings that could inform a range of providers. A full external application process was not feasible due to time constraints, therefore NIoT worked with these partners recommended by relevant sector bodies. Of the four providers included in the pilot, LHU is the only HEI in the group, but HSS is associated with a HEI for issuance of PGCE. The two external providers received £20,000 each from the NIoT to support planning, delivery, and engagement with the evaluation.

The providers delivered their ITAP models during November and December 2022 to a total of 318 postgraduate trainee teachers and apprentices enrolled in their 2022/2023 cohorts.

1.4 The evaluation

This formative evaluation was designed to contribute to the evidence base on the role of intensive practice in ITT. We aimed to investigate the design, implementation, feasibility, promise and scalability of the four delivery models and particular elements within them.

NIoT intended that findings from the evaluation would inform their and other providers' development and scale-up of ITAP within ITT. Particularly, NIoT intended for the evaluation to provide learning on different elements of intensive practice that might best support trainee teachers and how potential challenges can be overcome. They initiated the pilot and commissioned this evaluation so that findings would be available to the sector in time for learning to take place before the ITAP policy rollout.

We used a mixed-method approach to address the 13 evaluation questions presented in Table 2 below⁴. Full detail of the evaluation design is available in the evaluation protocol (NIoT, 2022) and Annex C of the appendices. Further information on our methodology is provided in Section 2.

⁴ One question was added after the evaluation protocol (NIoT, 2022) was signed off – EQ 2 above. One question – “What evidence does the evaluation find to support the models' logic models?” was dropped. This is because as the design phase progressed it was decided that the evaluation team should support providers to articulate their models against a framework developed by NIoT and against intended outcomes (see Section 3.2.1), rather than to support them to produce fully expressed logic models or Theories of Change.

Table 2: Evaluation questions

Evaluation domain	Evaluation question
Design & implementation	1. What four models were developed to deliver ITAP?
	2. How were the four models developed?
	3. To what extent were the four models delivered as intended/as appropriate?
Evidence of promise	4. What theory and evidence have the models and approaches to delivery been based on?
	5. How different were the models to business as usual?
	6. How do providers, trainees and teacher educators rate the models' promise with regard to key agreed outcomes?
	7. How acceptable were the different models to trainee teachers?
	8. Are there particular features of the piloted models that show best promise?
Feasibility	9. What are the challenges involved in planning for this new element of teacher training? How did providers overcome these challenges?
	10. What are the challenges involved in successfully delivering the different models? How did teacher educators and providers overcome these challenges?
	11. What are the challenges involved in taking part in the different models? How did trainees, teacher educators and providers overcome these challenges?
Scalability	12. What are the barriers and potential solutions to taking the piloted models to scale?
	13. What is the cost of delivering the different models and how do they compare?

1.5 Structure of this report

We have structured this evaluation report as follows: In Section 2, we describe our methodology. In Section 3, we report our findings, arranging these by domain and research question⁵. To facilitate navigation and synthesis, we also summarise key results under each domain, focusing on what might be most useful to ITT providers and other stakeholders. In Section 4, we outline limitations of the research. In Section 5, we provide our conclusions and in Section 6, we outline key considerations for future ITAP delivery and development and for evaluations of ITAP.

This main evaluation report is supplemented by detailed appendices, available online at <https://niot.org.uk/research>. These appendices contain further details on the design of the four models, our methodology and its limitations, and the research tools we employed.

⁵ We address Question 13 in Annex B of the appendices due to limited comparability of cost information across the four providers.

2. Methodology



We provide a summary of our methodology here. More information can be found in the evaluation protocol (NIoT, 2022) and in Annex C of the appendices.

2.1 Overview of design

We designed our methodology to:

- be learning-focused, to help future providers of ITAP identify models, tools and approaches to adopt and how they can overcome challenges that they might face
- be responsive to the evolving designs of the four models
- minimise respondent burden
- address variability across and within the four delivery models

The design included a desk-based review, observations of ITAP delivery and primary research with three stakeholder groups:

- **providers:** core members of the design teams who were members of staff at the ITT provider organisations
- **teacher educators:** other educators involved in delivering the pilot, including tutors, subject leads, mentors and coaches
- **trainees:** teacher trainees and apprentices

Figure 1 sets out how these activities addressed the 13 evaluation questions.

Figure 1: Evaluation overview

		Literature review	Project partners			Trainees		Teacher educators	Training observations
			Design workshops	Follow-up workshop	Cost template	Pre-post surveys	Interviews	Interviews	
Design & implementation	What four models were developed to deliver ITP?								
	How were the four models developed?								
	To what extent were the four models delivered as intended/as appropriate?								
Evidence of promise	What theory and evidence have the models and approaches to delivery been based on?								
	How different were the models to business as usual?								
	How acceptable were the different models to trainee teachers?								
	How do providers, trainees and teacher educators rate the models' promise with regard to key agreed outcomes?								
	Are there particular features of the piloted models that show best promise?								
Feasibility	What are the challenges involved in planning for this new element of teacher training? How did providers overcome these challenges?								
	What are the challenges involved in taking part in the different models? How did trainees, teacher educators and providers overcome these challenges?								
	What are the challenges involved in successfully delivering the different models? How did teacher educators and providers overcome these challenges?								
Scalability	What are the barriers and potential solutions to taking the piloted models to scale?								
	What is the cost of delivering the different models and how do they compare?								

2.2 Data collection

We employed a range of qualitative and quantitative methods, including observations, qualitative interviews, participatory workshops and quantitative surveys (see Figure 1) and engaged over 250 stakeholders in the evaluation.

We provide detail on how we collected the quantitative and qualitative data referred to in this report in Annex C of the appendices. Surveys and workshops were carried out online, while qualitative interviews and observations entailed a mix of online and in-person delivery.

The surveys of trainees were census-based: we asked the providers to share a link with all trainees who took part in the pilot before and after the end of ITAP delivery. For our interviews with trainees and educators, we sampled purposively, deliberately selecting individuals from those willing to be interviewed to ensure that we spoke to a diverse group of trainees and educators involved in the pilot. More detail on our sampling strategies and achieved samples can be found in Annex C of the appendices, and we provide response rates in Section 3.

2.3 Analysis and reporting

We used thematic analysis to analyse our qualitative data. Using pre-identified themes covered in the topic guides and themes that emerged during the data collection process, our team assembled matrices by interviewee category and data collection mode and summarised the data in these matrices. Once the data were collated, we undertook a process of coding and classification of the content into higher levels. We sought to explore and explain convergence and dissonance in implementation and in experiences between the four models. We also employed illustrative quotes to support our findings.

We carried out descriptive and longitudinal analysis using the quantitative survey data, following the analysis plan agreed with NIoT and included in the study protocol (NIoT, 2022). We report disaggregated figures by provider if there were a minimum of 20 responses for that provider, as per the protocol (*ibid.*). This means that we report disaggregated quantitative findings for trainees from HF and LHU.

To understand changes in intended outcomes over the course of the pilot, we matched trainees' responses to the pre- and post-pilot surveys. We were able to match 83% (115 out of 138) of the responses to the post-pilot survey to the pre-pilot survey. For these trainees, we used the sign test, a simple non-parametric statistical test, to examine if the median values of outcomes in the pre- and post-pilot surveys were statistically different from one another. We also produced Sankey diagrams to depict changes over time for the matched samples.

We also used cross-tabulations and chi-square tests to check for differences between trainees who were and were not included in the matched analyses and to examine whether experiences and outcomes differed significantly between sub-groups of trainees (for more detail, please refer to Annex C of the appendices and evaluation protocol: NIoT, 2022).

3. Findings

3.1 Participants

There were generally high response rates, except for our post-pilot survey with trainees. Achieved samples against targets are set out in Table 3 below, and further information can be found in Annex C of the appendices.

Table 3: Sample achieved against target

	Activity	Sampling strategy	Target n.	Achieved n.
Providers	Design workshop	One workshop with <6 stakeholders for each model	4	4
	Cost template	One template to be completed for each model	4	3
	Learning workshop	One workshop for each model, same stakeholders as design workshop	4	4
Trainees	Pre-pilot survey	Census of all trainees	318 ⁶	243
	Post-pilot survey	Census of all trainees	315 ⁷	138
	Qualitative interview	Purposive sample of six trainees per provider	24	27
Teacher educators	Qualitative interview	Purposive sample of six educators per provider	24	24

⁶ Total number of trainees providers told us were eligible for the pilot (their current cohort size).

⁷ Total number of trainees providers told us took part in the pilot (only HHS and LHU confirmed absences).

We met the sampling targets for all qualitative research activities. We did not set targets for the observations. We observed a total of 45 sessions, including the core training days for all four models and at least some school-based activities for each.

We achieved a 76% response rate for the pre-pilot survey of trainees (n=243) and a 43% response rate for the post-pilot survey (n=138). As shown in Table 4 below, achievement against targets differed dramatically by provider, with post-pilot survey response rates ranging from 9% (for SI) to 76% (HHS).

Table 4: Provider wise pre- and post-pilot survey response rates

	HF		HHS		LHU		STS		Total	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Eligible	140		21		30		127		318	
Completed survey	112	90	12	16	24	21	94	11	243	138
% of eligible	80%	64%	57%	76%	80%	70%	74%	9%	76%	43%
% of total sample	46%	65%	5%	12%	10%	15%	39%	8%	100%	100%

3.2 Design and implementation

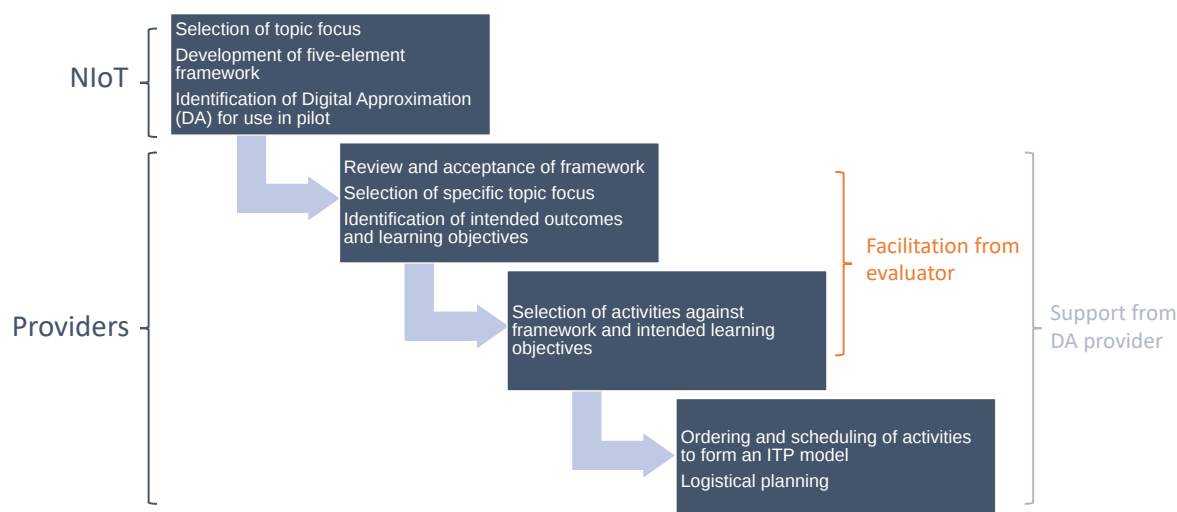
Key findings: solid design with implementation as planned across all four providers

1. Harris Federation, the Huddersfield Horizon SCITT, Liverpool Hope University and Star Teachers SCITT designed and piloted four new models of ITAP with 318 teacher trainees in 2022.
2. Each pilot was well-designed and employed activities to address five core elements of ITAP set out in a framework developed by NIoT.
3. The models differed from one another on six parameters: rhythm of work, location, use of guest educators, use of digital approximations, delivery mode, and incorporation of phase/subject.
4. Factors that drove design decisions included the resources available and practical considerations, priorities and intended learning outcomes for the pilot and providers' existing approaches.
5. Delivery of all models was as intended - we observed only minor tweaks during the delivery period except in the case of the Star Teachers SCITT pilot which changed its delivery mode due to staff absence.

3.2.1 How were the four models developed?

Figure 2 summarises the design process that occurred for the pilot. We explain this further below.

Figure 2: Design process



NIOT initial decision-making

As explained in Section 1.2 of this report, NIoT selected the topic for the pilot. They also developed a framework informed by the work of Grossman (2018, see also Teaching Works, no date) to support the four providers in designing ITAP models to pilot. The framework aims to support providers in meeting ITAP guidance and designing models that bridge pedagogical theory and teaching practice. It articulates five elements of ITAP:

- **Introduce:** activities to support trainees' learning about the theory of teaching and learning
- **Analyse:** activities to support trainees to analyse expert teaching
- **Prepare:** activities to provide opportunities for trainees to use approximations to practice and get feedback
- **Enact:** activities to support trainees to apply their learning in the classroom
- **Assess:** activities to monitor trainees' knowledge and skills

NIoT provided the framework in the format shown in Figure 3. They provided examples of the types of activities that might be included under each of the five elements, but these were not intended to be exhaustive. Instead, NIoT expected that providers would select activities based on the topic, intended outcomes, practical considerations and knowledge of their cohorts. NIoT also emphasised that the framework was not intended to represent five days or even to be understood chronologically. Rather, they expected the five elements to be iterative and reinforce one another to support the link between theory and practice.

Figure 3 Five-element framework developed by NIoT and informed by Grossman (2018)

INTRODUCE	ANALYSE	PREPARE	ENACT	ASSESS
<i>Learning about the theory of teaching and learning.</i>	<i>Using representations to analyse expert teaching.</i>	<i>Using approximations to practice and get feedback.</i>	<i>Receiving support to apply learning in the classroom.</i>	<i>Tracking trainees' growing knowledge and skills.</i>
e.g. Lectures	Lesson observations	Instructional rehearsals	Instructional coaching	Lesson observations
Seminars	Video deconstructions	Scenario planning	Team teaching	Classroom artefacts
Assigned readings	Classroom artefacts	Role plays	Small group teaching	Quizzes
Podcasts	Lesson transcripts	Case studies	Co-planning	Portfolios
Interviews	Expert modelling	Digital approximations	Lesson study	Approximations

NIoT also identified “digital approximations” that could be used as an activity under the *prepare* element (see Box 1). They offered these to their partner organisations (HF and STS) and not to the two external providers (HHS and LHU), aiming for a range of different approaches to be piloted across the project.

Box 1

By “digital approximations” we refer to initiatives that make use of digital technology to provide a simulated opportunity for trainees to practice applying what they have learned in a simulated environment.

The digital approximations used in the pilot were provided by Proxima Education. They were designed to provide opportunities for trainees to practise, reflect on and receive feedback on a range of different questioning strategies. For example, one approximation used an “action maze” to require trainees to pose questions to a group of virtual pupils to inform a pedagogical decision about whether to re-teach a topic or move on to new material. A second approximation enabled trainees to practise asking different types of questions to extend pupil understanding (e.g. probing questions or questions designed to support peer dialogue), before providing opportunities for assessment and peer feedback.

Provider decision-making

The four providers led on the design of the four models, selecting the activities that they would use and how they would combine them into a cohesive ITAP model. For each provider, a core team of up to three individuals led both on model development and delivery. We observed all core design teams to consist of highly motivated experts with extensive experience in designing teacher training programmes. These teams received support during the design phase from the evaluator (Oxford MeasurEd). NIoT’s partner organisations HF and STS also received support from the digital approximation provider (Proxima Education) and NIoT.

The evaluation team supported the partners to review the five-element framework (Figure 3) in a design workshop. Box 2 describes what happened in these workshops.

Box 2

In the design workshops we:

- **Re-introduced the five-element framework.** All providers found that the framework resonated with their understanding and intentions for ITP and provided a useful prompt to start thinking about different types of activities to include. We therefore used the framework as a facilitation tool for later activities.
- **Supported providers to articulate the intended outcomes of the pilot.** Providers articulated high-level outcomes for the pilot (knowledge, understanding, confidence and behaviour with regard to questioning practice) as well as the learning objectives from their respective curricula that they would refer to in planning.
- **Supported providers to (begin to) select the activities that they would use in their model.** We used the five-element framework and providers’ intended outcomes as prompts for this. Providers referred to the examples in the framework and the existing activities they use in their ITT provision.
- **Supported providers to think about how they would order these activities.** For providers who already knew which activities they planned to use, we supported them to think about how and why they would order activities based on their understanding of how best to support trainees’ learning as well as practical considerations.

The core provider teams had initiated their thinking about the pilot before the design workshops. Some came to the workshops with clear plans for what they would do and focused on articulating this, whereas others used the workshop as a more exploratory space. All had conversations about the points covered in Box 2 before and/or after the workshop itself. They went on to map their plans against their intended outcomes in more detail and think more about the practical considerations for their model design before finalising their detailed plans. Notably, because of the timing of the pilot, the four providers had determined their curricula and plans for the 2022/23 year and had to consider how they would fit the ITAP model into the autumn term.

NIoT and Proxima Education attended the design workshops with HF and STS and also supported their internal design discussions.

Involvement of other stakeholders

The core teams brought in their wider teams to assist in the detailed planning. This included soliciting input from phase leads or subject leads. The main stakeholders who were less involved in this initial planning process were school leaders and mentors. Mentors (referred to as coaches in the HHS model) were, however, included in the lead-up to delivery. This included providing mentor training and other communications such as emails, information posted on shared Teams platforms, ad-hoc telephone calls and in-person briefings. The training and communications covered objectives, content and logistical plans for the ITAP, and received varying levels of engagement and uptake.

3.2.2 What four models were developed?

The four models covered trainees specialising in early years, primary and secondary phases, as well as apprentices (see Table 5). The models used a combination of school and centre-based (or virtual) activities focused on the pre-decided topic of **questioning**.

Table 5: Pilot delivery by the four providers

	Harris Federation (HF)	The Huddersfield Horizon SCITT (HHS)	Liverpool Hope University (LHU)	Star Teachers SCITT (STS)
Pilot recipients	25 primary and 115 secondary trainees	21 trainees in primary and secondary	30 primary (3-7) trainees	98 trainees and 29 apprentices in early years, primary and secondary
Geographic location of pilot	London	Huddersfield and surrounding areas	Liverpool	Virtual plus Northwest and Midlands

The four models are described in detail in Annex A of the appendices. Below we present key dimensions of similarity and difference across the four models.

Similarities in the models

All providers designed models that relied on a mix of lectures and assigned readings to introduce pedagogical theory and/or research around questioning, a blend of observations and lesson deconstructions to understand the practice of questioning, and a variety of practical applications including small group teaching and live approximation/deliberate practice⁸. All providers also relied on experts to support practice and encouraged self-reflection to assess performance.

⁸ Deliberate practice refers to highly structured activities isolating specific teaching practices and offering trainees opportunities to rehearse and improve these practices through targeted feedback (for more information, see Farndon (no date)).

These similarities reflect the use of the framework developed by NIoT (Figure 3) as well as a shared understanding of ITAP as intending to explicitly and promptly link theory and practice. Table 6 summarises the specific activities adopted under each category of the five-element framework by each provider. More information on each model and how these activities were ordered and organised is provided in Annex A of the appendices.

Table 6: Provider activities mapped against the five-element framework

	HF	HHS	LHU	STS
Introduce	Lecture, assigned readings, expert modelling	Introduction of core questioning concepts in earlier PCE training days on classroom talk, lecture, assigned readings	Lecture, seminar, assigned readings	Lecture, assigned readings, practice workshops
Analyse	Lesson observations, video deconstruction, expert modelling	Lesson observations, video deconstruction, lesson deconstruction with mentor, essay reflection	Lesson observations, lesson deconstruction with mentor, expert modelling in simulation suite	Lesson observations, lesson deconstruction

	HF	HHS	LHU	STS
Prepare	Deliberate practice of layering and role playing, instructional rehearsal with mentor, digital approximation	Live approximation and lesson planning in small groups	Co-planning with mentor, instructional rehearsal with small group	Digital approximation, lesson planning, role play
Enact	Instructional coaching, co-planning with mentor, team teaching, lesson observation and feedback	Lesson planning, small group teaching, peer lesson observation and feedback	Lesson observation, small group teaching, subject-specific planning, trainee planning	Instructional coaching, team teaching, small group teaching, solo teaching, lesson observation
Assess	Digital approximation, lesson observation, quizzes, evidence bundles, mentor meeting, self-reflection	Lesson observation, group reflection, self-reflection	Lesson observation, quizzes, assessment, self-reflection and target setting	Portfolio, quizzes, assessment against weekly targets

Differences in the models

While the theoretical five-element framework allows us to easily spot the broad similarities between the pilots, deeper qualitative analysis suggests that the designed pilots also differed on six distinct parameters. We explain these parameters below and summarise them in Table 7. We then consider the factors that explain why these aspects differed across the providers.

1. **Rhythm:** All four providers designed their models to be more intensive than a usual week or “business as usual” (BAU) for the rest of their ITT programme in 2022/23⁹. Core to all four interpretations of “intensiveness” was the consistent focus on questioning across centre- and school-based activities. Despite this shared aim of increased intensity, the extent to which activities and the rhythm of the week differed from usual practice varied across the four models. For trainees of HF and STS, the week followed largely the same pattern as usual, albeit with a focus on questioning. For trainees of HHS and LHU, the pilot followed a uniquely different rhythm to BAU with changes in activities and pace. Another difference was that the HHS and LHU models were delivered within a single week, whilst the STS model was spread over six school days – Friday to Friday and the HF model had centre-based activities spread over three weekly core training days (scheduled as per BAU) and placement school-based activities between the second and third training day activities.

⁹ Or that term.

2. **Location:** All four models involved activities delivered to an entire cohort¹⁰ in a core training location in addition to school-based activities. The element of difference was whether activities took place in the locations that trainees attended in a usual week or BAU for the rest of their ITT programme in 2022/23¹¹, such as in the case of the HF and STS pilots or in different centres and/or schools used specifically for the piloted model, such as in the case of HHS and LHU.
3. **Use of digital approximations:** All models incorporated the use of approximations to provide opportunities for trainees to practice questioning techniques and get feedback outside of a classroom environment. While all four models included live in-person approximations (e.g. educator-trainee and peer-to-peer deliberate practice), the two providers delivering to larger cohorts, HF and STS, also included the use of digital approximations.
4. **Delivery mode:** Three models were delivered entirely in person, while STS had a hybrid approach. This hybrid approach entailed core introductory input and digital approximations being delivered remotely while school-based activities were carried out in person.
5. **Subject/phase specificity:** The final parameter of difference between the four models was the extent to which content and activities were phase- and/or subject-specific. The LHU model was entirely phase- and subject-specific. This model was developed for and delivered to a primary (3-7) cohort and focused specifically on questioning for phonics and early reading. At the other end of the spectrum, the STS model had core training that provided content intended to be applicable to all, with subject- and phase-specific input being delivered by trainees' mentors at their placement schools. HF and HHS provided core training elements that introduced broader pedagogical approaches to questioning that could be applied in different phases and subjects and included dedicated phase- and subject-specific activities.
6. **Use of "guest" educators:** All providers worked with a range of educators in the delivery of their models, including the regular educators involved in their wider ITT (including senior provider staff, tutors, mentors, coaches, subject specialists and consultants) and external educators who the trainees did not usually work with. The different models used different types of external educators for different purposes.

Table 7: Model differences by provider

	HF	HHS	LHU	STS
Rhythm	As per BAU: Followed the timetable/pattern of usual training: 4 days at placement school and one day in central training per week.	Different to BAU: 2 centre-based days within the week rather than the usual one for the primary cohort. In addition, trainees taken off timetable at their placement schools to attend an intensive programme of activity at specific pilot host schools.	Different to BAU: Pilot included core training sessions that would not usually happen at this time of the year. Business as usual at this time of year involves trainees spending all week on placement.	As per BAU Followed the usual timetable/pattern of usual training: 4 days at placement school and one day in central training per week.

¹⁰ For LHU it was a specific course cohort – those taking the 2022/23 PGCE Primary (3-7) course – and not all teacher trainees enrolled at LHU.

¹¹ Or that term.

	HF	HHS	LHU	STS
Location	As per BAU: Training days took place at different academies in the HF chain and school-based activities took place in placement schools.	Different to BAU: Training took place at usual venue, but school-based days took place at pilot host schools.	Different to BAU: School-based activities took place in usual placement schools but training and reflection at the university would not usually happen.	As per BAU: Core training was online (hybrid is usual practice) and school-based activities took place in placement schools.
Use of digital approximations	Yes	No	No	Yes
Delivery mode	In person	In person	In person	Hybrid
Incorporation of phase/subject	Due to timetabling, there was a dual focus on questioning and behaviour management throughout the duration of the pilot. The second training day was phase- and subject-specific.	All activities were delivered to phase-specific cohorts, with core input on Day 1 being the same for the two groups. Primary cohort focused on English and writing. Secondary cohort had a mix of general and subject-specific activities.	The pilot was delivered to a specific primary (3-7) cohort only. All activities focussed specifically on questioning in phonics and early reading.	Tailoring came through mentoring in placement schools.
Use of "guest" educators	External provider supported delivery of the digital approximations.	Specialists from the Trust delivered core training for secondary cohort. An external speaker delivered subject-specific training content for the primary cohort. Coaches delivered school-based content at the pilot host schools – trainees worked throughout the week in small groups with a coach assigned to their group.	A "lead mentor" – an experienced teacher with expertise in phonics – modelled good questioning practice with a small group of pupils in the University's simulation suite. Trainees were then supported by their school-based mentor and LHU educators to analyse this expert practice.	External provider delivered the digital approximations and accompanying content.

Factors driving differences in models

In Table 8 we outline the key factors shaping the differences between the models and explain how these manifested for the four models.

Table 8: Factors shaping six parameters of difference

	What dictated choices	How choices manifested
Rhythm	- Focus on feasibility – particularly resources available and trainee teaching commitments - Interpretation of “intense”	All providers felt that their models were more intense than a usual week or “business as usual” (BAU) for the rest of their ITT programme in 2022/23 and were feasible to deliver in the context of the pilot. HF and STS had a strong focus on what would be feasible in the future when the ITAP requirement is rolled out. They made use of existing teaching resources (except the digital approximations and accompanying content and materials, which were new) and human resources involved in delivering their usual ITT provision and did not vary rhythm beyond intensity of focus on questioning. As well as a strong focus on future feasibility, this reflected the fact that these providers had already booked training venues, staff availability and so on for the academic year before the pilot was initiated. It was also driven by the fact that trainees with HF and STS have timetabled teaching commitments that these providers felt could not be abandoned for the pilot. HHS and LHU saw the pilot as an opportunity to try things out and explore what <i>might</i> be feasible in the future. They responded to the opportunity by providing an experience that was very different to their wider ITT provision in 2022/23, the term that the pilot took place.
Location	- Feasibility - Size of cohort - Desire to demonstrate best practice	Location was very closely linked to training rhythm (i.e. those providers who followed the rhythm of BAU used the same locations as BAU and <i>vice versa</i>) and many of the factors above shaped decision-making about location. HF and STS also had larger cohort sizes, making it difficult to change venues, particularly in the context of the pilot’s limited lead-in time. HHS and LHU, who had smaller cohorts, moved trainees to new host schools for school-based activities (HHS) and back to the university (including the university’s simulation suite) for centre-based activities (LHU). These changes were intended to enable trainees to observe specific best practice from selected practitioners or in identified contexts.

	What dictated choices	How choices manifested
Use of digital approximations	- Whether offered - Interest in education technology and proposed pedagogical benefits - Cohort size	NIoT offered the digital approximations to their partner organisations, HF and STS, and both providers decided to adopt them. Motivations for adopting the digital approximations included an interest in education technology, an understanding of the proposed pedagogical benefits of the approach and the ability of digital approximations to cater to their large cohorts. The two external providers designed their programmes using other forms of approximation. We are unable to say whether they would have adopted the digital approximations used by HF and STS if offered them.
Delivery mode	- Resources available - Standard practices	STS switched to hybrid delivery due to staff illness. This was a feasible and smooth transition due to their existing hybrid model.
Incorporation of phase/ subject-specific input	- Cohort type - Standard practices	The LHU model included only the primary (3-7) cohort. This enabled the entirety of the ITAP to be phase-specific and facilitated a subject-specific focus. Other providers followed their standard practices of incorporation.
Use of "guest" educators	- Desire to demonstrate best practice - Use of digital approximations	HHS employed an external speaker to deliver subject-specific training to demonstrate best practice in their phase-specific model. The LHU model invited a "lead mentor" to demonstrate questioning in the university's simulation suite to ensure that trainees saw strong expert practice. HF and STS worked with Proxima Education to deliver the digital approximations and accompanying content. They made use of the Proxima team's technical expertise as developers but also their pedagogical expertise and ability to articulate the "mental models" of questioning underpinning the approximations.

3.2.3 To what extent are the four models delivered as intended/as appropriate?

Delivery against intention

There were tweaks made to the four models in response to external circumstances or practical considerations during delivery, but delivery was generally as planned. The most significant change we observed was the delivery mode for STS, which switched from in-person delivery to remote delivery on the first day of their pilot due to staff illness and venue challenges. Consistent with the rest of their ITT programme in 2022/23, STS continued to deliver a hybrid model throughout the week.

Educators made several minor live adjustments to the schedule to tailor delivery to trainee needs. For instance, educators from STS introduced a recap session on the following training day, which was not originally planned. This was done in response to a need identified following professional tutor visits and observations with trainees. Similarly, educators from STS omitted a role play initially planned for after the digital approximation because the trainees were tired from the exercise. Such adjustments, akin to programme adaptation in a short pilot context, did not significantly affect the

overall shape and intent of the models. All models still included their intended activities under each of the intended categories of the five-element framework.

Delivery against guidelines

The design and implementation of the four piloted models could largely be considered compliant with the DfE guidance (as per DfE, 2022; ITT Market Review Expert Advisory Group, 2021). In the table below, we summarise the extent to which the four models can be considered to comply with each guideline explained in Table 1 in the introduction to this report.

Table 9: Model compliance against the DfE guidance

Principle	Compliance of the four models
1. Align to the CCF	Yes, with explicit references to CCF made by HHS, LHU and STS
2. Leverage own expertise and evidence	Mostly, use of high-quality design team with experience for design and delivery, but less explicit use of evidence of what works in design process
3. Focus on pivotal topic	Yes, with pre-agreed focus on questioning
4. Deploy expert support	Yes, all used internal experts and guest educators
5. Be distinct from school placement but include classroom exposure	Yes, all models created distinct approaches to existing school placements
6. Allow observance of expert practice	Yes, common use of observations and deconstructions
7. Allow opportunity to apply theory	Yes, extensive use of approximations, with digital approximations used by HF and STS
8. Provide feedback	Yes, with multiple models of immediate feedback employed

3.3 Evidence of promise

Key findings: cause for cautious optimism as the underpinning framework for the pilot was evidence-based, all four models were acceptable and perceived to have met intended outcomes and indicative quantitative evidence showed positive change in these outcomes over time.

1. All providers drew on the five-element framework developed by NIoT to design their models. This framework, based on the work of Grossman (2018), suggested activities that providers could use **to bridge pedagogical theory and teaching practice** in ITAP. The providers also drew on curricula frameworks and on their professional expertise. There was less evidence that providers referred to evidence of “what works” when selecting specific activities for their curricula. This reflects reliance on their professional expertise and a limited evidence base on ITAP.
2. Qualitative evidence suggested that the piloted models were sufficiently distinct from business as usual to have the potential to make a difference to intended outcomes, with the HHS and LHU models being most different to business as usual. However, most trainees who responded to the post-pilot survey reported that the ITAP was neither better nor worse than the rest of their teacher training so far.
3. Survey data showed an overall improvement in trainees’ self-perceived knowledge, understanding and confidence around questioning over the course of the pilot. There was no counterfactual to enable interrogation of whether this improvement was higher than would be expected in the absence of the pilot. Still, in qualitative encounters all stakeholder groups linked improvements in these intended outcomes to ITAP.
4. Stakeholders raised questions about whether ITAP would benefit some groups of trainees less than others. This included more introverted or less confident trainees and trainees with some SEND.
5. Most trainees who took part in the post-pilot survey rated the piloted models positively, finding them enjoyable, and relevant, although a sizeable number of trainees reported the pace may have been too fast.

3.3.1 What theory and evidence have the models and approaches to delivery been based on?

Theory and evidence informing the piloted models

In conceptualising the pilot, NIoT drew on the Deans for Impact's "Building Blocks" model (Deans for Impact, 2017). They also decided to focus on questioning based on the evidence that effective questioning is a core teaching skill that can enhance pupils' learning and teacher planning and delivery.

The design of the piloted models was centrally informed by Grossman's work on effectively integrating theory and practice in teacher education (Grossman 2018, see also Teaching Works, no date). As explained in Section 3.2.1 above, NIoT developed a five-element framework based on this work, and the providers developed their models in reference to that framework. All four providers found this framework useful and in line with their own principles and understanding of effective teacher training.

NIoT also referred to the research literature in emphasising the importance of using approximations (Reich, 2022; Klassen et al., 2021). They specifically referred to Klassen et al.'s (ibid.) testing of an online scenario-based learning tool as evidence for their decision to include digital approximations. Klassen et al. found some evidence that their online SBL tool supported trainee teachers' self-efficacy and sense of preparedness (ibid.).

In addition, providers' design thinking was also informed by other models and frameworks of teaching theory that they use more widely. For instance, STS and HHS referred to the CCF to articulate intended outcomes for the ITAP and identified the activities that they expected to address these. LHU similarly referred to their own curriculum, which explicitly links to the CCF. HHS respondents uniquely highlighted how Kolb's theory of adult learning (cited from Zhou and Brown, 2015) had informed their design and planning.

Beyond these frameworks, the four providers relied on their existing experience and expertise in teacher training. They drew on their understanding of what would be feasible and what they expected to be effective for their cohorts to select specific activities to address the five elements of the framework and address their intended learning outcomes, and to order these activities. Providers did not explicitly refer to evidence of what works in PBTE or reference to the evidence for the efficacy of specific activities or approaches – for example, effective ways to approximate practice or to support trainees to analyse expert practice. This could be explained by two things. First, as described above, the evidence base for what works in PBTE is limited. Second, the providers delivering the pilot are experts in ITT and may have implicitly relied on their knowledge of the theory and evidence behind adult learning.

3.3.2 How different are the models to business as usual?

For these pilots to have the potential to make a difference to trainee teachers' knowledge, understanding, confidence and behaviours, they need to be sufficiently different from business as usual (BAU). Here we consider BAU to mean the rest of the providers' ITT provision in 2022/23 and their provision before the introduction of ITAP.

Difference to BAU

Our analysis suggests that the models were distinct enough from BAU to have the potential to make a difference to intended outcomes¹².

The primary difference to BAU was the sustained and singular focus on the topic of questioning. The narrower focus on questioning was identified by trainees as a notable change from a usual week in the rest of their teacher training so far, where for example, they might cover questioning and behaviour management in training and then work on providing useful feedback to pupils with their mentor. Trainees from all providers explained how this focus helped them understand the area of questioning. For example, one trainee noted:

"...it gave me a focus, and it gave me one thing to look out for. Sometimes I find that when I go into observations, I'm just focusing on, like, the entire lesson rather than one specific skill. So, this was a really good way of doing it."

Trainees also felt the singular focus for all trainees at the same time helped build a joint sense of purpose and provided opportunities for peer learning and support. One trainee observed:

"It was also easier to maintain the core focus that week when everyone was working on the same goal which is not always possible."

A second way in which the ITAP provided during the pilot was perceived to differ from BAU was the more direct link between theory and practice. This involved:

- a greater alignment between core training days and in-school provision (implicitly linking "academic" inputs and "practical" training in school)
- more explicit links being made between theory and practice by educators
- more opportunities to observe practice that directly reflected recent theoretical inputs
- more opportunities to practice applying the theory through approximation

The immediate application of the theory being taught was noted as a big advantage by trainees who worried that their teacher training so far had not always allowed for timely real-life application, making it challenging for some to successfully transfer the theory to practice. Along a similar vein, several educators applauded the direct link between theory and practice in their interview responses, highlighting that in ITT, theory and practice were often divorced due to timing and school/mentor priorities. For example, one educator noted:

"It [the school practice] marries with what they're training and how they then put it into practice. I feel the model is more secure than what we've done previously." (Educator)

Section 3.2 of this report highlights how each pilot differed from the relevant provider's BAU in distinct ways. For those involved in the HF and STS pilots, for example, the use of digital approximations was the main divergence from BAU. For those in the HHS pilot, change was more palpable because trainees were taken out of their usual school to practice in special host schools. Likewise, for those involved in the LHU pilot, going to the campus at this time of year was unusual and distinct from BAU. Because of this, the two models delivered to larger cohorts could be considered marginally less

¹² Different does not equal impact but difference is needed for impact.

distinctive from BAU than those of the two external providers. In addition, a handful of mini-innovations related to how providers delivered their activities, summarised in **Error! Not a valid bookmark self-reference.** below, set the pilot models apart.

Table 10: Provider-specific differences to BAU

Type of difference	Nature
Model-specific differences against BAU	- HF and STS: digital approximations - HHS and LHU: change in rhythm and location
Mini-innovations against BAU	- HF: Pre-reading - HHS: Video deconstruction and observation of other subjects - STS: Guided observation form - LHU: Use of simulation suite to analyse expert practice demonstrated by lead mentor with a group of pupils, joint observation with mentor, joint lesson deconstructions

Similarities to BAU

Despite the differences noted above, all four piloted models had several similarities to BAU. First, our observations and interviewee responses suggested that the broad type of activities employed (except for digital approximations) and resources deployed were similar to those used at other times of the year or before ITAP was introduced. It was clear that some elements of PBTE (for example, deliberate practice, structured observation and analysis of best practice) were already part of all four providers' existing ITT provision before the introduction of ITAP. Moreover, educators noted that they re-used content from previous training modules on questioning rather than developing all new material from scratch – focusing on varying their mode of delivery rather than content.

Overall, there was a spectrum in providers' and educators' perceptions of how different the piloted models were to their BAU. This was driven by design choices and by the extent to which they used practice-based approaches or intensive focus periods in their wider ITT provision. Three questions emerged from the qualitative analysis of provider and educators' narratives: about the extent to which the new ITAP mandate required ITAP to be different to business as usual (a lack of clarity about the requirement), the extent to which the new ITAP mandate would result in provision that was different to business as usual (a lack of agreement on or understanding of how different this was to their BAU and how different it would be to other ITT providers' BAU) and the extent to which the ITAP delivered from 2024/25 should be different to BAU (a lack of consensus on whether change is needed, particularly where ITT is already high quality and to some extent practice-based).

3.3.3 How do providers, trainees and teacher educators rate the models' promise with regard to key agreed outcomes?

Evidence of promise

As explained in Section 3.2.1 of this report, the four providers in the pilot agreed on four intended outcomes: improved knowledge, understanding, confidence and behaviour regarding questioning practice. We explored perceived changes in these intended outcomes in our quantitative and qualitative analysis.

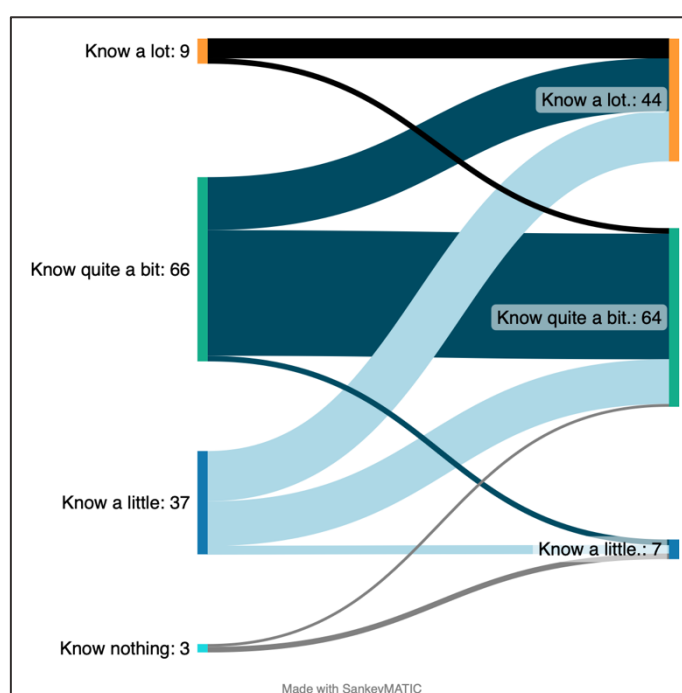
Our pre-post survey analysis shows cause for cautious optimism about changes in the intended outcomes. We matched individual trainees' answers from before and after the pilot to analyse

change in five self-reported outcomes from our pre- and post-pilot surveys of trainees¹³, all rated on a four-point scale:

- knowledge about *how and why teachers use questioning in the classroom*
- understanding about how to use questioning to *establish what learners know*
- understanding about how to use questioning to *extend and challenge* learners
- confidence to *plan in advance* how to use questioning in a lesson
- confidence to *adapt and ask appropriate questions in response* to what happens during a lesson

Figure 4 to Figure 8 below depict the change in these self-reported outcomes among our matched sample, between the pre-pilot and post-pilot surveys. These show a positive trajectory across all five indicators. For each indicator, we see that the proportion of trainees reporting higher levels of knowledge, understanding or confidence increased while the proportion of reporting lower levels decreased.

Figure 4: Changes in trainees' knowledge about how and why teachers use questioning in the classroom



¹³ The full wording of the relevant questions – and all survey questions – can be found in Annex D of the appendices.

Figure 5: Changes in trainees' understanding about how to use questioning to establish what learners know

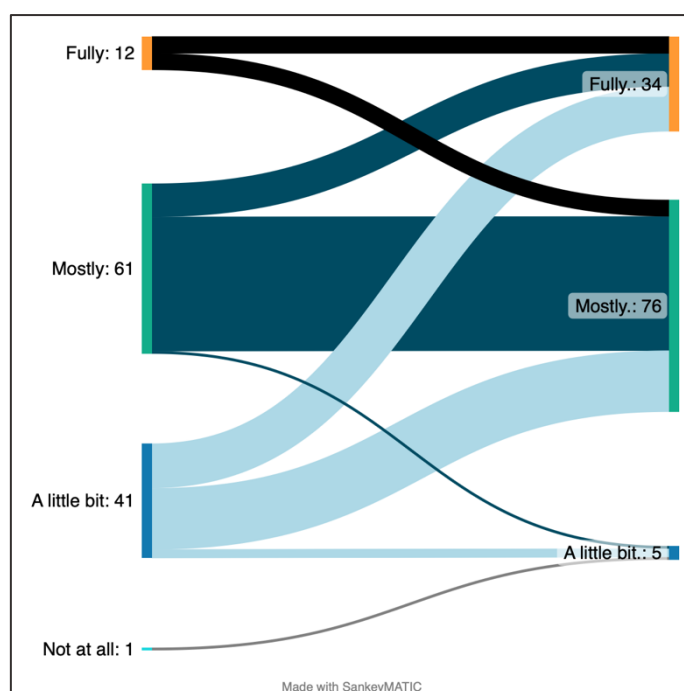


Figure 6: Change in trainees' understanding about how to use questioning to extend and challenge learners

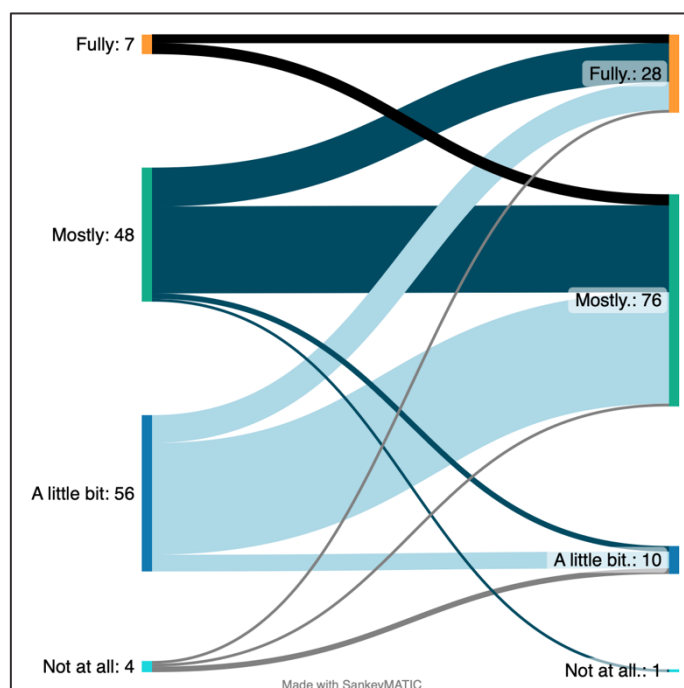


Figure 7: Change in trainees' confidence to plan in advance how to use questioning in a lesson

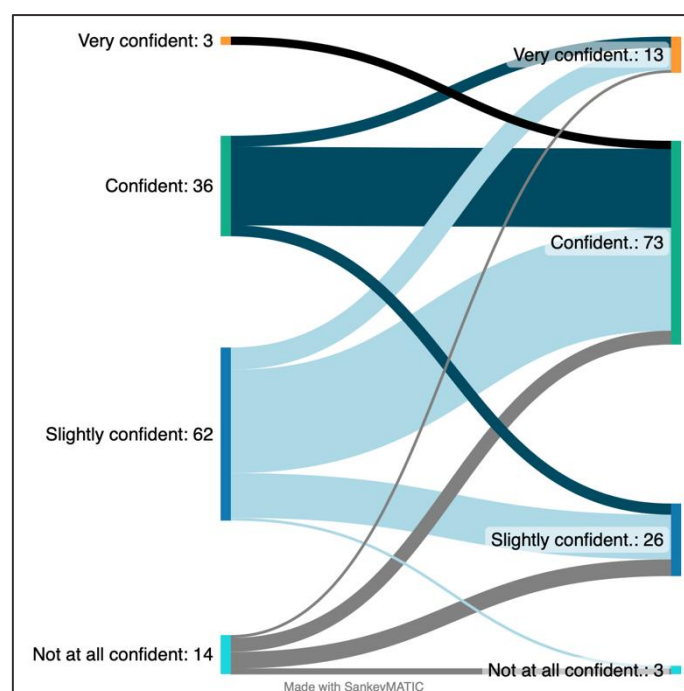
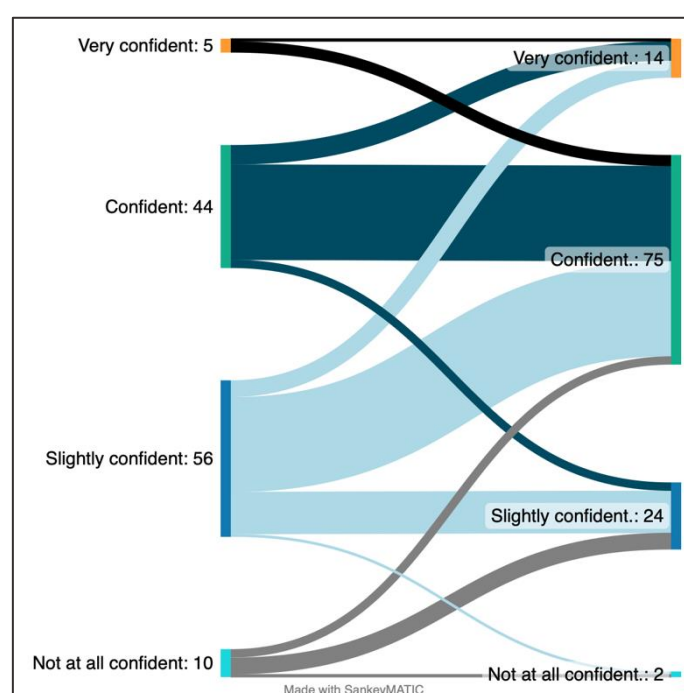


Figure 8: Change in trainees' confidence to adapt and ask appropriate questions in response to what happens during a lesson



In Annex E of the appendices, we present findings from our sign test analysis. This analysis confirmed that the positive change in each indicator from before to after the pilot was statistically significant for all outcomes and that only a very small number of participants experienced a negative change. Our further analyses show that most trainees who experienced zero change already reported high knowledge/ confidence/understanding and that positive changes were most common among trainees with the lowest self-reported outcomes at the pre-pilot survey.

Our sub-group analyses (available in Annex E of the appendices) found few notable differences in outcomes. However, we did find that:

- Secondary trainees were less likely than primary/early years trainees to have a positive trajectory regarding knowledge of questioning but reported higher levels of knowledge before the pilot began.
- Younger trainees reported most confidence in adapting questioning in response to what happened in the classroom in the post-pilot survey but were not significantly more likely to have a positive trajectory for this outcome.
- Female trainees reported less confidence in adapting questioning than male trainees and were more likely to have a positive trajectory for this outcome.

Without a comparison group, we are unable to compare this to the progress that we might have expected trainees to make in the absence of ITAP. These quantitative results should also be viewed with caution because not all trainees could be included in the matched analysis.

Nevertheless, stakeholders appeared to corroborate the quantitative findings around knowledge and understanding during our interviews and learning workshops. Educators noted that they felt trainees had gained the necessary knowledge and understanding related to questioning, including theoretical and practical aspects. They suggested that the fact that trainees could use the terminology accurately and had met their course targets through assessments supported this perception. Trainees concurred, and highlighted that their understanding of the role, purpose, and impact of questioning had deepened through the course. A few trainees elaborated that they also understood how to adapt questioning for different contexts and subjects and, importantly, how to make their teaching more child-oriented. One provider speculated on the relative impact, saying:

"With this being in its nature, more intense, you would naturally see a bigger impact with this than with your average week, probably"

Educators and trainees were positive not only about the promise of improvements in confidence and behaviour but also in wider skills. Educators reported that trainees were more likely to employ questioning earlier than before and had learned how to self-reflect on their teaching practice. Evidence from interviews with trainees pointed to increases in confidence and in self-perceived abilities to apply what they had learned at the conclusion of the course. For example, one trainee commented:

"...if anything, it just made us more confident and going back, more able [to use questioning]. I feel like having more confidence in ourselves as well and how to approach [questioning]."

Among educators, there were also additional perceived benefits for trainees in core teaching skills beyond the focus of the pilot, including lesson planning, relationship building, and child-centred teaching.

We found additional suggestive evidence that the outcomes may have gone beyond the pilot's original intent and affected teacher educators positively as well. Several educators suggested that their own knowledge and understanding of questioning had improved through the course as it gave them the opportunity to refresh and reflect on their own practice of questioning. Educators from different models also highlighted how useful they found unique aspects of the course, such as (a) access to new learning materials (HF), (b) the opportunity to engage in digital approximations (STS and HF), (c) the opportunity to deepen coaching relationships with trainees (HHS and LHU) and (d) the incorporation of additional research into applying questioning in special needs settings (STS).

Elements that could limit promise

Some respondents suggested that it might be too early to gauge whether outcomes were achieved. For a few educators, to confirm promise, it was critical to see trainees in more practical settings in the future as they argued that effective questioning will likely require sustained practice. One educator also pointed out that because cohort sizes were large, they were unable to assess every trainee's approach individually. These limitations are expected for a short-term pilot such as this. However, it should be born in mind, and concern about whether learning would be sustained was raised, as well as just a lack of evidence for this. For example, one educator noted that what had been learned could dissipate over time, and therefore follow-up training would be needed to ensure that trainees sustain the skill of questioning

One element to consider when gauging potential promise is whether all trainees and students benefited from the models in the same way. Some educators and trainees worried about whether all trainees were benefiting equally from the course. They noted natural differences in personality, such as some trainees being more extroverted or confident than others, which may affect perceptions and uptake. They also highlighted that a course designed to be intensive may have placed a potentially greater burden on trainees who have particular special educational needs or disabilities (SEND), resulting in higher levels of stress and scheduling disturbance for this select group. These challenges are discussed further in Section 3.4.3 of this report.

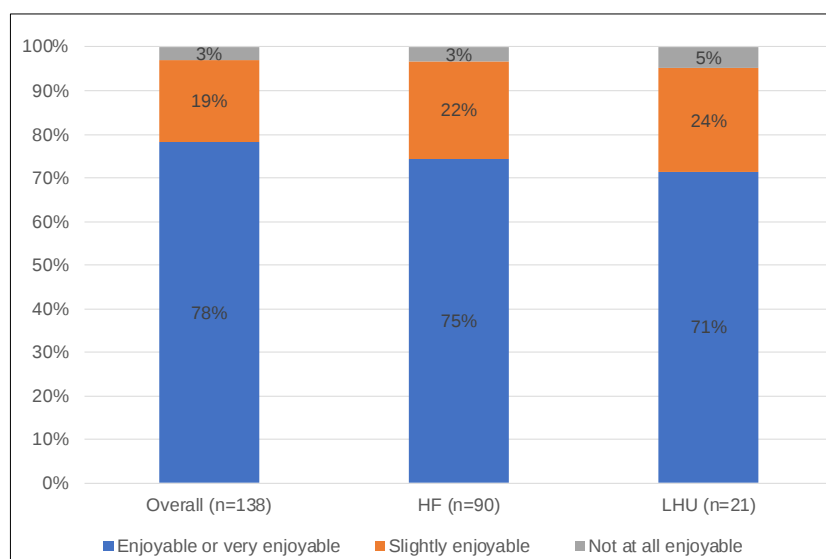
3.3.4 How acceptable are the different models to trainee teachers?

All four models seemed acceptable to trainee teachers, with many respondents rating the pilots highly on enjoyment, quality, and relevance. To gauge acceptability, we examined trainee responses from the post-pilot survey.

Enjoyment

Overall, most trainees who responded to the post-pilot survey reported positive experiences of ITAP. Just under 80% of trainees reported that the course was enjoyable or very enjoyable. Proportions did, however, vary across providers, as shown in the figure below. Enjoyment was found to be slightly lower than the average for respondents from HF and LHU at 75% and 71% respectively.

Figure 9: Post-pilot survey: trainees' enjoyment of the ITAP pilot



We found no significant differences between sub-groups of trainees regarding how enjoyable they found the ITAP.

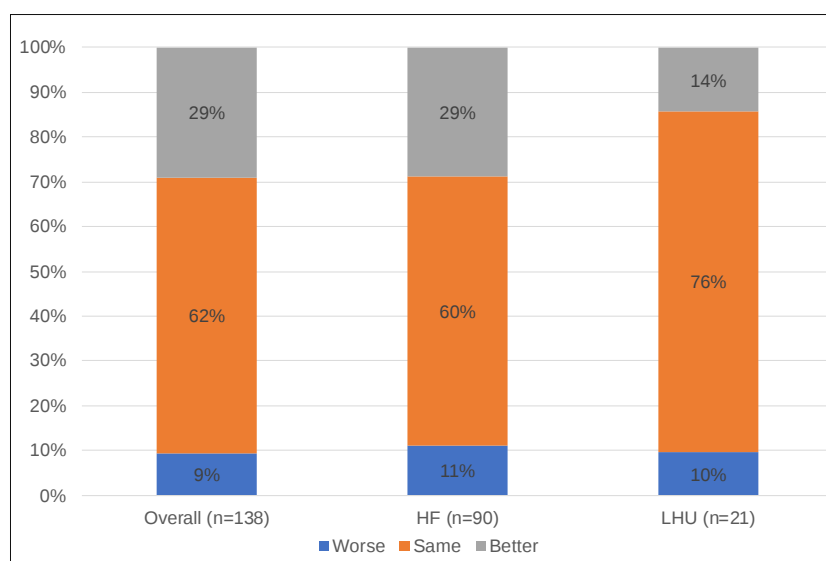
Our qualitative analysis suggests that the high levels of enjoyment were likely driven by several factors. In interviews, trainees commented on the novelty of the week, how beneficial they perceived it to be, and social benefits related to interacting with their peers and different professionals. One trainee noted:

"The week was amazing. I've said to family and friends that it was the best week in teaching so far, um, for many reasons. Just due to obviously the school, actually the fun of it as well. It was quite fun."

Comparison to the rest of ITT so far

Despite the predominantly positive experiences of the pilot, when asked to compare the course to the rest of their teacher training so far, most (62%) trainees who completed the post-pilot survey rated the ITAP as "about the same" as the rest of their ITT (Figure 10).

Figure 10: Post-pilot survey: trainees' ratings of the ITAP pilot compared to the rest of ITT



Almost three in ten (29%) of respondents rated the programme as better than the rest of their ITT, though this figure was lower (14%) among respondents who took part in the LHU model. When asked to explain their answer, the most common reasons that trainees gave for preferring the ITAP to the rest of their teacher training so far were having **opportunities to put learning into practice** (n=8), **the relevance of the questioning topic** (n=5) and **the usefulness of digital approximations** (n=4). Other reasons included the direct links between theory and practice (n=3), the narrow focus of the week (n=3), having more opportunities to observe expert practice (n=3), understanding the purpose of the activities (n=2), receiving feedback or learning from peers (n=2), receiving expert feedback (n=2), having a positive relationship with or valuing support from staff (n=2), finding the week enjoyable (n=1), liking the diversity of activity types (n=1), and valuing repetition throughout the week (n=1).

Across providers, a notable minority of around one in ten trainees who took part in the post-pilot survey rated the programme as worse than the rest of their ITT. The proportion was the same across providers. The most common reason for rating the programme as worse than the rest of their ITT, mentioned by five HF trainees, was that they found the pilot repetitive or too long to focus on one topic. Other reasons included being taken away from teaching opportunities in their placement school (n=2, both from LHU), a lack of structure or clear purpose for the pilot (n=2), a lack of feedback (n=1), finding the content less interesting than the rest of their ITT (n=1), finding that the ITAP added to their workload (n=1) and finding the digital approximations "not appropriate" or helpful (n=1).

We found no significant variation between sub-groups of trainees regarding how they rated the ITAP compared to the rest of their ITT so far.

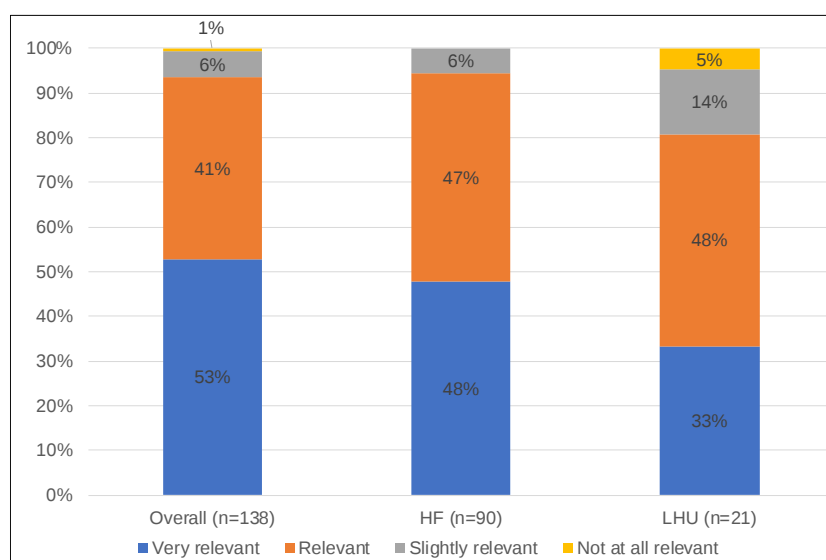
Relevance

Most trainees who took part in the post-pilot survey considered the course relevant, with 94% reporting they found it relevant or very relevant. Importantly, relevance was a strong theme across the qualitative data from all providers. Trainees often reflected positively on the usefulness of the topic and content for application in the classroom for checking pupils' understanding and depth of knowledge.

Only one trainee from LHU did not find the training relevant at all, while a further three trainees (14%) from the same provider found it only slightly relevant. This may be related to one of a couple of challenges that emerged in the qualitative research. Because different schools follow different – and very prescriptive – phonics schemes and accompanying lesson plans, some trainees in the LHU model found it difficult to envisage how they could apply the questioning techniques they were taught about in the ITAP pilot in their lessons. Another theme emerging from the qualitative interviews with LHU trainees was that they had already received a lot of phonics input in their ITT to date.

We found no significant differences between other sub-groups of trainees regarding how relevant they perceived their ITAP to be.

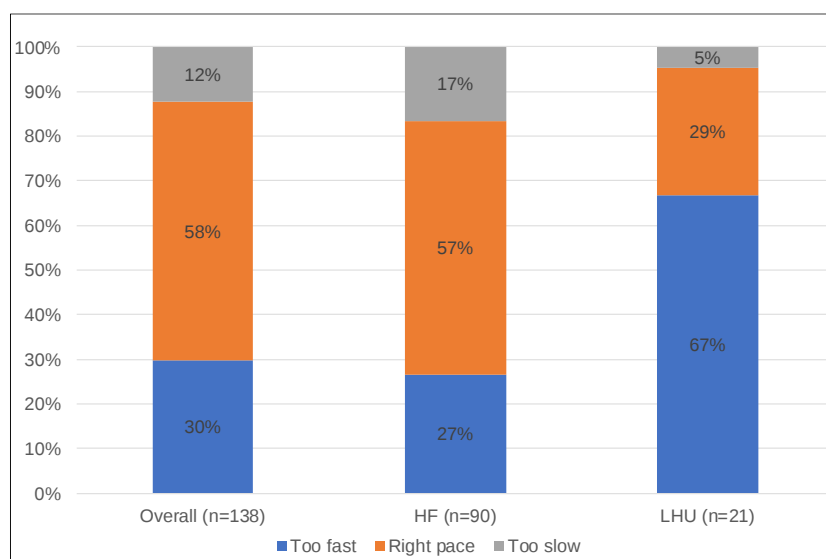
Figure 11: Post-pilot survey: trainees' feedback on the relevance of the ITAP pilot



Pace

Although most of the trainees who completed the post-pilot survey (58%) felt the pace of the course was just right, a sizeable proportion (28%) felt the pace was too fast. The concerns over pace were particularly common among LHU respondents – 67% felt the pace was too fast. As noted in Section 3.3.2 above, the LHU model involved a notable change in rhythm compared to a usual week for LHU trainees in their ITT so far. It also took place within a single week, compared to the HF model, which was spread over a longer duration. However, our qualitative analysis suggests that all trainees noticed a change in pace as was intended by the providers. Trainees reported that there was a lot of material to cover, and that that material was more in-depth than they were used to, but that they also generally considered it to be adding to their skillset.

Figure 12: Post-pilot survey: trainees' feedback on the pace of the ITAP pilot



Our further quantitative analyses (available in Annex E of the appendices) found that female trainees were significantly more likely to find the pace of the ITAP too fast and male trainees were significantly more likely to say it was too slow. Our qualitative data are unable to explain this difference and it is an interesting question for future research to explore.

3.3.5 Are there particular features of the piloted models that show best promise?

Promising elements based on model similarities

On balance, we found support for the basic premise and logic of ITAP. As discussed above, our qualitative analysis indicates that the singular focus and the weaving in of theory and practice were core to the perceived usefulness of the course.

Across the four models, the features of expert demonstration and expert feedback were viewed as beneficial for trainees' understanding and confidence. Each pilot incorporated expert demonstration and feedback in multiple ways, and in general, these features of course delivery were discussed repeatedly in interviews as positive aspects of the pilot models. In qualitative interviews, trainees commented on the variety of demonstration, from lesson observations to video deconstructions. They also appreciated the immediacy of feedback, the guidance and commitment of mentors and coaches, and the variety of feedback sources including their own peers. In our survey, 75% of trainee respondents felt the course offered the right amount of expert feedback. Our own observations corroborated these views – each provider employed experienced, high-quality course leads who expertly delivered the course while ensuring that trainees were shown good practice and offered structured feedback on their questioning practice – through approximations and/or observations of classroom teaching.

Promising elements based on model differences

The use of digital approximations in the pilot showed evidence of promise. Providers, educators, and trainees expressed positive views about the digital approximation innovation. Providers and educators commented that it was the novelty of the approach which made it engaging. Providers, educators and trainees also noted pedagogical benefits from the digital approximations relating to their deliberate and focused nature and the immediacy of opportunities to apply learning and receive feedback, as exemplified in the following two quotes from trainees:

"I feel like with all the other training that we've done, it's very lecture based, which is fine. But we don't have explicit examples where we get a chance to consider what we would do at that time and in that moment. And that's what that approximation gave me specifically. [...] Because when we have the normal trainings, we're told the theory this is how you should do something and then a question might be posed and then we'll look at each other and have a quick discussion and kind of work out. But to be given an example, a very explicit example, and then just a few options where we had to pick in that moment, this is what I would do explicitly."

"I was focusing on it a lot more as well. Instead of just listening constantly, I was actually doing activities as well and it was sort of like testing my own understanding and my own learning as well. Just like if I gave a lesson to a student and I was giving them questions to check their understanding [...] then I was confident that I could go back and do this [in the classroom] because I've done it in the lesson [...] and I know how I can improve because I've got feedback from other peers as well."

Nonetheless, we did observe some challenges with the delivery of digital approximations (discussed in Section 3.4.2 of this report) that may need to be addressed if this feature is used in the future.

We found that the six elements that differed amongst providers involved trade-offs, and that the optimal balance may need to be struck between these trade-offs to enhance outcomes. For example, a useful result of our qualitative analysis was that respondents considered "novelty" a feature of enjoyment, and that this enjoyment was often perceived to lead to higher levels of efficacy if the novelty was not considered to be too disruptive to the schedule or workload of trainees. This can be illustrated through the results of the HHS and LHU models, where respondents appeared to enjoy the course and its intensity but perceived the change in rhythm and location to be more disruptive than ideal. Likewise, for the HF and STS models, similar trade-offs are involved in offering digital approximations. Considering the trade-offs of design choices will be important for ensuring high-quality ITAP delivery.

Trainees were largely satisfied with the nature of the content presented in all four models, except for the amount of independent study. Asked about lectures and presentations, independent study, opportunities to try out questioning techniques outside of the classroom (i.e. approximation), chances to practice questioning in the classroom and expert feedback, trainees who responded to the post-pilot survey were most likely to say that they received "about the right amount" of lectures and presentations (85%), though almost one-tenth of trainees said they received too much of this – more than for any of the other activities (see Figure 13). At least three-quarters of trainees also felt that they received the right amount of opportunities for approximation (78%), classroom practice (79%) and expert feedback on their teaching (75%) (see Figure 13).

Trainees' reflections on the nature of the content in their ITAP model reflected providers' design decisions. For example, 79% of survey respondents who took part in the HF model, which included digital approximations, felt they had sufficient opportunities to practice questioning outside of the classroom (Figure 14), compared to 67% of LHU respondents (Figure 15). Conversely, 95% of survey respondents who participated in the LHU model, which had a small cohort and included lesson observations and feedback from both LHU tutors and school-based mentors, felt they received sufficient expert feedback, compared to 66% of HF respondents.

The consistent challenge appeared to be trainee perceptions of the amount of independent study – around three in ten (29%) of trainees across the providers felt this was too little or not provided at all (Figure 13). This likely reflects both a deliberate design decision by providers and the contained duration of delivery that may need to be reconsidered in the future.

Figure 13: Post-pilot survey: trainees' feedback on amount of activities included in ITAP (all providers)

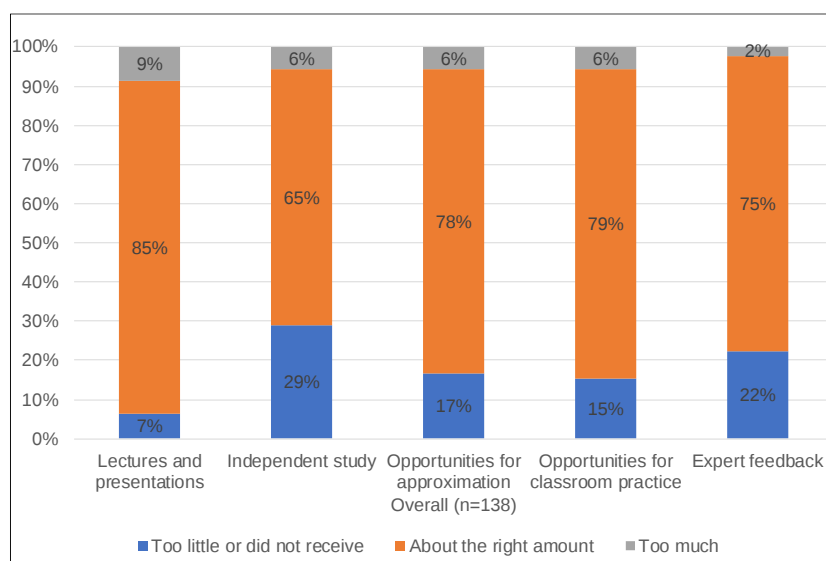


Figure 14: Post-pilot survey: trainees' feedback on amount of activities included in ITAP (HF)

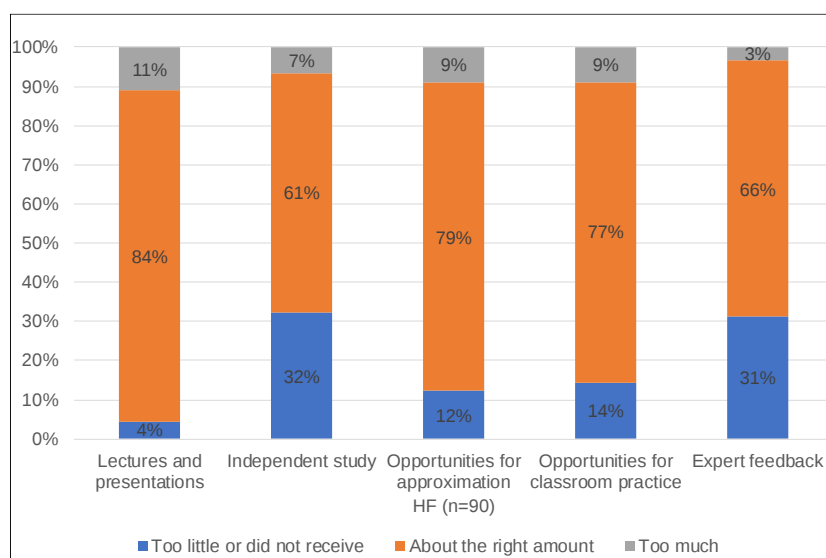
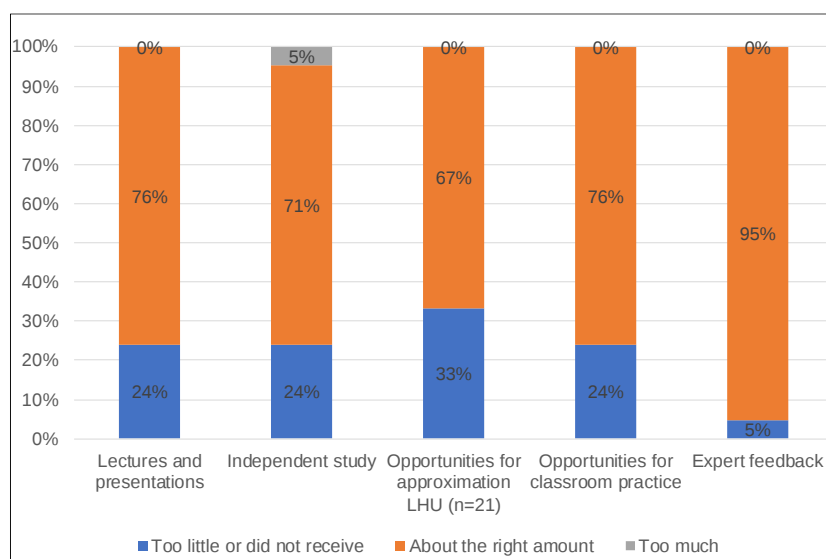


Figure 15: Post-pilot survey: trainees' feedback on amount of activities included in ITAP (LHU)



3.4 Feasibility

Key findings: *models appear to be feasible if a handful of challenges are proactively managed*

1. Ambiguity around the ITAP concept and a short time for designing were the main challenges related to planning. Earlier planning and additional support for design helped address both.
2. Burden on schools and a lack of sufficient guidance for adaptation were the main challenges in delivery, although some models saw specific challenges related to their design as well.
3. There were also potential challenges for trainees, including cognitive overload, emotional pressure and lack of clarity around how the content applied to their specific context.
4. Most challenges were overcome during the pilot and there was important learning for the future. Proactively supporting trainees will likely be vital for attraction and retention.

3.4.1 What are the challenges involved in planning for this new element of teacher training? How did providers overcome these challenges?

Challenges

Providers reported that a key challenge for planning revolved around the ambiguity of the concept of ITAP. The guidelines and remit provided by the government were considered by respondents to be broad and vague, with educators noting that this meant that they initially did not know where to start. A particular point of confusion was around the level of "intensive practice" expected, and how any model designed would be differentiated from ITT programmes that already include elements of PBTE.

Practical considerations such as a short time for planning in this pilot and the impact of introducing ITAP on the existing curriculum were also notable challenges. Across all four providers, educators reported that it took considerable energy to bring everyone in the right place to develop the right model, especially given that planning in the context of the pilot needed to be undertaken in a short period of time. A related concern was around the adjustments that were needed in the standard curriculum to allow for an ITAP pilot week, and how this may play out once it is scaled up to be delivered multiple times a year. NIoT had chosen the topic of questioning for this pilot, but questions were raised around which topics providers should and would select for future ITAP delivery, and whether and how other curriculum areas would need to be dropped as a result.

Solutions

We summarise the specific solutions employed for each planning challenge below. Our analysis suggests that the solutions adopted for planning challenges were adequate but may require further streamlining as scaling occurs.

Table 11: Planning challenges and solutions

Challenge	Solutions employed	Adequacy for future
Ambiguity of concept	NIoT framework, Oxford MeasurEd facilitated workshops, Support from digital approximation provider	Adequate with room to streamline further
Time and disruption	High quality educators, NIoT framework, Oxford MeasurEd facilitated design workshops, Support from digital provider	Adequate with room to streamline further

Centralised guidance and assistance were the main ways challenges around planning were addressed. This included NIoT designing the five-element framework, Oxford MeasurEd facilitating the design workshops and additional support from Proxima for digital approximations. Another solution arose from the quality of educators leading design and delivery for all providers. We observed that these educators were highly experienced and relied on their extensive expertise in ITT to work quickly to devise feasible, effective model solutions. Providers reflected on the following strengths in their teams that they felt contributed to successful design and delivery:

- combined experience and expertise in leadership, adult learning and phase- and subject-specific content
- leaders keeping on top of the theory of good adult education and supporting colleagues to implement this theory in practice

As other providers move to implement ITAP, a handful of factors will be different to the context of this pilot that will need to be accounted for to enhance feasibility. On the one hand, we expect that providers will have much more time for planning and communicating effectively than they did in this pilot. On the other, the additional assistance provided by NIoT, the digital approximation provider and evaluation team will not be available. As a result, other providers may want to consider whether they want to seek out additional planning support early on during their design phases, for instance, to select topics or finalise curricula. Without NIoT's coordination, providers may have to find alternative means of organising and sharing to build a robust evidence base on ITAP design and delivery.

3.4.2 What are the challenges involved in successfully delivering the different models? How did teacher educators and providers overcome these challenges?

Course level challenges

Some challenges were model-specific, relating to the design choices that had been made and activities that were used.

For the models involving digital approximations, there were challenges associated with this. Challenges included:

- adopting a new technology requiring familiarisation time and additional organisation and coordination
- lack of clarity/shared expectations among educators about who would deliver the session and how the sessions would fit into the day
- technical challenges logging in

Findings from our research with educators and providers suggest that these challenges were overcome during the pilot and were largely related to initial set-up/using a new technology for the first time. We would therefore expect them to be unlikely to be repeated for the same providers as their familiarity with the digital approximation grows. However, similar challenges may well arise for other providers adopting the digital approximations used in the pilot and/or the providers involved in the pilot if they adopt different technologies in the future.

For the two external providers, there were challenges related to their design choices around location and difference to BAU. For example, we found reports of mentors from the LHU pilot finding it difficult to manage changes in the location of their support because they had to travel to the university. Staff at HHS, meanwhile, highlighted that the success of their model depended on strong partnerships with and buy-in from the pilot host schools.

A handful of other model-specific challenges were raised by trainees that may need to be ironed out before formal launch. In the HF pilot, some trainees noted there weren't enough lessons to observe and that trainees sometimes had to observe subjects they did not teach. There were indications from the STS pilot that not all targeted observations happened within the allocated week, and some were pushed to a later date.

School level challenges

Three sets of school level challenges were raised in interviews that applied to all models and appeared to be of a serious, systemic nature.

The first was related to the additional demand being placed on mentors. Mentors generally are not paid extra for mentoring responsibilities and often have a full teaching load themselves. This meant that not all mentors attended training or engaged as fully in the pilot, limiting support for trainees, and providers suggested that this challenge might heighten as other requirements on mentors and schools increase (for example, the minimum time requirements for mentor training required by DfE ITT reform (DfE, 2022)). As will be the case for much ITT provision, trainees' experiences of the ITAP provision that took place in placement schools depended heavily on their mentor. Trainees within the same models had inconsistent experiences of the ITAP pilot due to differences in the quality of their mentor meetings and of the observation feedback that they received, for example.

The second challenge or concern was around trainees being taken away from their teaching placements for a week in the LHU and HHS models. Among educators, there were concerns about disruption to trainees' routines and to their adapting and becoming embedded in their placement school. This was particularly linked to the pilot being early in the year. For ITT courses – or at later stages of ITT – where trainees have a higher teaching load, a week away from teaching placements could also lead to concerns about impacts on pupils and/or on requiring supply cover.

The final challenge was related to school commitment. Senior staff from schools understandably had competing priorities and without a deeper understanding of the ITAP programme were less willing to accommodate shifts in schedules of trainees and mentors. Many noted that ensuring that schools and mentors were not overburdened had to start right from the planning stage, and that a lack of timely communication in the context of the pilot's short timeframe meant that commitment had the potential of being variable.

Solutions

Providers responded to delivery challenges with a varied set of solutions that are summarised in the table below. These ranged from training for mentors and coordination with schools, to assistance from providers. Our main observation related to delivery challenges that were course specific is that they appear to be the consequence of the trade-offs in design that we noted in the previous section. For example, while digital approximations were enjoyed by trainees and understood to have pedagogical benefits and potential for scale, they introduced challenges around access and

familiarity. Likewise, varying the rhythm and location of the course against usual ITT provision in 2022/23 offered trainees a unique experience that allowed them to interact with new schools, professionals, and peers, but introduced greater resource requirements and logistical challenges. Each design difference brings unique advantages and challenges that must be addressed.

Table 12: Delivery challenges and solutions

Nature of challenge	Solutions employed	Adequacy for future
Course level	Support from core design team, Support from Proxima, Sharing digital technology with peers	Generally adequate with room to streamline further Ability of course to facilitate adaptation may need greater focus in future
School level	Timely communication to schools, Training for mentors, Support from core design team	May need greater focus in future

Overall, we found that course level challenges were addressed appropriately with some room to refine these further in the future, but that school level challenges may need deeper engagement by all providers. School level challenges were largely related to capacity and engagement and therefore thinking through more effective ways of communicating and incentivising schools may be important. We noted indicative evidence for instance that communicating with schools verbally was more effective than just communicating in a written format. Examples included in-person and virtual briefings for mentors in advance of the training, inviting mentors to attend the core training day and making phone calls to support mentors during the school-based activities. Simple solutions such as these may alleviate delivery challenges in schools to a certain extent. What would still need to be considered is potential costs of arranging teacher and mentor cover, and if and how more effective planning for these ITAP weeks could reduce this impact.

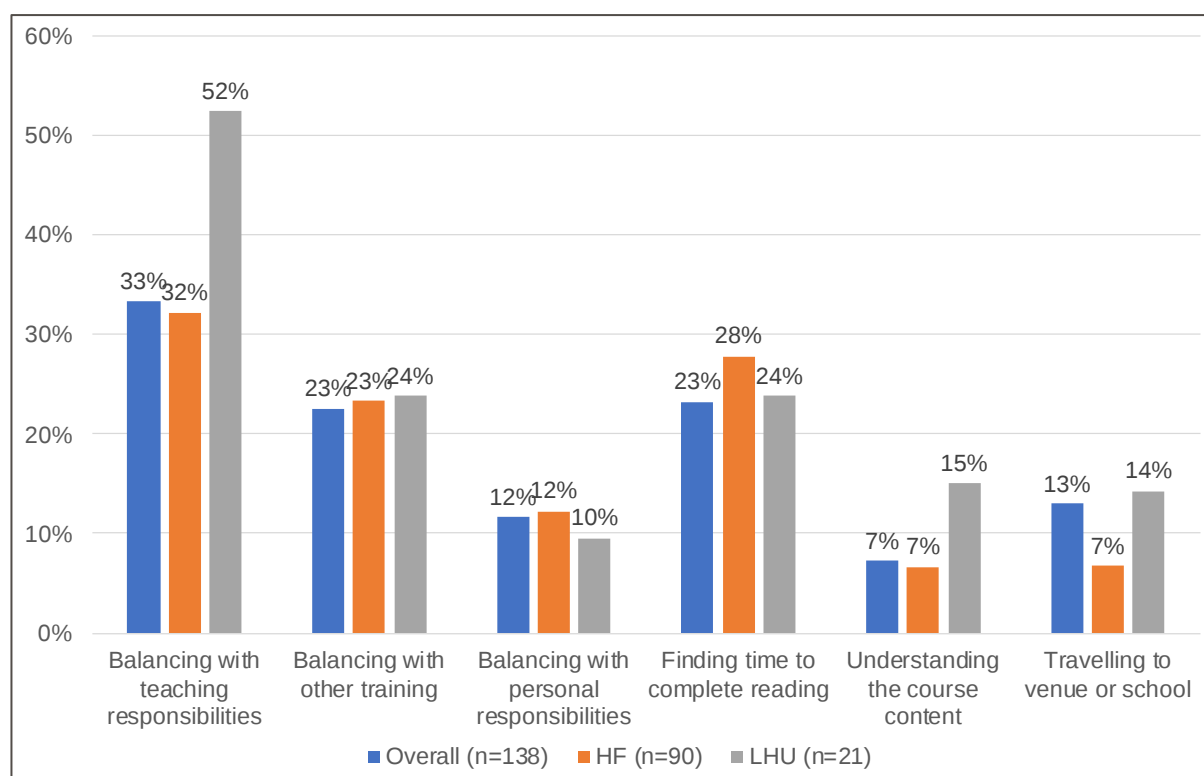
3.4.3 What are the challenges involved in taking part in the different models? How did trainees, teacher educators and providers overcome these challenges?

Challenges

The majority of trainees who responded to the post-pilot survey did not report any challenges in participating in ITAP. However, a notable 33% did feel that balancing the demands of the ITAP with their regular teaching workload made it difficult for them to take part in the pilot (see Figure 16). In addition, 22% reported difficulties balancing the demands of their ITAP against other training.

We can see that balancing the ITAP with teaching commitments was more commonly mentioned as a challenge by trainees in the LHU model, which involved a break in their current placement activity, than by HF trainees who continued with their usual teaching commitments. However even for HF trainees who stayed in their usual placement school and followed the usual pattern of a week as per their ITT to date, this was a challenge for around a third (32%) of trainees, suggesting that this challenge relates to the content and focus of ITAP as well as location or routine changes.

Figure 16: Post-pilot survey: Proportion of trainees reporting challenges in participating in ITAP pilot



Challenges around balancing the ITAP with personal responsibilities such as parenting or caring were more common among younger trainees (under 34 and particularly the 25-34 age group) and among those not receiving a salary during their training¹⁴. There were no significant differences by gender (see Annex E of the appendices for all sub-group comparisons).

Just 7% of trainees said in the post-pilot survey that they had had difficulties understanding the content (see Figure 16). Nevertheless, the potential for cognitive overload during ITAP was a key theme emerging from our qualitative analysis. Trainees generally found the pilot to be intense but also beneficial. However, trainees from across all four providers commented that the first day of training was perhaps too intense. For example, one trainee stated that they were “overwhelmed by amount of theory and workload...” on the first day. There was also a suggestion that some participants were so tired from lengthy sessions earlier in the week that they could not focus on or attend later sessions. Our own observations on attendance do not suggest this was a major problem, but that some specific individuals may have been affected.

Related to the theme of cognitive overload, some trainees reported struggling with having to learn the topic of questioning while not having much experience managing other aspects of teaching, such as student behaviour. Educators added further nuance to this point by noting that all trainees had had to switch topics in the middle of their curriculum while some – like those in the HHS pilot – had had to switch schools. This, according to them, created a high cognitive burden on trainees.

In addition, trainees and educators raised concerns about the emotional burden and pressure caused by the nature of the pilot. One trainee explained why they were feeling anxious or nervous by noting:

¹⁴ It is not possible to know whether more or fewer trainees – or the same groups – experience similar challenges in balancing their wider ITT participation with these responsibilities, as the survey only asked about the ITAP pilot.

"Especially because we were only in our third or fourth week and have all these people in the room with us. And then also it was like, this is a new thing as well. It's pressure to get it right."

Educators appeared to be conscious of the enhanced emotional burden on trainees as well, especially given that they were being expected to undertake ITAP very early in their training:

"They found it quite challenging in terms of their resilience, their professionalism and I think their emotional intelligence. I think what we're asking them to do at this early stage of this intensive practice is to [...] put themselves in positions where they're teaching in front of other people, where they're pulling together all those different aspects of training and trying to deliver on them. I think for some of them that has felt quite high stakes in some ways. It's different for each of them, you can't generalize, but I think they've needed more reassurance than I would have anticipated. As I said at the beginning, it's really accelerated their progress, but for some of them, at what cost in terms of [...] emotional aspects?"

The educator from the quote above raised concerns about the wider implications of placing this pressure, worrying that if ITAP was "too much" for trainees early in their training and this might negatively affect retention.

Other challenges related to the lack of a fully supportive environment for trainees. Our evidence shows that some mentors did not attend training sessions or did not have enough time to prepare for the course. This meant that not all trainees received the best input from their mentors, and some may not have been supported fully. In models that relied heavily on peer feedback, such as the HHS model, trainees suggested that not having good relationships meant they could feel frustrated by the feedback process.

A final trainee-level challenge emerged where trainees felt unable to see how the content of the ITAP on questioning was relevant to or could be applied in their specific teaching context. Examples included applying the content in different phases, in specific subjects, within the context of specific phonic curricula as discussed above and in special schools or other contexts working with pupils with special educational needs or disabilities [SEND]. While providers felt that the content of the ITAP was relevant for their cohorts, and educators and trainees linked this challenge to the specific topic covered by this pilot, overcoming this perceived lack of relevance and the challenge of providing content that is suitably tailored will likely be a challenge for all ITAP.

Of concern is indicative evidence that there may be sub-groups of trainees who may have experienced greater challenges in participating than others. The first sub-group was trainees with some special educational needs or disabilities (SEND). According to providers, this group experienced higher levels of anxiety, especially during activities such as deliberate practice, and needed additional support to overcome barriers such as difficulties engaging with approximated scenarios without full contextualisation. The second sub-group includes those trainees with less or no experience. Approximately three-quarters (77%) of trainees participating in the pilot had prior classroom experience as either a teaching assistant or teaching in other schools. According to educators, in some cases the pilot required those with less classroom or teaching experience to demonstrate professionalism and emotional. Finally, educators pointed out that the primary school teachers may have less of a culture of evidence-based practice and research than secondary school teachers, and that this may mean they had less familiarity with terminology than secondary trainees going into the pilot, limiting their effectiveness. The survey found no significant differences between the experiences of trainees with and without prior classroom experience and very few differences by school phase but was limited in the challenges and areas of experience that it asked about.

Solutions

All providers employed a variety of proactive and reactive solutions to aid trainees in participating. These generally focused on providing specialised material and access to mentors to manage the

cognitive overload, facilitating peer support and advertising pastoral care to address a potentially increased emotional burden. We summarise these solutions in the table below.

In our view, the solutions for the risk of cognitive overload were largely adequate, as borne out in the survey results. And with greater time for planning and consideration of the lessons learned from this pilot, solutions to this could easily be optimised further for future delivery. The solutions for the challenges of emotional burden, on the other hand, may need a greater focus in the future. Communications to trainees around assessments and whether they were high stakes or not, and dedicated coaching for confidence combined with reminders of pastoral care support available may be useful. Similarly, while providers did try to address the perceived lack of relevance to specific contexts – whether through further explanation and discussion in the core training or through the one-to-one support trainees received in their placement schools – our qualitative and quantitative findings suggest this challenge was not fully addressed. For the perceived lack of a fully supportive environment, it is hard for us to assess adequacy at this stage – the main solution for such instances in the future may be a solid and speedy feedback loop that facilitates adaptive programming to enhance the experience of all trainees.

Table 13: Solutions to participation challenges

Nature of challenge	Solutions employed	Adequacy for future
Cognitive overload	Focus of ITAP on questioning, support from mentors, specialised materials	Adequate with room to streamline further
Emotional burden	Peer support, support from mentors, anonymised digital approximations feedback, advertising pastoral care support	May need greater focus in future
Lack of tailoring/perceived lack of relevance	Explicit explanation of how pedagogies could be applied in different contexts, use of subject-/context-specific mentors in placement schools	May need greater focus in future
Perceived lack of a fully supportive environment	Training for mentors, support from mentors, support from core design team	Hard to assess

3.5 Scalability

Key findings: scaling is possible with additional planning and resources

1. Scaling this pilot will entail two types of processes: scaling by repetition (by including more ITAP in a year) and by volume (provision of ITAP for all trainees).
2. Key barriers to scaling by repetition include upfront investment in planning and design efforts, increased demands on schools and mentors, and enhanced pressure on trainees.
3. Key barriers to scaling by volume include maintaining quality of delivery as the volume of trainees increases and logistical readiness to cater for larger cohorts.
4. Barriers to scaling by repetition and volume may differ by model design but can be addressed by additional planning, resources and usage of digital technology solutions.
5. Tight budgets may prove to be a notable barrier to scaling and providers in the pilot lacked clarity around the nature of future funding for ITAP.

3.5.1 What are the barriers and potential solutions to taking the piloted models to scale?

Scaling the piloted models to comply with the government's requirement will require scaling both by repetition and by volume. By repetition, scaling entails delivering the ITAP course approximately four times a year, but with different topics of focus. By volume, scaling entails delivering the same

four ITAP courses a year to all trainees, and as a result, involving more schools, more educators, and more mentors. Estimates indicate that approximately 30,000 teachers enter ITT in England in a year and are therefore expected to be affected by the 2024 policy change. In the below table, we summarise key barriers to scaling raised by respondents alongside potential solutions to those barriers.

Table 14: Barriers and solutions to scaling

	Scaling by repetition	Scaling by volume
Barriers	• Additional planning • Increased demand on schools • Pressure on trainees • Novelty wearing off • Costs • Timeline	• Quality assurance • Logistical barriers • Additional planning • Increased demand on schools • Costs • Timeline
Potential solutions	• Additional resources • Upfront planning investment • Using ITAP specific mentors • Training for mentors and schools • Recruiting school coordinators • Offering schools flexibility • Revisiting ITT curriculum • Designing distinctive courses • Digital approximations	• Additional resources • Upfront planning investment • Using ITAP specific mentors • Training for mentors and schools • Recruiting school coordinators • Digital approximations • Online delivery

Scaling by repetition

The upfront investment in planning that is needed is a major challenge. According to providers, substantial initial planning and design efforts are required to design the remaining three ITAP courses. Educators noted that developing a calendar for the full ITAP mandate ahead of time would be important as this will allow a more coherent curriculum and smoother flow of activities in the year. They also highlighted the importance of selecting the topics for ITAP delivery in advance, with some stressing that continued government flexibility in topic selection would best allow providers to tailor their curricula to the needs of trainees. To address these additional planning needs, inserting supplementary resources temporarily, as was done in this pilot, might be one solution. This is because once the four ITAP courses have been designed, planning needs will likely reduce to more streamlined levels for all providers.

A second challenge is the increased demand on schools. More ITAP courses necessarily mean more practice in schools, more mentor support, and more coordination with school staff. One solution suggested by providers for this barrier was to develop a pool of specialist ITAP tutors who could reduce the burden on in-school mentors. A second solution offered was to engage schools proactively through dedicated ITAP-specific school coordinators whose role would be to communicate to schools, ensure buy-in, and facilitate smooth delivery. For instance, some providers used existing Professional Learning Coordinators to perform a similar role successfully. A third suggested solution was increasing the support and training provided to schools and mentors to aid ITAP delivery. Because the mandate is likely to allow for some flexibility around timing of ITAP courses, another suggested solution was to use this flexibility to address the potential overburdening of schools. This could for instance take the form of offering schools flexibility in when to deliver ITAP in the school calendar or a blocked approach in which the intensive week is

delivered over a longer period¹⁵. A final suggested solution that is not within the control of providers was that the DfE could reduce the requirement for ITAP to fewer hours (and by implication fewer cycles within an academic year or PGCE course).

A further challenge associated with scaling by repetition is related to the impact scaling may have on trainees. Many educators noted that the added pressure on trainees of four intensive weeks could pose a barrier and have consequences on attraction into and retention in the profession, unless this pressure is actively managed. One provider noted that the pressure might be particularly acute for those pursuing PGCE curricula. The main way to address this barrier – outside of a reduction in the ITAP requirement – would be for providers to revisit their ITT curricula and streamline these to keep the total workload manageable.

Due to scaling by repetition, the novelty experienced by trainees in this pilot may wear off, posing a barrier. Unless the four rounds are made to be sufficiently engaging, several educators warned, learning outcomes may suffer. Some respondents felt that this was inevitable, but not necessarily problematic. One solution offered to address this was the use of digital approximations which may reduce stress related to logistics, while at the same time offering novelty over a longer period.

Scaling by volume

A major barrier to scaling by volume is maintaining consistency in the quality of delivery. Providers noted concerns over ensuring that quality of delivery would be maintained once scaling by volume occurred. This, however, is an issue with delivering quality ITT for all providers, and once the core models and their associated processes have been developed, cascading of the material and quality assurance of programming can occur in a way similar to that for other elements of the ITT. This would nonetheless require a one-off redesign to allow integration of the ITAP into the ITT.

Logistical readiness, such as finding enough centre- and school-based space for the increased volume of trainees, is also a barrier. Educators from HHS, which used ITAP specific host schools, highlighted the need to rely on a bank of schools which may or may not have the space or size to accommodate the increased volume of trainees. Meanwhile, educators from the two providers delivering to larger cohorts, HF and STS, noted capacity constraints, reporting that finding enough physical space would be a barrier for them. Like scaling by repetition, scaling by volume will also require additional upfront planning, as well as proactive management of the increased demand on schools. For these barriers, the same solutions noted above are also likely to apply.

One solution to consider for scaling by volume barriers is the use of digital technologies. The digital approximations used in this pilot, for instance, were originally designed and tested because they lent themselves to providing opportunities for practice, feedback, and assessment, including at scale. Other digital solutions for scaling discussed by providers included online or hybrid delivery and the development of additional digital teaching resources, such as videos for demonstrating and analysing expert practice. Initial evidence on the use of digital approximations (HH and STS) and hybrid delivery (STS) from this pilot is promising. However, providers and educators noted that the use of digital technologies is not a straightforward or solve-all solution. There was a view among educators that moving too much of the course online has the risk of diluting the quality of the trainee experience. In addition, all providers may not have the technical capabilities to develop and deliver high quality online training. And developing digital resources such as digital approximations and videos requires financial investment. Relatedly, some providers worried about the potential costs or maintenance fees involved in using digital technologies and saw this a potential barrier to scaling up the use of digital approximations used in this pilot. However, while the cost estimates collected as part of this evaluation and direct comparisons between the costs reported for the

¹⁵ This could also help with concerns around cognitive load and perceived pressure for trainees, though it was not posed as a solution to that challenge by our participants.

different models should be treated with caution, the one model involving digital approximations that had reported costs appeared to be notably less costly than the alternatives piloted (see Annex B of the appendices for both these findings and why they should be treated with caution).

The final important considerations for scaling by either repetition or volume are financial. For all providers scaling ITAP, some additional upfront costs to facilitate design and planning will be required. Following this, providers will also need running costs for ongoing delivery such as to pay for school placement fees, teacher cover, training of mentors, recruitment of additional staff, and sourcing venues for teaching. To scale ITAP successfully, providers will likely need to revisit their entire budgets and carefully consider if the financing for ITAP can be sourced by simply reallocating resources, or if additional resources may need to be raised. It is currently unclear to providers if there will be avenues of funding available from the government, and without this, designing and delivering ITAP from 2024 may prove to be challenging for some, especially given the short time left to this deadline.

4. Limitations



A key limitation of our evaluation is that results may be affected by self-selection bias. Trainees self-selected both into interviews and into the survey. While we still obtained a diverse group of respondents, and triangulated our findings using other methods, it is possible that trainees who had particularly positive or negative experiences chose to participate in our research, skewing our findings.

Another limitation is that some providers are under-represented in our survey sample. There were logistical challenges for providers implementing the survey at both time points, including trainee absence and a lack of time within ITT curricula to allow protected time for trainees to complete the surveys¹⁶. The low post-pilot survey response rate was largely driven by very low responses among STS trainees. STS were unable to set time aside for trainees to complete the survey after the last training session for logistical reasons and this negatively affected the overall response rate. The low response rate at the post-pilot survey for STS trainees means that they are under-represented and that we cannot provide disaggregated findings for this provider. LHU and HHS had relatively small cohorts and also make up a minority of responses in our unweighted analyses. Unfortunately, the sample size for HHS was too small to carry out disaggregated analysis for this provider, but we do so for LHU. Together these factors mean that trainees from the other large-scale provider, HF, formed 65% of post-pilot survey respondents. To address this, where feasible and appropriate, we present findings by provider and triangulate findings using qualitative data. Nevertheless, we advise readers to exercise caution in interpreting the survey results, and specifically the results from our

¹⁶ We expected that response rates would be highest if allocated time was given for trainees to complete them – for example on the morning their ITAP pilot began and in a core training day after the pilot. This played out in practice, but was not always feasible.

statistical sign test which relies on an even smaller subset of survey responses. In our view, these results are indicative only.

There is an additional limitation related to the nature of our evaluation that is worth stating explicitly. Our evaluation is intended to be formative and, by design, does not make causal claims about the performance of any model or recommendations on the most advantageous combination of cost and quality. Both are beyond our scope. Rather, the aim of this report is much more modest – we aim to facilitate learning, and shed light on how different models of ITAP can be designed and implemented, the challenges they might face, and the solutions that could make them better.

5. Conclusions



Our formative evaluation supplies reasons to be optimistic not only about the promise of all four pilot models, but also about the promise of ITAP. Though our survey findings should be treated with caution in line with achieved response rates (see Sections 3.1 and 4 of this report), most trainees who responded to the post-pilot survey rated the models positively, finding them enjoyable and relevant. The qualitative research found that all stakeholder groups perceived the pilot to have met its intended outcomes and benefits for trainees were linked to the singular focus on questioning, and the marriage of theory and practice including timely opportunities to practically apply learning. Indicative quantitative evidence also showed positive change in knowledge, understanding and confidence over time. Moreover, we found some evidence that benefits for trainees may have extended beyond the models' focus on questioning practice and that the pilot may have contributed to learning and development for teacher educators as well. While our evidence is indicative rather than causal, our early findings suggest cause for cautious optimism.

We found the designs of all four models to be appropriate, and feasibility of implementation to be reasonable. Our evidence suggests that high quality, experienced educators from the four providers designed the models and led delivery in an effective manner. In addition, our formative evaluation offers rich evidence on a varied set of implementation challenges experienced in the pilot as well as a diverse set of solutions to consider when addressing these challenges. We note three areas of challenge that providers should prioritise addressing to best support trainees:

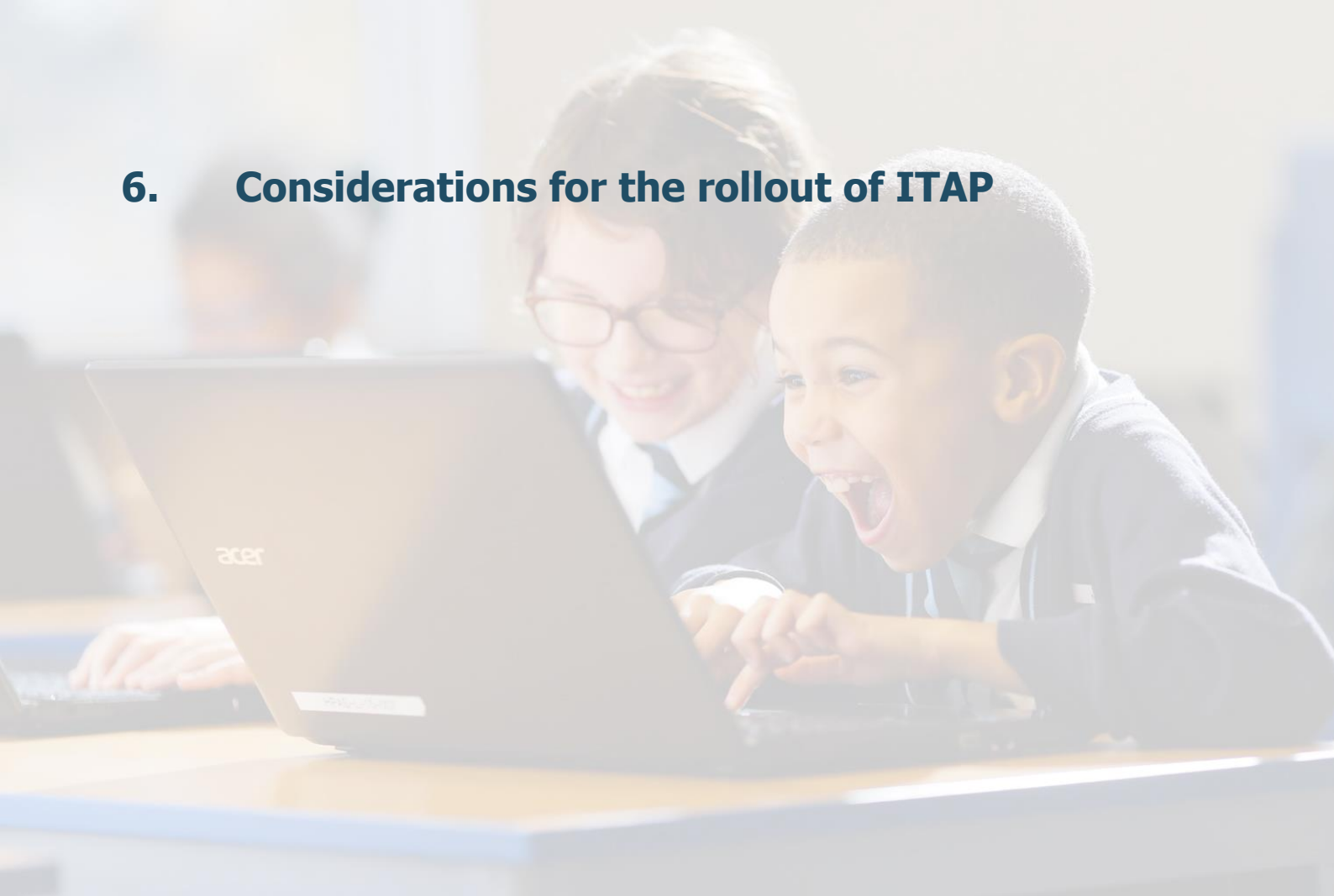
- providing reassurance and emotional support to address any risk of added pressure on trainees
- considering how to tailor content and how to demonstrate trainees that core ITAP content is relevant to the contexts they work in

- avoiding overburdening and securing buy-in from schools and mentors to ensure their vital contributions to ITAP are maximised

An important finding was that design features that often set the provider models apart involved trade-offs that enhanced the potential for improving outcomes on one hand while also increasing the potential for challenges on the other. For example, while digital approximations were enjoyed by trainees and understood to have pedagogical benefits and potential for scale on one hand, they introduced challenges around access, familiarity and hesitance on the other. Likewise, varying the rhythm and location of the course against usual ITT provision in 2022/23 offered trainees a unique experience that allowed them to interact with new schools, professionals, and peers, while simultaneously resulting in greater resource requirements and logistical challenges and in some cases even concerns over the fast pace. There was also a trade-off between the perceived pedagogical benefits of intensive training practice and risks around overwhelming trainees cognitively and emotionally.

Our analysis indicates that scalability by repetition and by volume is possible, albeit with additional planning and resources. How able providers are to reallocate existing resources to facilitate ITAP design and delivery will ultimately determine the number of additional resources needed. Because funding support for ITAP remains unclear in the minds of many providers, the cost implications of scaling on one hand, combined with worries about the short time left to develop these programmes on the other, may be perceived as barriers that are hard to overcome. That said, digital technologies and sharing of information to aid planning and design could go a long way in addressing these fears and supporting the upcoming policy change.

6. Considerations for the rollout of ITAP



Below we outline some key considerations for ITT providers as they prepare to design and deliver ITAP, based on the learning from this pilot. We also outline considerations for organisations like NIoT who aim to support ITT providers as the requirement is rolled out, and for future evaluations of ITAP.

6.1 Considerations for ITT providers

As ITT providers prepare to implement ITAP from 2024/25, we suggest that they consider seven things:

1. Plan ahead. The sector has two years to prepare for the ITAP requirement and providers will have the support of a DfE associate. The earlier that providers begin preparations the more they will be able to communicate effectively with all stakeholders and ensure all practical considerations are addressed. Planning for how and when ITAP will be incorporated into wider ITT curricula will also be important.
2. Proactively and carefully identify the topics that will be delivered using ITAP. Planning for scaling by repetition will require proactive consideration of which topics be selected for ITAP. Learning from this pilot suggested that topics should relate to core or critical teaching competencies and should be narrow in that they enable a singular focus but relevant to all teacher trainees in a cohort.
3. Ensure that ITAP provision explicitly explains the relevance of and teaches trainees how to adapt core content to different teaching contexts. Though providers felt the questioning content in their piloted models was relevant for all trainee teachers, some trainees did not see this

relevance and there were concerns among educators about how to best tailor the curriculum for specific contexts. This included for instance applying the content in different phases and for different subjects, but also applying it in special needs settings.

4. Communicate early, clearly and comprehensively to all stakeholders. Scaling these pilots successfully will require engagement of schools and trainees. Developing and implementing a comprehensive communication strategy that is clear and timely may help providers to proactively manage resistance and apprehension in both sets of stakeholders. Trainees would benefit from reassurance around the “stakes” behind feedback and assessment during ITAP, while schools would gain from learning about potential professional development benefits for their staff. Avoid overburdening schools wherever possible. Burden on schools and particularly mentors may pose a serious threat to scaling unless they are addressed through for instance digital options, resourcing support, or other similar means.
5. Avoid overburdening schools wherever possible. Burden on schools and particularly mentors may pose a serious threat to scaling unless they are addressed through for instance digital options, resourcing support, or other similar means.
6. Ensure trainees are aware of available pastoral support. A selection of trainees may feel the intensity of such programmes more deeply, such as for instance those with special needs. It is important that any avenues for pastoral care are advertised widely, ensuring access by trainees to mental health support is available.
7. Seek out and use resources available to support design thinking and delivery. Providers in this pilot found the five-element framework useful for thinking about the types of activities they could include in their ITAP model and why. The evidence on what works in ITAP is limited but will grow as the ITAP requirement is rolled out in England. We also encourage providers to consider seeking out information on specific, evidence-based CPD activities and approaches that are relevant to ITAP to inform the design of their bespoke models.

6.2 Considerations for organisations supporting ITT providers

Here we share four pieces of advice for NIoT and other organisations looking to support ITT providers as they prepare for the ITAP requirement.

The first three considerations relate to supporting providers to share with and learn from each other. As noted above, the evidence base and resources available for ITAP are limited but will grow over time. Organisations looking to support ITT providers could consider:

1. Build a bank of ITAP resources for sharing. This could include resources for use in ITAP delivery - for instance videos or other materials that trainees can analyse to understand teaching practice or digital approximations that can be used to support trainees to practice implementing their learning and get feedback outside of the classroom. It could also include the framework used in the evaluation as a tool to support design thinking.
2. Assist providers to share evidence and learning on ITAP implementation and effectiveness. The new mandate will pose a learning curve for the entire teacher training sector in England. Openly sharing learning about what options are available to providers, what is feasible and what works and what is/does not will be crucial to improving design and delivery over time. NIoT intend that this report informs this learning, and we encourage them to consider how it and other learning can best be disseminated and built on in coming months and years.

3. Invest in building the evidence base. We encourage investment in collecting the evidence needed to understand what works and why – for trainees and for their learners and schools. Organisations such as NIoT can also play a key role in building the evidence base of what works in ITAP.

Our final consideration relates to advocacy. Organisations looking to support ITT providers may want to consider:

4. Consult on and advocate for the right balance of guidance and flexibility from the Government. Many considered the guidance to be vague. As other providers implement ITAP, there are chances that interpretations will vary greatly. Thus, advocating for clearer guidance may ensure that all providers have confidence that they are meeting the requirement and maintain some consistency in ITAP design and delivery. Having said that, clarity does not mean prescriptiveness. There is a danger that guidance may become too strict thereby limiting flexibility, innovation, and cost-effectiveness for some. Advocating for the right balance of guidance and flexibility will be critical and organisations such as NIoT can play a key role in this.

6.3 Considerations for future evaluations of ITAP

Evaluators and funders like NIoT should consider:

1. Employ both formative and summative evaluations. A mix of formative and summative evaluations will need to be considered to build a robust evidence base on what works and how in ITAP.
2. Invest in impact evaluations. While it is important to understand what is feasible and acceptable, ITT providers want to know that the design choices they make will lead to effective teacher training. Rigorous impact evaluations such as randomised controlled trials (RCTs) or quasi-experimental evaluations will be needed in order for providers to make evidence-based decisions.
3. Which impact questions are of most importance. For example, there will be a choice to be made about whether to measure the impact of ITAP as freely interpreted by providers or to try and measure the impact of specific elements of or approaches to ITAP. There will also be questions to answer about which topics and areas teaching practice to measure the impact of ITAP on. All are important questions but will have implications for trial design and resource requirements.
4. Which trainees benefit the most or least. Evidence from our formative evaluation suggests that some groups of trainees may experience particular barriers to participation, but the evidence on whether this translated into differential benefits is limited and mixed. Considering who is included by and benefits from ITAP will be a critical question for process and impact evaluations.
5. Long-term impact. Evidence on long-term teacher- and pupil- outcomes and specifically teacher retention is crucial to the evidence and could also uncover interrogate some of the trade-offs that providers were concerned there might be between ITAP and other elements of ITT.
6. Value for Money (VfM). Different design decisions will involve trade-offs around impact and sustainability. By conducting robust VfM analysis for different models, providers will be able to select the optimal trade-offs for themselves and for trainees.
7. How to capture diverse perspectives from key stakeholder groups. ITAP design and delivery is going to involve a variety of stakeholder groups as well as different types of providers who vary by size, nature, and location. Evaluations should involve the range of stakeholders – including trainee teachers, educators and provider staff – as well as a variety of providers including SCITT and HEIs.

References

- Burn, K., and Mutton, T. (2013) "Review of "Research-informed Clinical Practice" in Initial Teacher Education". *BERA Paper 4*. Available at <https://www.bera.ac.uk/wp-content/uploads/2014/02/BERA-Paper-4-Research-informed-clinical-practice.pdf>
- Burn, K. and Mutton, T. (2015) "Review of 'research-informed clinical practice' in Initial Teacher Education" *Oxford Review of Education*, 41 (2). <https://doi.org/10.1080/03054985.2015.1020104>
- Carter, A. (2015) *Carter review of initial teacher training (ITT)*. London: DfE. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/399957/Carter_Review.pdf
- Cohen, J., Wong, V., Krishnamachari, A. and Berlin, R. (2020) "Teacher coaching in a simulated environment". *Educational Evaluation and Policy Analysis*, 42(2), 208-231. 24.
- Deans for Impact (2016) *Practice with Purpose. The Emerging Science of Teacher Expertise*. Austin, TX: Deans for Impact. Available at https://deansforimpact.org/wp-content/uploads/2016/12/Practice-with-Purpose_FOR-PRINT_113016.pdf
- Deans for Impact (2017) *Building Blocks Framework*. Austin, TX: Deans for Impact. Available at https://deansforimpact.org/wp-content/uploads/2017/11/Building-Blocks_Framework.pdf#:~:text=Building%20Blocks%20Framework%20Deans%20for%20Impact%20set%20out,four%20common%20elements%20--%20modeling%2C%20practice%2C%20feedback%2C%20and
- DfE [Department for Education] (2019a) *Early Career Framework*. London: DfE. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/978358/Early-Career_Framework_April_2021.pdf
- DfE [Department for Education] (2019b) *ITT Core Content Framework*. London: DfE. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974307/ITT_core_content_framework_.pdf
- DfE [Department for Education] (2021) *Government response to the initial teacher training (ITT) review report*. London: DfE. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1059746/FOR_PUBLICATION_Government_response_to_the_initial_teacher_training_ITT_market_review_report.pdf
- DfE [Department for Education] (2022) *ITT Reform: accompanying document*. London: DfE. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1079080/ITT_Reform_Accompanying_Document.pdf
- Farndon, S. (no date) "What is Deliberate Practice" *Ambition Institute Blog*. Online at <https://www.ambition.org.uk/blog/what-deliberate-practice/>
- Grossman, P. (2018) (ed.) *Teaching Core Practices in Teacher Education* Cambridge, MA: Harvard University Press
- ITT Market Review Expert Advisory Group (2021) *Initial teacher training (ITT) market review report*.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/999621/ITT_market_review_report.pdf

NCATE [National Council for Accreditation of Teacher Education] (2010) *Transforming Teacher Education Through Clinical Practice: National Strategy to Prepare Effective Teachers*. Washington, DC: US Department for Education. Available at <https://eric.ed.gov/?id=ED512807>

NIoT [National Institute of Teaching] (2022) *Intensive Training and Practice Pilot: Pilot Evaluation Plan*. London: NIoT. Available at https://niot.s3.amazonaws.com/documents/Evaluation_of_the_IPP_Protocol.pdf.

Perry, E., Booth, J., Owen, D., Bower, K. (2019) *The Curriculum for Initial Teacher Education: Literature Review*. Sheffield: Sheffield Institute of Education.

Sahlberg, P., Broadfoot, P., Coolahan, J., Furlong, J., Kirk, G. (2014). *Final Report on the Structure of ITE in Northern Ireland*.

Sims, S. (2021) "Making good on the initial teacher training market review" *CEPEO Working Paper No. 21-09*. London: Centre for Education Policy and Equalising Opportunities, UCL. Available at <https://repec-cepeo.ucl.ac.uk/cepeow/cepeowp21-09.pdf>

Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Van Herwegen, J. and Anders, J. (2021) *What are the Characteristics of Teacher Professional Development that Increase Pupil Achievement? A systematic review and meta-analysis*. London: Education Endowment Foundation. Available at <https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/teacher-professional-development-characteristics>

Teaching Works (no date) "Teacher Education Pedagogies" *Teaching Works Curriculum Resources*. Online at <https://www.teachingworks.org>

Zhou, M. and Brown, D. (2015). *Educational Learning Theories*. Education Open Textbooks. 1.