



Department
for Business
Innovation & Skills

RESEARCH

BIS RESEARCH PAPER NUMBER 191

Evaluation of English and Maths
Pathfinder Pilots

DECEMBER 2014

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Business, Innovation and Skills.

Department for Business, Innovation and Skills

1 Victoria Street

London SW1H 0ET

www.gov.uk/bis

Research paper number 191

December 2014

Contents

Contents	2
Executive summary	5
Context	5
Method.....	5
Findings	7
1. Background and context	8
Project aims	9
The Pathfinder pilots	10
Structure of this report.....	10
2. Research methods used.....	12
Provider workshops.....	12
Qualitative analysis	12
Fieldwork in Pathfinder pilots.....	13
Critical review of assessment tools.....	14
Quantitative analysis	15
3. Qualitative evaluation	19
Characteristics of Pathfinder providers.....	19
Learner cohorts included in the Pathfinder	26
Providers' choices of assessment tools and implementation models.....	28
Challenges arising from the adoption of a distance travelled approach	29
The suitability of existing tools to measure distance travelled.....	29
Collating learner information.....	30
Availability of appropriate ICT facilities	30

Administering a post-test.....	31
Tutor buy-in to the new approach.....	31
Learner buy-in to the new approach.....	32
Implications for the future development of a distance travelled approach.....	32
Appropriately incentivising providers.....	32
Effectively motivating learners.....	33
Developing a distance travelled tool/model.....	33
Data collection requirements.....	34
Impact of adopting a distance-travelled approach.....	34
Financial Implications of a distance travelled-approach.....	35
Start-up costs of rolling-out a distance travelled approach.....	36
Fixed costs of rolling-out of a distance travelled approach.....	37
4. Review of assessment tools.....	39
Focus and content of the review.....	40
Outcomes of the review.....	41
Conclusions and suggestions.....	43
5. Quantitative analysis: initial findings.....	45
Data quality.....	45
Challenges to data completeness.....	45
Descriptive statistics.....	46
Demographics.....	47
Qualification level of cohorts.....	48
Relationship between assessments and courses?.....	49
Analysis of validity and distance travelled.....	51
Discriminant validity.....	52

Concurrent validity.....	52
Distance travelled.....	54
Predictive validity	56
In conclusion	57
6. Conclusions and implications.....	59

Executive summary

This report summarises research commissioned by the Department of Business, Innovation and Skills (BIS) exploring the feasibility of a funding system based on the distance travelled by learners on adult literacy or numeracy courses. The research was conducted by SQW, in partnership with NRDC, between September 2012 and September 2013.

Context

As the Skills for Life survey demonstrated, between 2003 and 2011 there had been a substantial improvement in the literacy skills of 16-65 year olds in England at Level 2 and above (a change of 12.4 percentage points). There was, however, no matched improvement in lower level literacy and numeracy skills; indeed, these showed a small overall decline. This highlighted the need to take account of the curriculum in use (and the teaching that delivers it); the way the curriculum is assessed; and the way that providers are incentivised to ensure learners are enabled to develop their skills.

Following the moves signalled in the coalition government's Skills for Sustainable Growth, a total of 17 Pathfinder pilots were established to explore the feasibility of a funding system based on the distance travelled (rather than just qualifications gained) by learners in order to incentivise providers to maximise skills acquisition. The Pathfinder pilots were based in a range of different provider settings, including further education colleges, community settings, private training providers and not-for-profit organisations. Providers identified, adopted or modified tools, both initial assessment tools (commonly used to provide an indicative level), and diagnostic tools (frequently used to establish a full skills profile) that they hoped would provide the evidence they needed of the distance travelled by individual learners, especially those working at a low level (including Entry level 1, Entry level 2, Entry level 3 and Level 1).

Method

The study involved a detailed process evaluation, examining the various contexts (such as provider settings and range and type of learners) and mechanisms in place (including the factors affecting the choice of assessment tools and the validity of those tools), This approach was adopted in order to assess which Pathfinder approach appeared most effective in measuring distance-travelled and to provide a basis on which a funding model could be developed.

The study included:

- Three **workshops for providers** (facilitated by BIS) during the course of the pilot, which were used to present an overview of the evaluation and circulate and discuss the draft data collection tools designed to collate data on learners, their courses, their pre-tests and post-test scores and their subsequent qualifications (September 2012); look at progress in implementing the Pilot, explore the available data and discuss the emerging findings relating to the use of different tools and the process of assessment (February 2013); and examine the implications of the research findings for policy and practice (September 2014).
- A **detailed qualitative study**, incorporating two tranches of fieldwork involving a series of visits to the 17 Pathfinder sites. A total of 121 interviews were conducted, some in one-to-one discussions, others in group settings.
- A **critical review** of 36 of the 41 **assessment tools** used by the Pathfinder pilots; five tools were not made available to the research team in time to be reviewed. This review examined the validity of the tools (whether the items measured (or appear to measure) what they were intended to measure); manageability for those being tested and for those administering and scoring the tests; usefulness - that the results provided information on achievement and (where appropriate) progress that was clear and meaningful; whether the tools existed in statistically equated parallel forms (and so could be used effectively to measure distance travelled) and the length and precision of the measurement scale
- A **comprehensive quantitative analysis**, which provided descriptive statistics on the demography, qualification levels and courses being followed by the participants and an analysis of:
 - the relationship of learners' scores in the pre-test tools and their previous attainment (the **convergent** and **discriminant validity** of the tools)
 - the extent of learner improvement between pre-test and post-test scores (the measurement of **distance travelled**)
 - the measurement of pre-test and post-test scores and distance travelled for learners with different skill levels (**concurrent validity**)
 - the relationship of learners' scores in the post-test tools and their subsequent attainment (the **predictive validity** of the tools)

Findings

It should be noted that, prior to the Pilot, none of the participating sites had specifically administered a post-test, relying (in general) on external qualifications to measure learner progression. Moreover, as some providers noted, the primary outcome many were seeking to achieve for their learners was entry into employment, rather than measurable progression in maths or English per se. The study has, therefore, highlighted a number of issues that have implications for the longer-term use of a distance-travelled approach to funding.

- Providers' experiences, the critical review of the assessment tools they deployed and the statistical assessment suggest that none of the tools used in the pilot is suitable for use as the metric by which distance travelled should be assessed.
- The statistical analysis indicated that none of the tools had good concurrent or predictive validity and so could not be used in any reliable way as a pre-test alongside learners' subsequent formal qualifications for measuring distance-travelled on a larger scale.
- Online and adaptive tools would, arguably, be the most cost-effective way of conducting pre-test/post-test assessments, but raise significant logistical (and other) challenges in cases where access to ICT facilities is limited.
- The level of preparedness for data collection, recording and collating varied widely across the pilot sites, all of whom were volunteers and eager to make a contribution to this process. This highlighted the guidance and support in the recording and sharing of data means for adult learning providers.
- If distance-travelled is to be used as the basis for a future funding mechanism, pre and post assessment results will need to be recorded on the ILR as compulsory fields alongside data on course completion and qualification achievement. Moreover, additional data fields (such as language used in the home) may be needed on the ILR, while existing data fields (including employment status) may need to be more rigorously completed.

The study subsequently considered a number of different options for addressing these issues were a distance-travelled approach to be adopted at any point in the future. These findings suggest that the development of bespoke tools for the measurement of distance travelled would be the **most robust option** in establishing an assessment process that reliably assesses changes in a learner's skills profile over the length of the course.

1. Background and context

- 1.1 As the Skills for Life survey has demonstrated, between 2003 and 2011 there has been a substantial improvement in the literacy skills of 16-65 year olds in England at Level 2 and above (a change of 12.4 percentage points);¹ but there has been no improvement in lower level literacy, and numeracy skills, indeed, these have shown a small overall decline.
- 1.2 In order to focus on the development of adult skills, there is a need to take account of the curriculum in use (and the teaching that delivers it); the way the curriculum is assessed and the way that providers are incentivised to ensure learners are enabled to develop their skills. The Skills for Life curriculum has gained wide acceptance, and the assessments used to support it have been updated; since August 2012, the Literacy and Numeracy adult qualifications have been phased out, replaced by the English and Mathematics Functional Skills, which are designed to meet the needs of adult and working life. Alongside the Functional Skills Curriculum, concern around the appropriateness of existing qualifications to meet the needs of some learners prompted the introduction of standalone maths and English Qualifications and Credit Framework (QCF) Units in 2012, allowing providers greater flexibility in offering learners a more personalised offer. The next stage (highlighted in a Government review, in 2011², aimed at making the provision more effective) is to incentivise providers to ensure that learners make progress.
- 1.3 The established system involves payments to providers on the basis of qualifications achieved by learners. The official guidance states that:
- "Learners must be enrolled on a level of learning that is beyond that to which they are assessed. For example, if a learner is assessed as being at entry level 3 they must be enrolled on at least a level 1 qualification....Learners must not simply be accredited for knowledge they already have.... [Providers] must carry out a thorough initial assessment to determine the level at which the learner is currently working so [providers] can decide which level they will enrol onto."*
- 1.4 Even so, the system of payment by learner achievement (regardless of the skill level they started from) has raised concerns that some learners are not being sufficiently stretched, and there are reports of 'teaching to the test' by providers rather than fully developing the skills of learners. As a result providers could be paid for enabling learners with unaccredited Level 1 skills to achieve a Level 1 qualification, whilst at the same time making no real gains in skill levels.

¹ <https://www.gov.uk/government/publications/2011-skills-for-life-survey>

² Department for Business, Innovation and Skills (2011) New Challenges, New Chances [online] Available. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/145452/11-1380-further-education-skills-system-reform-plan.pdf

- 1.5 Following the moves signalled in the coalition government's Skills for Sustainable Growth³, a total of 17 Pathfinder pilots,⁴ based in a range of different provider settings (see Chapter 3) were set up to explore the feasibility of a funding system based on the distance travelled (rather than just qualifications gained) by learners in order to incentivise providers to maximise skills acquisition. It is accepted good practice for adult learners to receive an initial assessment,⁵ prior to starting an adult literacy or numeracy class and there are assessments that accredit learners' skills level at Entry level 1, Entry level 2, Entry level 3, Level 1 and Level 2. However, there are no recognised national tools that enable measurement of distance travelled by learners, in terms of skills gain, from the starting point to the end of their funded learning.
- 1.6 In establishing the pilots, providers identified, adopted or modified tools (both initial assessment³ tools (commonly used to provide an indicative level), and diagnostic tools (frequently used to establish a full skills profile) that they hoped would provide the evidence they needed of the distance travelled by individual learners, especially those working at a low level. These learners may make significant progress in a year but yet not move up a level (such as from Entry level 3 to Level 1). While there are many competing initial assessments on the market, recent unpublished research (Brooks et al., 2012)⁶ indicates that they are not consistent and that the skills levels indicated by one test are not necessarily matched by those apparently demonstrated by others. For the providers, however, there is a need to be sure that initial assessments are accurate and that post-tests record the exact skills gains. In a funding system based on payments by distance-travelled, for instance, colleges could be penalised not because their work did not lead to skills gains, but because that gain could not be measured.

Project aims

- 1.7 The main aims of the study were to:
- undertake a detailed process evaluation of the Pathfinders, examining the tools they used and the practices they had adopted (in order to understand their rationale and the basis for their anticipated outcomes) and, ultimately, to assess their impact and effectiveness
 - assess the extent to which the tools being used were a) fit for purpose and b) enabled distance travelled to be measured accurately

³ Department for Business, Innovation and Skills (2010) Skills for Sustainable Growth: Strategy document. [online] Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/32368/10-1274-skills-for-sustainable-growth-strategy.pdf

⁴ There were initially 18 but one provider withdrew in the early stages of the pilot.

⁵ These are often supported by more detailed diagnostic tests.

⁶ Unpublished advice note to BIS.

- draw conclusions as to the comparative robustness of the approach for a funding model based on learner distance travelled
- make recommendations on what is needed to enable distance travelled to be captured for funding purposes

1.8 In order to meet these aims, the study:

- designed a data collection and analysis strategy that would enable the research team to assess the extent to which the tools used by the Pathfinders were able to capture distance travelled
- implemented a programme of qualitative fieldwork to develop a deeper understanding of the process of capturing distance travelled and the anticipated benefits of such an approach
- developed an approach to capturing cost data from colleges to contribute to an understanding of how (and if) different funding approaches might benefit learners in the future

The Pathfinder pilots

1.9 A total of 17 Pathfinders were established in September 2012, comprising colleges of Further Education, independent training providers and voluntary and community based providers, covering discrete geographical settings in the north, south and Midlands, as well as those with a national coverage. A range of tools were used in the pilots, including established commercial tools, and those developed in-house. An overview of these pilot provider settings is presented in Chapter 3.

1.10 Following an inception workshop in early September and initial exploratory period, during which many of the pilots re-visited the tools they planned to use (in some instances adopting a different tool) and the numbers of learners with whom they decided to use the approach, the study team began the collection, collation and analysis of both qualitative and quantitative data.

Structure of this report

1.11 The research approach and methods adopted for the evaluation are outlined in Chapter 2. The respective findings are explored in Chapter 3 (Qualitative findings); Chapter 4 (Assessment of tools) and Chapter 5 (Quantitative analysis). Chapter 3 looks at the rationale providers gave for their involvement as well as their perceptions of the learner cohort; the thoughts behind their choice of assessment tools; the approach providers have adopted towards assessment; any changes that were made in relation to the assessment process, and the recording and sharing of data; and the impact of the use of a distance-travelled approach (whether on staff or learners, on costs, or on course provision). Chapter 4 provides a critical appraisal of

the assessment tools that were being used by the Pathfinder pilots and the implications this appraisal has for both the quantitative analysis and for their longer-term use in a distance-travelled approach.

- 1.12 Chapter 5 provides an overview of the characteristics of the learners involved in the study, providing details on demographics, course-related information and some socio-economic information. It also includes a critique of the data collected and collated by the Pathfinder pilots and insights into the ability of that data to contribute to a robust assessment of distance travelled.
- 1.13 The final chapter, Chapter 6 looks at the implications of these findings (both qualitative and quantitative) for the wider roll-out of a distance travelled approach in the future, with recommendations for providers, for BIS, and the Skills Funding Agency (the Agency).

2. Research methods used

- 2.1 The study involved a detailed process evaluation, examining the various contexts (such as provider settings and range and type of learners) and mechanisms in place (including the factors affecting the choice of assessment tools and the validity of those tools), This approach was adopted in order to assess which Pathfinder approach appeared most effective in measuring distance-travelled and to provide a basis on which a funding model could be developed.
- 2.2 The research approach is described in the remaining sub-sections of this chapter and, in brief, includes:
- contributions to provider workshops (organised by BIS)
 - quantitative analysis of learner-level and provider-level data to explore the impact and effectiveness of different measures of skills gain
 - qualitative analyses to look at process, implementation, fidelity of approach, including case studies of providers
 - examining the impact of using a distance-travelled approach on provider funding

Provider workshops

- 2.3 Three workshops for providers were organised by BIS during the course of the pilot study. In the first of these, in September 2012, the research team presented an overview of the evaluation and circulated and discussed a draft Excel workbook, which included a series of spreadsheets designed to collect data from the colleges on learners, their courses, their pre-tests and post-test scores and their subsequent qualifications. These spreadsheets were refined following the workshop and then circulated (for completion) to all Pathfinder pilots (see quantitative analysis below). The second workshop (February 2013) looked at progress in implementing the Pilot, explored the available data and discussed the emerging findings relating to the use of different tools and the process of assessment. The final workshop (September 2013) examined the implications of the research findings for policy and practice.

Qualitative analysis

- 2.4 This element of the study had three main foci: evaluating processes, evaluating impacts, and evaluating tools. In particular, the project sought to understand how different approaches had (or had not) worked, in what situations, with which particular groups of people, and why. In order to do so, we adopted a phased approach, with two tranches of face-to-face interviews with each of the pilot

providers (including the project lead/Pathfinder coordinator and staff implementing the testing or working with the in-scope learners) and a critical review of the tools being used by the Pathfinder pilots. This report provides a synthesis of our findings from these activities.

Fieldwork in Pathfinder pilots

- 2.5 Over the course of the evaluation we conducted two rounds of fieldwork involving a series of visits to the 17 Pathfinder sites. A total of 121 interviews were conducted, some in one-to-one discussions, others in group settings. A detailed breakdown of the roles and responsibilities of interviewees is provided in Table 2-1.

Table 2-1: Roles and responsibilities of interviewees

Role/Responsibility	Number of Interviewees		Total
	Round 1	Round 2	
Admin Staff	1	2	3
MIS Officer	3	2	5
Tutor	7	12	19
ICT Staff	1	2	3
Finance Staff	1	1	2
Senior Management	10	11	21
Middle Management	32	33	65
Skills Funding Agency Advisor	3	0	3
Total	58	63	121

Source: SQW

- 2.6 As illustrated in the table above, interviewees included a wide range of staff from provider settings and, in two areas, local advisors from the Skills Funding Agency. While over half (54%) of our total interviews were conducted with staff at middle management level, we interviewed representatives of all the provider groups likely to be affected by the pilot process including administration staff, those leading on Management Information Systems (MIS) and ICT staff, course tutors, finance officers, and senior leaders. Staff interviewed at a middle management level were responsible for a wide range of activity including Skills for Life, Employability, Family/Community Learning and specific maths and/or English responsibility (sometimes referred to as Adult Literacy and Numeracy).
- 2.7 The range of interviewees, and the variety of their professional responsibilities (which included employer provision, professional development and the development

of assessment materials), meant that the study was able to gain an overview of the different ways in which the pilot was perceived, including the rationale for taking part, the impacts, and the challenges encountered.

- 2.8 The interviews were conducted using semi-structured interview guides, exploring context, early issues with implementation, providers' experiences of capturing distance travelled, including: resource use; the perceived strengths and weaknesses of the approach; the perceived strengths and weaknesses of each measurement tool; benefits of measuring distance travelled; implementation challenges; challenges to policy integrity; unintended consequences; and a range of other impacts and experiences.

Critical review of assessment tools

- 2.9 Alongside the qualitative fieldwork, each of the tools used by the Pathfinder pilots (a total of 41 different instruments) was accessed (in paper or online form). Initially, it appeared that there were 52 instruments, but closer inspection revealed that some of the materials provided were answer booklets and guidance materials, and others were recording sheets for in-house tests. These tools (once accessed) were subjected to a critical academic review. It should be noted that five of the 41 were not made available by the publishers in time for the review, so the review was based on 36 tools in total.
- 2.10 Normally, such a critical review would include an examination of reliability in the statistical sense so that, if tested twice on the same instrument (or parallel forms), individuals would, within narrow limits, achieve the same score. Assessing the reliability of test items is part of the development process for psychometric tools (such tools would only include reliable items). This aspect can only be judged statistically (and requires access to the development data). Given the status of the tools (some of which were developed in-house) and the fact that most had not been designed as pre-test/post-test tools, this element could not be done for this review and so was excluded (see Chapter 4).
- 2.11 This review was designed to ascertain, for each tool:
- **Validity** – did the items measure (or appear to measure) what they were intended to measure?
 - **Manageability** for those being tested and for those administering and scoring the tests
 - **Usefulness** - that the results provide information on achievement and (where appropriate) progress that is clear and meaningful to those tested and their teachers, providers, funders and for policy-makers.

2.12 The review sought to ascertain whether the tool existed in statistically equated parallel forms (and so could be used effectively to measure distance travelled) and the length and precision of the measurement scale. For a national survey of adults' skills a short scale with (preferably) publicly intelligible steps is often sufficient (in the current context in England, the relevant scale for surveys would be the Qualifications and Credit Framework - QCF levels). For useful measurements of progress, however, (and given that many Entry level adult learners make very slow and modest progress, for a variety of reasons) a more finely gradated scale is essential. The outcomes of this review of assessment tools are summarised in Chapter 3.

Quantitative analysis

2.13 The approach adopted for the quantitative analysis was designed not only to measure the distance travelled of the learners in the study but also to provide information on the ability of the tools selected by the Pathfinders to measure that distance travelled accurately and reliably. Unless a decision is taken at a national level to adopt a single tool approach to initial and post-course testing, any future funding model based on distance travelled would need to know that the outcomes identified using different tools were comparable with those obtained using other tools (and have a method whereby this could be assessed).

2.14 Assessment tools are normally evaluated with regard to their construct validity (that is, are they valid and do they measure what they should measure) based on four main criteria:

- **convergent validity** - comparing tests with an established measure, such as an existing research tool
- **predictive validity** - which looks at whether the tests accurately predict outcomes in formal examinations – in this instance, the study would need to look at whether the distance travelled between pre-test and post-test accurately predicted future outcomes such as completion and achievement in Functional Skills or GCSEs, for example,
- **concurrent validity** - which assesses whether the tests can differentiate accurately between different groups of learners or between different sub-levels
- **discriminant validity** – which would look at correlations between results of tests for two *different* skills, such as literacy and numeracy. Where *low* correlations occurred this would be evidence of discriminant validity, indicating that the tests measured what they were intended to measure.

2.15 The evaluation methods that could be deployed for the study were partly dictated by the way in which the pilot was established. The Pathfinders were set up with

providers administering and collecting test scores for their chosen tool in order to see whether any of these tools could be used for measuring distance-travelled under any new funding model. They were not required to use an established research tool as well, since that would have meant learners taking two further tests in addition to those they would complete for the pilot. This means, however, that the assessment of the convergent and discriminant validity of the tools that colleges were using was very difficult. Equally, none of the chosen tools (as yet) have been equated with GCSE scores or the new Functional Skills outcomes, which means that their predictive validity has not yet been established.

2.16 The decision was taken, therefore, to focus on identifying concurrent validity and on overall construct validity (whether the test measured what it claimed to measure). In order to carry out these analyses, the team needed access to a range of both quantitative and qualitative data, including information:

- **on learners:** such as gender, age, ethnicity, first language, self-reported or known disability, prior highest qualification overall, highest qualification in English and maths and employment status. Some of this data is already included in the Individualised Learner Record (ILR), but the timing of the study meant that we needed to approach the providers directly for individual learner-level data extracts. Other data, such as first language, is not normally part of the ILR dataset, so was collected by the providers directly from the learners before being passed on to the research team.
- **on the pre-/post tests and test scores:** in addition to the raw test scores, which included collecting some supporting qualitative information on the ways in which tests were administered (under what conditions and in what form - electronic or paper), data on the length of time tests took and the resources required to implement them, and a critical assessment of the range of levels and sub-levels they measured
- **on the teaching process:** the quantitative data included information on the number of guided learning hours at a course level, the frequency of classes and the intensity of the courses, as well as some qualitative data on type of teaching/learning strategy used and an exploration of the teaching and subject-based (maths and English) qualifications of staff (where known)
- **on providers:** quantitative indicators that are being collected over the course of the study include Ofsted ratings, success rates (that is, the proportion of starting learners who stayed on the course to the end and achieved a qualification – however the qualification is defined) and measures of deprivation for the area. These measures were augmented by additional qualitative data.

- **on costs and benefits to providers:** the study collected information (where available) on actual costs (of purchasing the tests, of staff time for administration and marking, overheads etc.), any hidden costs that could be identified (such as the cost of providing support for LLDD learners) and indicators of impact that providers identified, (such as reduced drop-out, increased attendance or increased achievement of qualifications).
- 2.17 Following the first provider workshop in September 2012, all providers were sent an Excel workbook for completion. This included an agreed series of spreadsheets on learners and on their courses.
- 2.18 At that stage in the study we had access to some of this data, though not all. Some of the data was time dependent (post-test scores, in particular were not available for a number of months after the courses were completed) and some had not been fully collated by the providers (including information on courses; not all learners for whom we had pre-test data had been assigned to a course when the initial data was sent to us, for example) and some proved more problematic for providers to collect (including data on first language, which is not a field in the ILR, and employment status, which is a standard field in the ILR, though not always comprehensively reported).
- 2.19 These various data were used in a number of different ways to enable:
- the profiling of learners in the study
 - an evaluation of construct and concurrent validity
 - an assessment of distance travelled, exploring the mean distance travelled for each provider and/or type of tests used and looking (where possible) at comparisons both within providers (differential progress by level, subject or learner group, for example) and between providers (where comparison of test outcomes proves possible)
 - the contextualisation of statistical analyses; in order to assess which approach appears most effective (and most economic) at measuring distance-travelled it is important not only to look at the trajectories for different learner types, but also to include (where possible) an examination of the impact of teaching processes and provider settings, for example. This approach was intended to enable us to examine the relative impact of the Pathfinder approach on learner outcomes.
- 2.20 While providers were asked about their costs during the initial qualitative interviews (see above), most said they had not been monitoring costs (whether in monetary or time terms) in detail at that stage and had no data to hand. During the second round of fieldwork in March to May 2013 we explored this with the providers who still found

it difficult to attach costs to the pilot. Instead we asked providers to identify any additional activities that they had undertaken as a result of the pilot (these included marking test scripts, collating/validating test data, and managing the implementation of the pilot) and to estimate the additional time they had spent on these activities. This data was analysed by looking at the relative time spent (per staff member and grade), and calculating the relative costs per activity and per learner, based on the salary levels of the members of staff⁷.

- 2.21 This report provides an overview of the profile of learners in the study and discusses the extent to which the data which the research team have been provided enabled the conduct of the planned analyses.

⁷ Salary levels were estimated on the basis of information submitted by the providers.

3. Qualitative evaluation

3.1 The findings from the 34 visits (two per provider) that took place between October and December 2012 and April to June 2013 are explored below. In particular, this chapter reviews:

- the key characteristics of the providers engaged in the Pathfinder
- the rationale providers gave for their involvement (including any identified local priorities)
- providers' choice of learners to include in the pilot
- the thoughts behind their choice of assessment tools (and the extent to which these had been used before by the provider)
- the approach providers have adopted towards assessment, highlighting any differences in approach
- any changes that have had to be made in relation to the assessment process, the recording and sharing of data and any challenges that have arisen
- the impact of the use of a distance travelled approach (including on staff and learners, on costs, and on course provision)

3.2 It should be noted that, at the time the initial field visits were conducted in 2012, many providers were still in the early stages of developing their overall approach to Pilot activity. Not all had a final and definitive target list of learners to be included in the Pilot, for example. Some were still confirming which assessment tool they intended to implement, and finalising the process for implementing the pre- and post-assessments, with further amendments taking place as late as February 2013. This reflects the fact that, while providers are used to conducting initial (and diagnostic) assessments of learners, the use of a pre-test/post-test approach was not common, nor was the wider collation and reporting of such data outside the course team (or teams) in provider settings.

Characteristics of Pathfinder providers

3.3 In Table 3-1: below we outline the characteristics of the 17⁸ providers that agreed to participate in the Pathfinder.⁹ The pilot sites comprised five in the East Midlands¹⁰,

⁸ 18 providers originally agreed to participate in the Pathfinder,

⁹ The names have been included here as they are known to the steering group. We will need permission from the providers to include them in the final (public) report.

¹⁰ One of the Pathfinders was a national provider but their pilot activity is focused on sites in this region

four in the North East, two in each of West Midlands and North West, two national providers, and one site in both London and Yorkshire and Humberside. The majority of pilot sites were Further Education colleges/providers, with five sites representing other types of provider (private training providers and not for profit organisations).

Table 3-1: Characteristics of Pathfinder providers

Pathfinder site	Nature/size of the pilot provider	Catchment profile
Newcastle and Gateshead Skills for Life Service	This is a joint Pathfinder between Newcastle upon Tyne City Council and Gateshead Councils. In Newcastle responsibility for adult and community learning is devolved to Newcastle City Learning. The partnership is responsible for enrolling around 10,000 adult learners each year. A wide range of courses are available including apprenticeships, entry-to employment, work-based, ESOL, Train to Gain, and Skills for Life. Gateshead Adult Learning and Skills Directorate is responsible for the enrolment of around 9,500 learners each year across over 100 venues. Courses include family learning, Train to Gain, first steps and neighbourhood learning.	Newcastle and Gateshead cater for the needs of a wide range of learners from non-accredited learners through to those learners looking to progress into higher education. Most learners are at Entry Level. A significant proportion of learners are referred to the services by Jobcentre Plus.
Middlesbrough Community Learning	The Skills for Life Service is council-run community-based provision is offered in a variety of venues including community centres, schools and pre-school settings, The service delivers a wide range of programmes that include apprenticeships, Foundation learning, family learning, learning for pleasure and employability courses. Over the course of 2012/3 it is expected that c.400 learners will have registered for English and Maths courses.	Courses are offered for non-accredited learners through to level 2. The majority are at the lower end of this attainment spectrum. Just under three quarters of all learners are self-referrals. Around one third are referred from Jobcentre Plus.
City of Sunderland College	Sunderland is a large college (c. 20,000 students) offering a range of 16-19, adult and HE provision, as well as work-based learning and modern apprenticeships. They have several sites across the city, including 'The Place' in the city-centre which focuses on the delivery of Skills for Life. Relationships are also in place with partner providers supporting delivery of entry level and above courses from a number of different locations.	One quarter of the city's working population is without a Level 2 qualification. There is also an emphasis on preparing people for employment. Most learners are self-referrals.
North Tyneside Council Adult Learning Alliance	A council run community-based provision delivered from 40 sites catering for c. 6,500 learners each year. The Alliance offers a wide range of courses including apprenticeships, family learning, modern languages and ICT, alongside Foundation Learning, employability, and adult literacy and numeracy.	Courses are delivered for non-accredited learners through to GCSE. The majority of learners are self-referrals although a significant minority of referred via Jobcentre Plus. The majority of these learners have low skill levels on entry.

Pathfinder site	Nature/size of the pilot provider	Catchment profile
Learndirect	Learndirect is a large national provider made up of 57 independent sites, collectively delivering courses to more than 200,000 learners each year. Most learners are enrolled on workplace learning courses. There is a substantial uptake of literacy and numeracy courses. A large number of learners also study for ICT qualifications.	A significant proportion of learners enrol on a course delivered by Learndirect following a period of disengagement from education or training. The majority of learners are self-referrals.
PPDG Training (Pertemps)	PPDG is a large Birmingham-based provider with c. 3,500 registered learners at any time. The majority of learner's access work-placed learning or apprenticeships, but literacy and numeracy courses are on offer through community-based locations.	A significant proportion of learners are referred by Jobcentre Plus, as well as workplace programmes or the National Offender Management Service. The majority of learners are at Level 2.
Derbyshire Learning and Development Consortium	The pilot site comprises Derby Adult Learning Services (DALs) and Derbyshire Learning and Development Consortium (DLDC). DALs operate from 10 sites within Derby City, with c.500 (of 6,000 learners) engaged in Maths and English courses. DLDC offer a wide variety of courses that aim to support those furthest from the labour market into employment, some of which comprise numeracy and literacy elements.	DALs have a diverse catchment area and this is reflected in their learner profile. DLDC receive a large number of referrals from Jobcentre Plus (many with ESOL needs), and also has many learners with learning difficulties and disabilities.
Hackney Community College	Hackney College is a large provider operating from a single site and offering a range of courses for both adult learners and 16-18 year olds. It is the only general further education college in the borough. While its main site is in Shoreditch, the college also works in over 40 community and employer venues. The college offers a variety of courses for adult learners including vocational courses such as hospitality and catering, apprenticeships, ESOL, and Train to Gain.	There is high demand for ESOL provision, and many students are bi-lingual. Many learners have existing qualifications from their country of origin and so are less in need of basic literacy and numeracy than language courses. The college also has a high proportion of learners with learning difficulties. The college caters for a wide range of learners from those looking to take entry level courses (the majority) through to those looking to enter higher education.
Hull College	Hull college is one of the largest general further education colleges in the country. It operates from a number of sites across the city. Courses are offered in 15 sector subject areas catering for learners at pre-entry level through to HE. The college has a	A large proportion of learners are highly disadvantaged, both socio-economically and educationally.

Pathfinder site	Nature/size of the pilot provider	Catchment profile
Telford College of Arts and Technology	<p>developed adult learning offer including apprenticeships, entry to employment programmes, and the DWP Work step initiative, the college caters for around 1,200 – 1,800 adult literacy and numeracy learners (aged 19+ at any time).</p> <p>Telford is the only large general further education college in the local are. Based on a single site, it also offers community and work-based learning. The adult learning offer includes Apprenticeships, ESOL and support for the unemployed, with a rolling programme for most courses.</p>	<p>There is a very diverse learner profile with a wide age range, and differing literacy/numeracy levels from pre-entry through to HE. They have a strong relationship with local businesses and provide bespoke English and maths training for them outside of apprenticeship programmes.</p>
Oldham College	<p>This is a medium-sized college with 10,000 registered learners. The college is based on one site and specialises in vocational provision from pre-entry to advanced level. There is a developed adult learning offer with provision for Apprenticeships and ESOL.</p>	<p>Learners are drawn in from a highly deprived catchment area. There is high demand for ESOL provision. A significant proportion of learners are referred by Jobcentre Plus.</p>
Blackburn College	<p>Blackburn college is a large general further education college with 15,000 registered students. A wide range of courses are delivered for learners from pre-entry to post-graduate level. Courses in Maths and English include part-time functional skills adult classes, cross-college courses to support vocational learners (mainly under 19), GCSE routes for adults and ESOL. Many courses are 'roll on, roll off'.</p>	<p>Learners are drawn from an ethnically diverse and highly deprived community. There is strong demand for ESOL provision. Some learners are referred by Jobcentre Plus.</p>
TBG Learning	<p>TBG Learning is an independent not-for-profit organisation that provides training, employment, health and social care, as well as commercial services. They are a national organisation, but pilot activity is focused on sites in Derby, Allerton and Chesterfield. They offer a range of courses for adults and young people, primarily centred on employability and employment skills.</p>	<p>The catchment varies site by site, Derby having a very ethnically diverse community compared with Chesterfield and Allerton, which are less diverse with an older population profile.</p>
Derby College	<p>Derby college is a large provider operating across four major campuses. The college offers vocational courses across 15 subject areas. The majority of enrolments are at advanced level.</p>	<p>The catchment area is ethnically diverse and many learners have English as a second language.</p>

Pathfinder site	Nature/size of the pilot provider	Catchment profile
Leicester College	<p>The college also offers Apprenticeships and Train to Gain Courses. Literacy and numeracy are integral to what the college offers, but the main focus is on employment skills and progression. Developing literacy and numeracy skills are deemed a key component in achieving these outcomes.</p> <p>The college operates across a number of sites, with three in particular focusing on the delivery of discrete adult skills programmes. For many learners their main programme of study revolves around English and Maths but many are also involved in other courses, or want to progress onto other courses.</p>	Learners reflect the ethnic diversity of the city, although the profile will vary between sites.
Access Training East Midlands (ATEM)	ATEM is a regionally focused organisation that has been providing 'on the job' training in the East Midlands for more than 30 years, primarily through work-based learning and apprenticeships. They have two sites, based in Derby and Nottingham. They offer a wide variety of courses including discrete functional skills courses, which can also be integrated with other courses when appropriate.	ATEM have a large catchment area, and a high proportion of learners are in employment or on apprenticeships, meaning they may be doing courses alongside working.
Tribal Education and Trades Unions Congress (TUC)	Tribal is a large provider of technology products and services to the education, training and learning markets, and they have developed an e-learning platform for literacy and numeracy (including assessment tools). The TUC has a large number of Union Learning Representatives (ULRs) who are responsible for helping union members achieve education and training goals. Unionlearn is the learning and skills component of the TUC and has a network of over 400 learning centres. The two organisations have a history of working together in the field of adult basic skills development; this history continues through Unionlearn's use of Tribal's e-learning platform.	All learners are union members in England and are accessing a work-based learning offer.

Source: SQW Fieldwork

Rationale for engaging with the distance-travelled Pathfinder pilot

3.4 Providers were asked to explain their motivation for volunteering to pilot the new approach as part of the Pathfinder. In most cases there were several factors that had influenced the decision to participate. The main reasons (listed in order of the relative frequency with which they were mentioned) were as follows:

- **Keeping up to date with national policy developments** – most providers believed that participation at this stage would be beneficial because it would keep them up to date with emerging policy, giving them an opportunity to influence policy and better understand its longer term implications.
- **Personalising learning** – pilot sites were interested in testing the concept of ‘distance travelled’ as a means of contributing to personalising learning, particularly in being able to show learners the progress they have made. Linked to this was a concern that existing training and education systems were too focused on learners achieving full qualifications, rather than a personalised learning approach which recognises learners’ achievements (such as self-efficacy) outside any formal accreditation process.
- **Current systems favour those learners most likely to achieve accredited qualifications** – providers highlighted a disincentive in the current funding system to work with the most disadvantaged learners (those that were furthest from being able to successfully complete accredited qualifications). This was perceived as being inherently unfair to those learners who were most in need of support with English and maths. At the same time, providers often felt pressure to work with those closest to being ‘qualification-ready’ because, from a financial perspective, this led to quicker and easier gains for their organisation.
- **A chance to test new assessment tools** – providers said that they had existing initial assessment and diagnostic tools (both in-house and commercial) which they were using to assess learners. In several cases, however, the Pathfinder was identified as an opportunity to test new assessment and diagnostic tools they had not previously used but wanted to review. For some, the Pathfinder also provided an opportunity to run two assessment tools alongside each other in order to compare them and the extent to which they might help with assessment or measuring distance travelled.
- **An opportunity to strengthen existing processes** – sites remarked that the pilot had provided an opportunity to strengthen their existing processes governing the selection and delivery of the pre-test, as well as data collection and management. Such processes are increasingly being scrutinised by

Ofsted and the Skills Funding Agency and the pilot was viewed as a way of showing staff across the organisation the value of them.

- **Improving Ofsted ratings** –four providers felt that participation in the pilot would be beneficial as it would help them demonstrate that they were taking a holistic approach to learner progress by looking at distance travelled, which would have a positive impact on elements on their Ofsted rating.

3.5 Providers often cited more than one of these reasons but, in the main, appear to have been motivated by a combination of wanting to be part of the policy development process and improve both the assessment of learners and provision that they offer them.

Learner cohorts included in the Pathfinder

3.6 Table 3-2 (below) sets out the number of learners with whom the 17 providers in Pathfinder sites anticipated carrying out pre-and post-assessments at the start of the pilot (4,599), and the assessment data subsequently received following learner entry and at course completion. While a broad balance was maintained between the numbers of learners taking an English (1,344) or maths (1,286) course, just over half (57%) of the assessments envisaged at the start of the pilot were completed at Wave 1 (2,630 at course entry). The proportion of the cohort for whom both pre- and post-test data subsequently became available was less than one-quarter (1,077) of those that providers had originally estimated, and included just two-fifths of those who completed the initial assessments.

Table 3-2: Learner assessments

Anticipated completions	English	Maths	Total
Anticipated completions	2,286	2,313	4,599
Wave 1/pre-assessment data available	1,344	1,286	2,630
Wave 2/post-assessment data available	502	575	1,077

Source: Pathfinder pilot sites

3.7 The reasons for this attrition were many. At the outset, some providers noted that they had been unable to bring on board the number of tutors they had initially anticipated, reducing the number of initial assessments collated from learners. Equally, providers reported that many of the courses they included within the pilot

were roll-on, roll-off in nature, which had made it difficult for them to predict both entry numbers and successful completions.

- 3.8 In general, the natural course attrition rate (which varied between providers) was commonly around 15% of those who started a course, but the attrition in post-test assessments noted at Wave 2 (course completion) was considerably higher, suggesting that in-course attrition does not entirely explain the lower numbers of post-test assessments. Even when learners had successfully completed their course, a number did not complete Wave 2 post-test assessments, partly as a result of logistical challenges, such as finding space for an additional post-test within an established scheme of work, administering a post-test outside of a classroom setting, or bringing learners back into settings after they had completed their course. While these are issues that would need to be addressed in the adoption of a pre-test/post-test approach, of more concern, perhaps, might be the impact of an additional test on learner engagement. Two of the pilot providers said that they felt that the requirement for learners to complete a post-test had contributed specifically to a higher attrition rate on their courses.

Table 3-3: Completed pre- and post-test assessments by level and subject

	English	Maths	Total
Entry	237	219	456
Level 1	159	205	364
Level 2	106	151	257
Total	502	575	1,077

Source: Pathfinder pilot sites

- 3.9 From the outset the split between English and maths, and also between skills level, was broadly even, with the number of learners completing Entry Level assessments in maths and English anticipated to be slightly higher than those who would be assessed at Level 1 or Level 2. A more detailed breakdown of the profile of the 1,077 learners for whom assessment data was finally received is presented in Chapter 5.
- 3.10 During the initial visits to pilot providers, the characteristics of the learners they planned to recruit to the Pilot, and the types of courses they would be studying, were discussed. This varied across providers, but, in general, the presumption was that learners would comprise a mixture of self-referrals and referrals from other agencies (included 'mandated' learners via Jobcentre Plus). The courses included in the Pilot also comprised a mix of learners from roll on, roll off provision, as well as those enrolled on longer courses with fixed start dates.
- 3.11 A few providers were explicit in their intention to focus on learners who were undertaking longer courses (e.g. 24+ weeks). The rationale for this was two-fold. First that it provided a long enough period for learners to make progress that might

be identifiable in any assessment of skills gain. Second, there were concerns that learners might be unwilling or reluctant to undertake a second assessment within a short timeframe (but would do so over a longer period). Providers also intended to include learners who were accessing an array of course-types, both by level, including those from Entry-Level through to Level 2, by mode of delivery, including provider-based, community-based (e.g. outreach location) and work-based provision, and by type of course (both discrete maths and English courses as well as integrated maths and English provision as part of work-based learning or vocational courses)

Providers' choices of assessment tools and implementation models

- 3.12 The providers in the pilot reported two main ways in which they had traditionally assessed the skills of learners. The first approach was to undertake a single initial assessment (in some cases using a diagnostic assessment rather than an initial assessment tool) and then to refer the learner to appropriate provision. The second was to conduct an initial assessment to identify the appropriate level for the learner, followed by a more in-depth diagnostic to identify in more detail the specific needs of that learner to personalise English and maths provision to their needs. In many cases, in-house tools had been developed either for initial assessment or the more in-depth diagnosis because the providers felt that existing products did not meet their needs, and in particular the demands of the Functional Skills Curriculum. Just under two-thirds of the providers, it should be noted, were using a mix of commercially available or public-sector developed tools, alongside in-house tools.
- 3.13 Providers were also asked about any existing mechanisms that were in place to assess learner's progress or 'distance travelled'. In practice, data on learners' progression had not previously been collected systematically or centrally by providers and any collated information has been largely qualitative in nature. Prior to the Pilot, none of the participating sites had specifically administered a post-test, relying on external qualifications to measure progression. Moreover, as some providers noted, the primary outcome many were seeking to achieve for their learners was employment, rather than progression in maths or English *per se*.
- 3.14 As outlined earlier in this Chapter, one of the factors that influenced the decision of some providers to participate in the Pathfinder was the opportunity to test new initial assessment and/or diagnostic tools for the purpose of measuring learners' distance travelled. Other changes that were made to assessment approaches as a result of engagement in the Pathfinder pilot, related primarily (as might be expected) to the introduction of a second assessment once a learner had completed their course, or mid-way through, with the aim of being able to identify their progress. In some instances providers had routinely undertaken diagnostic assessments with learners in the past (focusing on an initial assessment to direct learners into appropriate provision) but introduced a pre- and post- diagnostic assessment for the Pilot.

Challenges arising from the adoption of a distance travelled approach

- 3.15 Providers indicated that they had faced a variety of challenges in delivering the distance travelled approach. Some of these were specific to the pilot, but many had implications for any wider adoption of the approach and deserve fuller consideration. These included the suitability of existing tools to measure distance travelled; the challenges of collating learner information; the availability of appropriate IT infrastructure; the challenges of supporting the administering a post-test, and that of ensuring the buy-in of tutors and learners into the approach.

The suitability of existing tools to measure distance travelled

- 3.16 Providers used a wide variety of assessment tools, often including a mixture of in-house tools alongside those produced commercially. They indicated that in-house tools were mainly used to fill in the perceived gaps left by existing assessment tools and, for the purposes of an initial assessment, provide a solid basis on which to decide the level at which learners should be working. Just over half of the sites expressed concern that the majority of commercially available tools failed to account adequately for the type of skills required by the new Functional Skills Curriculum; others indicated that they were not sensitive enough to enable differentiation of skills between low-level learners (at Entry and pre-Entry levels),
- 3.17 Fifteen of the seventeen pilot providers felt that the process they were using as part of the pre-test provided an effective way of ensuring that learners were entered on the correct course. There was concern, however, around the use of these tools as the basis of measuring distance travelled. Setting aside issues around the comparative validity of the results obtained through the use of in-house tools, sites questioned the lack of parallel forms for the published tests. While some providers used computer adaptive tests (meaning that, in re-sitting a test, learners would be asked a range of different, but similarly levelled questions), most did not and so required learners to retake the same assessment for both initial and post-test (with the accompanying concerns about practice effects).
- 3.18 Variation in the approaches used by providers to select the course on which a learner should be enrolled, and how to establish their skills profile, also had a bearing on the subsequent selection of the test that would be used to measure distance travelled. While many had originally considered using their initial assessment tool at the post-test stage, some subsequently queried whether it would be sensitive enough to pick up learner progression (see Chapter 5), particularly in

cases where learners had a 'spiky profile'¹¹ and might be working at varying levels (or on different skills) within maths or English. In some cases, therefore, providers used the diagnostic tools both to identify learners' development needs and to assess the progress they had made.

Collating learner information

- 3.19 Sixteen of the seventeen providers (94%) indicated that the collation of learner information had led to an additional administrative burden. While the staff involved in undertaking these tasks varied by site, administrative, MIS staff and tutors commonly bore the majority of this additional burden.
- 3.20 The nature of the burden was also seen to differ depending on the nature of the site and the flexibility of existing data collection systems. Whilst the larger providers, mainly general further education colleges, collated learner data centrally (including assessment results), the smaller community providers were more likely to operate systems more tightly attuned to the demands of existing monitoring systems such as the ILR. For these sites the Pathfinder was generally seen to pose a much greater administrative burden, requiring the development of new data collection processes.
- 3.21 The scale of the overall administrative burden also seemed to be affected by the nature of the assessment tool(s) being used by particular sites. Eight of the sites opted to use online or web-based tests; this was felt to have eased the administrative burden, particularly for tutors who did not need to mark any papers. It is worth noting that half of these indicated that, were the pilot approach to be rolled out, this burden would be even less as the full functionality of these tools could be accessed. For example, one site noted that for the pilot they only had access to their tool on a trial basis, which meant that, instead of recording learner scores automatically (as would happen with full functionality), this had to be done manually.
- 3.22 Sites noted that they had been required to collect more contextual information for the purposes of the Pilot than would routinely have been required for the ILR or for normal course provision. While this had posed an additional administrative challenge for providers, and in some cases had required the development of new data collection processes it was unlikely that this information would be required were the pilot approach to be mainstreamed.

Availability of appropriate ICT facilities

- 3.23 Seven of the seventeen sites were using paper opposed to web-based tools¹² as their primary method of assessment. Across these sites, there was broad recognition

¹¹ 'Spiky profile' is term used when learners have different skill levels within a Level. In Entry Level 3 English, for example, a learner might be secure in spelling words of type A and type B, but have problems with words of type C. This is explored more fully in Section 4.

¹²One site used a mixture of both online/web-based and paper assessments

that the availability of appropriate ICT facilities could reduce the administrative burden posed by the need to measure distance travelled. That said, four sites questioned the feasibility/appropriateness of introducing such systems on their site. Those providers working in outreach and community locations noted that they did not have access to IT facilities at which learners could complete web-based/online assessments. In most cases it was considered unrealistic to expect learners to travel to a central location to complete them. It was also questioned if the level of investment to ensure access to IT facilities in all of these settings would be prohibitive.

- 3.24 Some sites also worried that while online or web-based assessment tools were accessible for the vast majority of learners, there were still those that required a paper form. For example, those with sight issues or low levels of computer literacy. Indeed providers queried whether learners' computer literacy had an impact on their test scores – particularly at a time when summative assessments continued to be administered in paper form. It was queried if this could potentially lead to discontinuity between the level achieved on the post-test and then through the summative assessment.

Administering a post-test

- 3.25 As indicated, none of the sites had administered a post-test prior to the introduction of the pilot. Following the pilot, only one of the sites indicated that they planned to continue administering the post-test, if the distance travelled funding model was not mainstreamed. Aside from the additional administrative burden posed by the need to collect and collate post-test information, a number of other issues were felt to impede its effective delivery. Central amongst these was concern around the appropriate scheduling of the test, bearing in mind the variety of courses covered under the heading of adult basic skills from dedicated English and maths courses through to Apprenticeships. It was noted that many of these courses have established schemes of work, making it difficult to re-allocate the teaching team for another purpose. There was also concern around the feasibility of asking learners to come back in to college following the completion of their summative assessment purely for this purpose. The majority of providers had decided to trial the approach on long-courses (lasting two or more academic terms). They questioned whether for shorter courses the need to administer a further test, in addition to those already administered, suggesting that the timescale would render the result meaningless and lead to a worrying erosion of teaching time.

Tutor buy-in to the new approach

- 3.26 Providers identified an additional burden on tutors from participating in the pilot because of the extra assessments that would need to be undertaken. In a small number of cases this was linked to concerns about tutors' ability to take on the additional activity and/or find the space within a packed curriculum to administer the

test. This was particularly the case when working on employer-based provision. Across the pilot sites there was evidence to suggest that reluctance to take on the burden of collecting and collating post-test data had resulted in poorer engagement than had previously been anticipated. In the absence of additional remuneration for the time spent by tutors on these additional tasks there was concern that many tutors would be resistant to implementation of the model.

Learner buy-in to the new approach

3.27 Eight of the seventeen providers) considered learner engagement to be a major issue facing them as part of the Pilot. A number of providers were concerned that, for many of their learners, tests were a considerable ordeal, and frequently the cause of disengagement - particularly for mandated learners. A number questioned how a funding model based-on distance travelled would take account of those learners who accessed learning but did not go on to take a post-test. Providers also indicated that the timing of the second assessment (whether pre- or post- external accreditation) would be critical. Two key points emerged:

- Many felt that any post-test needed to be timetabled to take place before assessment/exams linked to formal accreditation, as they felt learners might be more reluctant to come back and complete a second assessment once they have completed any formal or external qualification tests.
- Others however, suggested that such timetabling was not possible; they noted that, on many courses, the last few weeks were poorly attended, as learners chose to undertake their revision independently.

Implications for the future development of a distance travelled approach

3.28 A number of the challenges illustrated in the previous section have implications for the development of a distance travelled approach. These are discussed below.

Appropriately incentivising providers

3.29 Providers – largely those involved in outreach or community based activities – indicated that a major reason that they had chosen to get involved in the pilot was concern that the existing funding system dis-incentivised providers from enrolling those learners least likely to achieve a full level of progress across a single academic year, or those not yet ready to engage with accredited courses. It was hoped that a distance travelled approach could provide a fairer way of assessing learner outcomes that reflected the efforts of providers to engage those learners requiring more support to successfully reach the next level. Over the course of the round 2 fieldwork, while many of the providers still considered this a laudable aim, there was concern around the capacity of the pilot approach to achieve this outcome. In

particular there was frustration that the existing tools on the market did not provide for the level of granularity to measure progress adequately.

- 3.30 There was also frustration that the increasing attention given by the Skills Funding Agency/Ofsted to the data collected through an initial/diagnostic assessment might, in fact, have a detrimental effect. Instead of encouraging the more effective assessment of learners, it could possibly have the opposite effect, as tutors became increasingly risk averse and sought, through the levelling process, to ensure that learners would not only be placed on a course that they could manage, but also one on which they could be seen to succeed. It was noted that many learners – particularly those with lower level skills – often had a skills profile spanning more than one level, but that any such learners might be put on a Level 1 course that they could complete (so demonstrating some limited progression) rather than on a course that might stretch them. Providers were keen to incentivise behaviour that would promote stretching learners, rather than behaviours that would yield moderate returns.

Effectively motivating learners

- 3.31 Some providers expressed concern that, while the majority of those learners chosen to participate in the pilot had welcomed the opportunity to receive feedback on their distance travelled, the cohort was not representative of all adult basic skills learners, particularly those referred to courses by Jobcentre Plus or the National Offender Management Service. There was some apprehension that were this not to be a pilot, and were post-tests to be a requirement across all basic skills courses, there would be substantial reluctance amongst some learner groups to engage in the process, particularly where sitting the test did not contribute to their passing of the course. In seeking to develop a distance travelled model further, it was felt that greater consideration should be given to the relationship between the distance-travelled measure and the summative assessment. For instance, one provider asked if the distance travelled measure could be used to determine learner outcomes as well as supporting funding decisions. Alternatively, rather than replacing the summative assessment, both the distance travelled measure and the assessment outcome could jointly be used to ascertain the funding outcome.

Developing a distance travelled tool/model

- 3.32 Fifteen of the seventeen providers believed that the tools/processes that they had used (over the course of the pilot) to consider learner placement were effective for that purpose, but few considered those tools/processes to provide an effective model for measuring distance travelled. Most felt that the most sensible solution lay in the development of a new bespoke tool, explicitly designed for the purpose of assessing distance travelled through the Functional Skills Curriculum. A number suggested that substantive savings would be made if this was procured centrally and then made available to local sites. Others, however, questioned if one tool would ever be

sufficient to cater for the wide variety of learners accessing adult literacy and numeracy courses, and wondered whether a suite of tools would be required. This would undoubtedly lead to additional complexity in ensuring an effective read-across from test-to-test.

- 3.33 In any subsequent development of the distance travelled approach, there was also an appetite amongst providers for increased clarity around the process by which distance travelled should be measured, and when tests should be scheduled. Without concrete guidance on these issues, there was concern that the accuracy/validity of any measure would be reduced, and would not allow meaningful comparison.

Data collection requirements

- 3.34 Nine of the seventeen providers indicated that one of the challenges they had faced through their participation in the pilot had been the need to respond to the increased requirement for learner information. The majority indicated that any move towards a distance travelled approach (even if only some of this information was needed) would require changes to their existing monitoring and information system. In order to keep costs to a minimum, sites were keen that clear guidance was provided on precisely what information should be provided, and when and how it should be submitted.

Impact of adopting a distance-travelled approach

- 3.35 Providers were asked about the impact (on their institution and on their learners) of participating in the Pilot, and to consider what the impact would be of scaling-up a distance travelled approach. At the outset, providers anticipated more holistic/personalised learning, greater learner achievement/motivation and additional costs
- 3.36 There was optimism from providers that adopting a new approach to measuring distance travelled could lead to greater personalisation of learning because it would lead to an increased focus on learner need, rather than focusing on equipping learners with the knowledge to get through a specific exam or test (which might be unrelated to immediate or wider needs). By the end of the second round of the fieldwork in June 2013, initial optimism that a distance travelled approach could lead to improved learner outcomes had been replaced with some scepticism that the benefits that had arisen from the pilot were outweighed by the challenges (and indeed costs) of delivering it.
- 3.37 During the first round of fieldwork a number of providers indicated that adopting a new approach to measuring distance travelled could lead to more holistic personalised learning and improved learner outcomes. Over the course of the pilot, the majority of sites appeared to have come to the conclusion that it would be wrong

to overestimate the capacity of this approach – in its existing form - to improve learner outcomes in any substantive way. Sites indicated that involvement in the pilot had ensured a renewed focus on the development of effective Individual Learning Plans (ILPs), though suggested that this would have been something that they would have been looking to do in any case.

- 3.38 Some providers suggested that there was potential for this approach to yield positive gains for learner motivation and achievement, by providing real evidence of their progression, regardless of whether they passed an accredited qualification. Over half of the providers indicated that learners had benefited from the opportunity to discuss the outcome of their post-test with their tutor. A number indicated that learners had said that access to the result of this test had made them more confident about sitting the summative assessment. However, as one provider noted, learners were only likely to feel more confident about sitting the summative assessment where the post-test indicated that they had travelled the required distance. Were the pilot to be scaled-up it is unlikely that all learners would end up in this position.
- 3.39 Just under one quarter of the providers indicated that a major reason behind their decision to get involved in the pilot had been the opportunity to improve the consistency/reliability of their existing assessment processes. Most indicated that they felt that the pilot had led to improvements in the consistency of approach used by tutors to assess the needs of learners, and the tools used to do so. One third of the sites also indicated that they had welcomed the opportunity to improve the quality of data collected on each learner. A few indicated that this could provide a resource in the monitoring of learner outcomes. Given the deficiencies in the existing tools available to providers, a number of tutors questioned if this information was accurate enough to be used for this purpose.

Financial Implications of a distance travelled-approach

- 3.40 Over the course of round one of the fieldwork, a number of sites indicated that they anticipated that implementing a distance-travelled approach was likely to have resource implications. Whilst providers could estimate the *types* of additional cost (if any) they might incur as a result of the new approach, they struggled to quantify this. The difficulty in quantifying additional costs is, perhaps, not surprising given the early stage providers were at when first visited and so further data was gathered during the second round visits. The potential costs of rolling-out a distance travelled approach were found to break down into two principal types; start-up costs associated with adapting existing systems/processes to meet the demands of the new approach, and fixed costs associated with administering the distance-travelled approach. The potential costs associated with the adoption of the approach are discussed below.

Start-up costs of rolling-out a distance travelled approach

- 3.41 Across the pilot providers there was significant variation in the extent to which they felt that the roll-out of a distance travelled approach would require up-front investment. This was found to vary both in terms of the type of provider (e.g. General Further Education (GFE) Colleges, or community providers) and the type of assessment tool used (paper or web-based/online).
- 3.42 Perhaps unsurprisingly the larger providers (often GFE Colleges) proved less anxious about the potential implications of gearing up for the roll-out of the approach. While it was acknowledged that adopting a distance travelled approach would require changes to both their existing monitoring information system (MIS) and their administration processes, most were confident, however, that their existing systems were sophisticated enough to handle the increased burden that would result from it, and that the potential costs of doing so would be limited. Smaller providers – often those engaged in community outreach – expressed much greater concern over the financial implications of moving to the new system. In particular there was concern that adopting a distance travelled approach would require substantial changes to be made to existing MIS systems and administration processes. To give a sense of the scale of the challenge, one provider noted that their existing system collected only the information that was currently required by the Skills Funding Agency (principally through the Individualised Learner Record (ILR)). This does not require the systematic collection of initial assessment or diagnostic assessment data. Both would need to be collected for the purpose of measuring a distance travelled approach. The providers found it difficult to estimate the costs of developing these systems (prior to confirmation of precisely what data would be required), but suggested that they would be likely to run to tens of thousands of pounds.
- 3.43 Start-up costs were also found to vary depending on the type of assessment tool (be it paper or web-based/online) used by a provider. While there was significant concern that, over the long-term, continued use of paper tools would lead to an increase of fixed costs, it was thought unlikely that adoption of the approach would require any start-up investment. However, four of the providers using a web-based/online tool noted that, over the course of the pilot, they had been unable to take-advantage of the full functionality of their assessment package. This was because they had accessed the tool on a trial basis, or had been unable to establish systems allowing the automatic communication of data between their chosen tool and their MIS system. In each case, learners' scores on tests had to be downloaded individually and then input manually onto the MIS system. Adoption of the tool (and upgrading of MIS systems, where necessary) would have enabled sites to gain access to the full functionality of the assessment package, meaning that results would automatically be collated and communicated to the sites' MIS systems. While this automation of processes has the potential to reduce the cost of delivery,

providers noted that achieving system integration was likely to require the recruitment of specialist IT contractors, attracting additional costs.

- 3.44 Three providers indicated that the complexities of managing the transition to a distance travelled approach would require the recruitment of a dedicated project manager. Due to the significance (and indeed scale) of the change some felt that this post would have to be at a senior level. In the short to medium term, five providers felt that they would also need to enhance their MIS/administration capacity. Most felt that a new full-time role would be created while the system was bedding-in.
- 3.45 Finally, providers indicated that they would want to purchase new assessment tools as part of any transition to a distance travelled approach. Depending on the tool selected, and the number of tests required, the costs of supporting this were found to vary markedly.

Fixed costs of rolling-out of a distance travelled approach

- 3.46 Over the course of round one of the fieldwork, a number of the providers indicated that the scaling up of a distance travelled approach was likely to pose additional demands on key staff. During the second round of visits, providers were asked to reflect on the additional tasks that staff would be required to perform, were the pilot to be scaled up, and what the cost implications of this would be.
- 3.47 Unfortunately, while all but one of the providers felt that administering a distance travelled approach would lead to an increased administrative burden, only one had sought to record this in a systematic way (and even then on a partial basis). As such most were only able to provide a rough estimation of the time spent on pilot related activities. In some cases, providers also found it difficult to disaggregate where additional time had been spent on pilot related activities as opposed to changes resulting from the introduction of other in-year developments such as the introduction of the Functional Skills Curriculum.
- 3.48 Nonetheless, despite these caveats, the analysis of this data (full details of which can be found in Annex A) is instructive. In particular, the average cost of undertaking the additional tasks required of administering the distance travelled approach was found to be in the region of £5 per learner. Were the approach to be bedded-in, however, it would be unlikely that additional support would subsequently be required from senior staff. This means that the day-to-day running costs (removing project management responsibilities) would therefore be lower, with an average cost per learner of around £3. This figure serves as a more likely approximation of the additional cost of administering a distance-travelled approach.
- 3.49 Over the course of round one and two of the fieldwork, providers expressed an interest in the benefits that could potentially result from the use of a web-based/online as opposed to a paper tool. As might be expected, the average total

cost of administration per learner where sites used paper tools was higher at around £6. However, further exploration identified a much more complex picture, especially when time spent by senior staff on project management tasks was discounted. In fact, the paper-based tool was found to be less costly at around £4 per learner (an average of around one pound less than for web-based tests).

- 3.50 A number of potential explanations can be given for this outcome. The most significant cost of participating in the pilot was found to be the requirement for tutors to mark, collate and validate the results of post-tests¹³. Of those providers who had used web-based/online assessment tools, two indicated that, in the pilot, they had explored ways to ensure the automatic transmission of test results from their chosen provider to their MIS system, but had been unsuccessful. As such, tutors had been responsible for uploading test information manually. Were the pilot to be scaled up, both providers said they would procure a solution to this problem. Another site indicated that for the purposes of the pilot they had used a tool on a trial basis. This meant that functionality had been reduced and tutors had had to mark the tests by hand. Were the tool in question to be purchased outright, tests would be marked automatically. In all three cases, were the pilot to be scaled-up, the cost per learner would be likely to fall.
- 3.51 Having said this, two sites (committed to the use of web-based/online tools) commented that following the introduction of the post-test, tutors had decided an additional session would be required in order to feed-back their results. This posed a significant additional time burden (around an hour per learner). While sites were resistant to the use of teaching time to conduct feedback on the post-test, all had nonetheless done so for the purposes of the pilot. It was unclear, if the approach were to be scaled-up, whether the sites would continue to do so and if, indeed, an additional session would need to be resourced. This would clearly have cost implications.
- 3.52 In conclusion, while use of web-based/online tools could provide major benefits (in terms of time-saving and allowing the use of adaptive tests), it remains unclear if use of such tools would reduce the overall cost of administering a distance-travelled approach, compared with a paper-based system. Indeed depending on the systems/processes used by individual sites, transition to web-based/online tools could potentially lead to an increase in the overall cost per learner.

¹³The collation and validation of post-test data accounted for on average 37% of the overall total cost of participating in the pilot (for all sites). For sites using a web-based/online tool these tasks accounted for on average 65% of the total cost.

4. Review of assessment tools

- 4.1 In order to measure distance travelled over a relatively short period of learning (courses may last one year or less), an assessment at the beginning (pre-test) and at the end of programme (post-test) needs a sufficient level of granularity in order to be meaningful. Tools that can only measure the move from one level to the next (from Entry Level 3 to Level 1 for example) may not be able to measure the skills gains made by some learners, which may be within a level (an increment in skills or knowledge) rather than between levels. As indicated in paragraph 1.4, no centralised national tests have yet been developed in the UK¹⁴ for measuring distance travelled as an accountability, rather than research, tool. The challenge faced by the pilots was whether the tools that they used to assess learners would be capable of assessing learners' skills at a level of granularity that would allow a reliable and accurate measurement of an individual learner's progress in order to act as the basis for a future funding model based on distance travelled.
- 4.2 The tools selected by the Pathfinder providers were not developed for the pilots, but instead included freely-available nationally-developed tests, commercially developed tools and those designed in-house by providers. They ranged in purpose from initial assessment tests in English and maths (designed to assign a functional level for learners) to diagnostic tests (which are designed to identify specific learning strengths and needs). The former relate to attainment against national standards and were used to confirm a learner's suitability for a programme of learning. They might show, for instance, that a learner had spelling skills at Entry Level 3, and is able to spell (or not spell) a selection of words at this level. The latter (the diagnostic tests) focus on identifying and shaping an individual's learning targets and the teaching and learning strategies that might be needed to achieve them. They might identify that a learner at Entry Level 3 was secure with words of type A and B, but had problems with words of type C (a learner's 'spiky profile'). This information would enable the provider to tailor a specific and personalised programme of learning for the individual.
- 4.3 These variations provide a challenge for the research team in equating the tools, since they had different aims and provided very different profiles of skills. Moreover, it became apparent that providers had mainly identified tools that would be used at the start of a learner's course (that is, the pre-test), relying on the achievement of a qualification or certificate to act as the measure of distance travelled (the post-test). Since no national testing has yet been done to equate any of these tools to external

¹⁴ New Zealand have instituted an online adaptive tool to measure Literacy and Numeracy for Adults– see <http://literacyandnumeracyforadults.com/Assessment-Tool-Support>

qualifications (even when they have been designed to reflect national standards), it became clear that, in the absence of any statistically equated post-test, distance travelled by learners in the pilot could only be measured by re-administering the pre-test, with all the concerns that raises about incorrect post-course assessments affected by learner familiarity with the items, for example.

- 4.4 The ways in which these challenges were addressed is explored more fully in Chapter 5. In order to underpin (and inform) the statistical analysis, however, a critical review of each of the tools was undertaken by Professor Greg Brooks. This review sought to go beyond the identification of whether the tools being used were initial assessment tools or diagnostic tools, and looked instead at a range of assessment criteria including validity, reliability, and manageability, as set out in Chapter 2.

Focus and content of the review

- 4.5 In considering these assessment criteria, there are further considerations, including establishing which elements of maths and English should, legitimately, be the focus of any assessment for distance-travelled. The change of nomenclature from ‘adult literacy and numeracy’ to ‘adult English and maths’ has the advantage of making the immediate labels more readily comprehensible. Nonetheless, it also raises questions as to what is currently assessed under these headings – and what should be assessed in the future.
- 4.6 It is recognised, for example, that maths is broader than numeracy. Initial assessments tended not to include those aspects such as algebra that do not form part of the core or Functional Skills curriculum. While many online tests provided onscreen calculators where appropriate, it was debateable as to whether some of the set problems had any real world context (or whether, indeed, they tested reading skills rather than calculation skills).¹⁵ In considering the adoption of a distance-travelled approach for funding for the future, it will be essential to decide which skills (and at what level and degree of complexity) should be included in any pre- and post-tests used for maths.
- 4.7 In addition, there is a further consideration as to the link between assessment and learning. Since teaching is primarily aimed at enabling learners to develop and apply their skills, there is a question as to whether those skills should not also be assessed in ways that reflect the various contexts in which the learners may feasibly need those skills (face validity).
- 4.8 This issue is particularly highlighted when considering the assessment of English, which is clearly broader than just reading skills (including listening comprehension,

¹⁵ Two of the tools avoided this by providing instructions in an audio file as well as on screen, although, in each case, the learner has the option of turning the audio off.

speaking and writing). Assessing such skills, however, is not straightforward and the role such assessments would play in a funding model based on skills gain needs to be carefully considered. In the case of English, for example, it has been argued that assessing listening comprehension in a test setting (through responding to aural stimuli using multiple response tests) can lead to the use of non-plausible or invalid assessments. Two of the tests that were reviewed (both commercially available) sought to assess listening comprehension; while the tests were set up well (with the both the contextual information and the questions provided only aurally) the tasks (with learners checking a multiple response item in response to the question) largely lacked face validity.

- 4.9 In terms of oracy (interactive speaking and listening), the prevailing view is that these are best assessed face-to-face and in context by teachers. None of the tools reviewed showed any attempt to assess speaking, for which there does not appear to be any current online assessment. It is, in any case, difficult to see how this could be implemented validly.
- 4.10 In order to make it easier to mark, the assessment of writing can focus too much on surface skills (spelling and punctuation, for example, including the correction of 'incorrect' English) and too little on authorial skills (including quality and appropriateness). This was evident in one of the three in-house tests that attempted to assess writing; the remaining two presented valid tasks, but did not make it clear how these tasks would be assessed. The reliability of marking for written tests is a particular challenge; generally this is overcome by double marking, although there was no indication that any such approach was advocated by the tool designers or adopted by the Pathfinder providers.

Outcomes of the review

- 4.11 A total of 36 assessment tools from a range of different sources (including the then DfES, commercial providers and in-house tools) were examined and reviewed in order to assess their suitability as tools to capture skills gain (distance travelled).
- 4.12 The majority of the tools reviewed were diagnostic tests (26 -the exceptions included two tests – one in English and one in maths – designed by a public-sector provider, and two tests – again in English and maths – designed by a commercial provider). Although the reviewer questioned the appropriateness of the use of diagnostic tests (which may be narrowly focused on specific skills rather than on all of the skills required within a Level) as a proper indicator of distance travelled, these tools were subject to the same level of review as the initial assessment tools.
- 4.13 The findings from the review are summarised below. It should be noted that the reviewer was not able to access five of the published tools in use in time for this report (these needed permissions from the publishers and were not granted in time for the review. It should also be noted that the tools being used by the Pathfinder

pilots were not always the most recent versions of tests (some were nearly a decade old); the commentary here relates primarily to the tools that were in use, although the reviewer referenced more recent versions where that was possible.

Table 4-1: Critical review of the tools in use by providers

Aspect	Comments
Validity	<p>Items in the 36 tests reviewed were considered mainly ‘valid’. Some good practice was evident in the use of graphics, or strategies to ensure that numeracy learners were being tested on their numeracy skills, not their literacy skills. Nonetheless:</p> <ul style="list-style-type: none"> • Some of the tests being used by the providers were out of date (although a number of these tests had more up-to-date versions that were not being used) • Two of the literacy tests included items that were designed to assess learners’ ability to correct wrong English, which is not generally accepted as good practice in assessing literacy • At least one numeracy test included as ‘wrong’ items (such as the presentation of dates of birth) that would have been correct in other countries • Some of items in the numeracy tests lacked ‘face validity’; testing instead other skills than the ones they set out to measure.
Reliability	As indicated in Chapter 2, this analysis could not be done for this review.
Manageability	It was judged that the majority of these tests would be too long to be used as an assessment tool for distance travelled. While eight were deemed acceptable or adequate in the anticipated length of completion time, most would take more than an hour to complete. Whether on line or on paper, most learners might find them too long to finish properly, while providers might find them more difficult to administer. Only four tools were deemed good.
Usefulness	<p>Seven of the tools being used by providers were very out of date, particularly those that were paper-based (though more recent online versions of some of these tools are available, they were not being used by the Pathfinder pilot providers).</p> <p>A number of the online tools had good features and were judged as potentially very useful with some modifications (such as the removal of items testing incorrect English)</p> <p>26 of the tools were labelled and intended as diagnostic tests, others were designed as level finders and only 4 as initial assessments.</p>
Statistically equated parallel forms	Only one of the tools in use is known to have statistically equated parallel forms. Three others tools (by virtue of being online, adaptive and with a long enough measurement scale) could potentially act as such. No others were identified.
Length of measurement	The majority of the tools that were being used by the providers measured only one level. One set of tools reported against five

Aspect	Comments
scale	levels and two against six, but were not sufficiently granular to show small steps of progress.
Precision of measurement scale	Around half of the tools had enough items to have produced more detailed assessments – and so, potentially, could report smaller steps of progress. However, this reporting relies on clear and justified rules for aggregating responses across items (not apparent for any of these tools).

Source: Brooks, 2013

Conclusions and suggestions

- 4.14 This review has shown that most of the instruments used in this project were rather poorly suited for measuring learners' skills gain, especially small amounts of distance travelled. Four instruments (all commercially produced) were judged more suitable than others for further development, but even these were not yet precise enough nor secure enough in the items they measured or the ways they measured them to be fully suitable.
- 4.15 If the field is to be rewarded for learners' distance travelled, the following recommendations, based on best evidence from previous comprehensive literature reviews in this field¹⁶, are worth consideration.
- There is a case for restricting online assessment to maths, reading and (provided plausible contexts can be found or devised) listening comprehension.
 - Writing should be assessed only in class by teachers. Clear guidance on this should be published, and the guidance should give more weight to the quality and appropriateness of learners' writing than to the surface features of spelling, punctuation, grammar and handwriting.
 - There is a strong case for eschewing tests of 'wrong English'.
 - There is a case for insisting that providers use only tools which:
 - are up-to-date
 - can assess across levels and measure small steps of progress
 - avoid practice effects, that is, either adaptive online assessments or paper-based tests with parallel forms (both should have evidence of statistical equating).

¹⁶ See, for example, Brooks, G (2005) Assessing adult literacy and numeracy: a review of assessment instruments. National Research and Development Centre for Adult Literacy and Numeracy London: Available: http://www.nrdc.org.uk/publications_details.asp?ID=23# The report was updated in an internal report for BIS in 2013.

- Providers would be justified in insisting that contextual factors be taken into account if they are to be rewarded for learners' 'value added' (skills gain/distance travelled).

4.16 On this basis, none of the tools currently in use by the providers appear to be entirely fit for measuring learners' progress – hence 'distance travelled'.

5. Quantitative analysis: initial findings

5.1 In the following subsections we provide:

- a discussion of the quality of the collated data, highlighting issues around local collation and the extent to which it could be seen as comprehensive, for example.
- an analysis of the collated baseline data, commenting (where possible) on concurrent validity and (where available) construct validity.

Data quality

5.2 Providers completed the spreadsheets on learners, courses and pre-test scores discussed in Chapter 2. These spreadsheets collated some data that would already have been collected for the ILR (but would not have been available through a central database in time to inform this project), along with some additional demographic fields (such as whether the learner spoke a first language other than English), course information and pre-test data.

5.3 The level of completion varied. While some data was received from providers on 4,599 learners, in many cases critical baseline assessment data was missing; pre-course assessment data, in a form that could be included in the baseline analysis (as raw scores or percentages, rather than simply as levels, such as Entry level 1, 2 or 3 or Level 1 or 2), was provided for only 57% of the known sample.¹⁷ Effectively, therefore, there were **2,599 learners** for whom there was at least some basic data from which to measure distance travelled. Post-test data, however, was made available for only a sub-set of these (**1,154** or 44% of the learners for whom pre-test data was available – see Table 5-1). As discussed below, however, some of the necessary contextual data was missing for a further seven per cent of these learners, reducing the sample to **1,077 learners** (41% of the pre-test population), of whom 456 were at Entry Level, 364 at Level 1 and 257 at Level 2 (see Table 3-3)

Challenges to data completeness

5.4 Many of the Pathfinder pilots faced challenges in the early stages of data collection and collation, not least of which was identifying the learners to include in the study. In some instances this was related to the type of courses learners accessed, with particular difficulties for those providers whose courses could be accessed on a

¹⁷ Data was also received on a group of 12 known ESOL learners; these were removed from the analysis.

flexible, non-term-linked basis. As a result, their initial estimates of learner numbers were often lower (as a result of a cautious view of recruitment) than the numbers for whom they provided data at Wave 1. In the case of multi-provider pilots, however, it was sometimes a struggle to recruit enough providers to the study to enable higher learner numbers.

- 5.5 Other challenges were related to the ability to obtain or access data for collation. While data was sometimes held centrally on management information systems, additional work was needed locally to obtain data on whether English was a learner’s first language, or to collate the data that was collected in different places. Some of the data that should have been routinely collected and collated for the ILR (including a learner’s employment status) was missing, reflecting the fact that these fields are often incomplete on the ILR itself. Pre-test data (whether initial assessments or diagnostic tests) were not routinely collated centrally in a number of providers, with some data recorded only in paper format, or held in different subject areas rather than centrally. In cases where assessments were conducted online, the pre-test data was often stored off-site and had to be retrieved before collation.

Table 5-1: Data availability on learners for whom pre- and post- assessment data was secured

	Data available on:	% of learners
Gender	1154	100
Age	1136	98
Course (English or Maths) currently being followed	1154	100
Ethnicity	1088	94
Prior qualifications in <i>any</i> subject	1034	90
Level of course being studied	1077	93
Disability/learning difficulties	989	86
Country of domicile	1029	89
Language spoken at home	907	79
Employment status (pre course)	1042	90
Employment status (post course)	905	78
Prior qualifications in English (all learners)	742	64
Prior qualifications in Maths (all learners)	778	67

Source: Pathfinder pilot data

Descriptive statistics

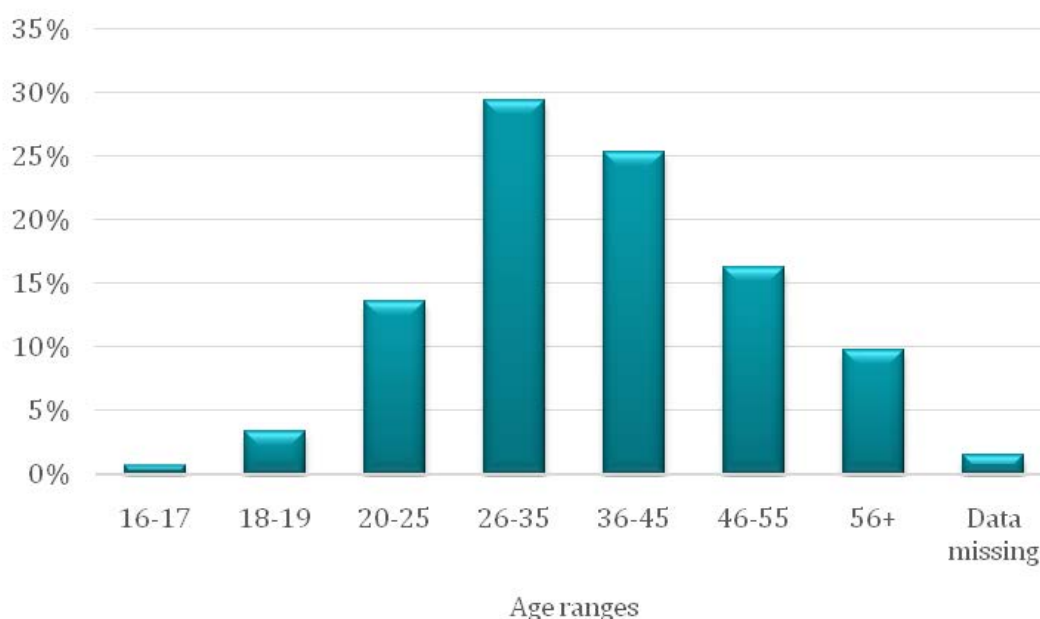
- 5.6 The tables and charts set out in this sub-section establish what is known about both the learners for whom we had more or less complete data (demographics, course

type and level, pre-test and post-test data and – in some cases – post-course qualifications).

Demographics

- 5.7 Female learners (66% of the 1154 learners) outnumbered males (34%), though the age distribution included learners from ages 18/19 to those aged over 56 (see Figure 5-1), with a mode of age 26-35. Some providers also included learners aged under 18 in their returns (less than one per cent of the total); these were removed from the analyses of distance travelled.

Figure 5-1: Age profile of learners

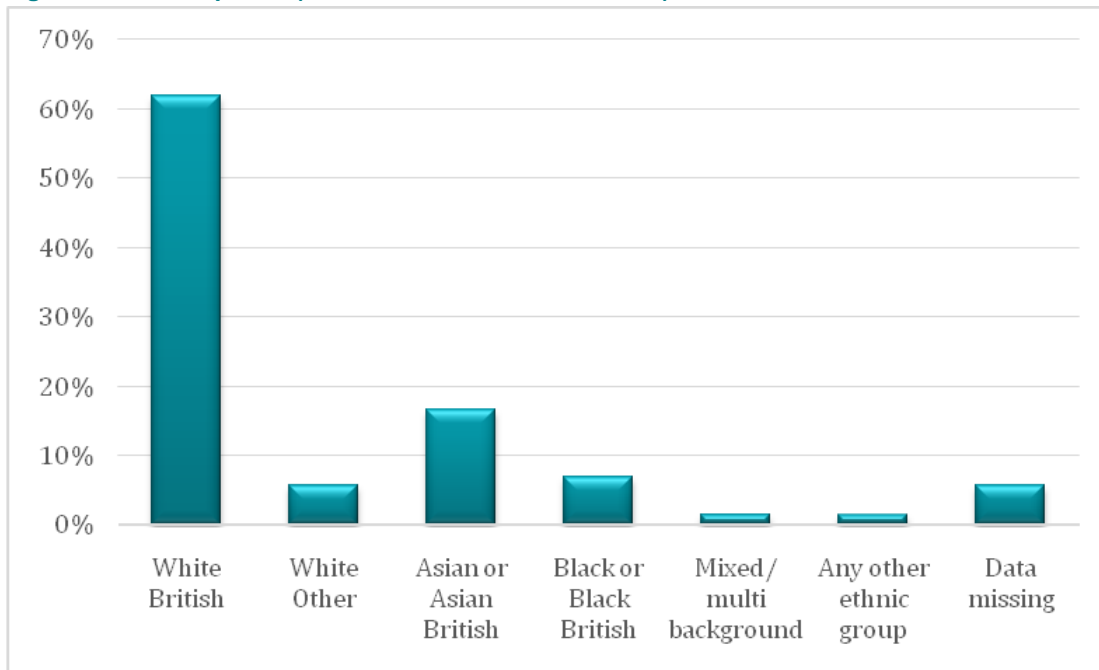


Source: Pathfinder pilot data

- 5.8 Over half of the study cohort (715) was made up of learners from White British backgrounds, although nearly one fifth of learners were from Asian backgrounds (17%) and just under one tenth from Black backgrounds (7%). The Census 2011 data indicates that just below eight per cent of England's 16+ population were from an Asian ethnic background and just over three per cent from a Black ethnic background, suggesting that the proportion of learners from Black and Asian backgrounds is higher than would be expected from a random population sample. It is close, however, to the profile of learners included in a recent study of adult literacy and numeracy learners (on-going unpublished study on skills gain for BIS) and, given the geographic locations of the providers, may be representative of the local population of such learners. Most (88% of those for whom language data was available) spoke English as a sole language at home. It should be noted, however, that this amounts to only 69% of the 1154 in the study; language data was missing for over one fifth of the learners.

5.9 Of those 989 (85% of the cohort) for whom disability status was provided, 284 had a known disability or learning difficulty. For some, this might be a key factor in determining their course and learning programme. Only 12 learners (1%) were noted as being in full time education, though the pre-course employment status was not recorded for over 10% of the learners. Amongst those for whom it was noted (1,042), most (55%) were in a job, 34% were unemployed and the remainder (10%) were identified as economically inactive; over one third of the economically inactive group was aged over 56.

Figure 5-2: Ethnic profile (based on data for 1154 learners)



Source: Pathfinder pilot data

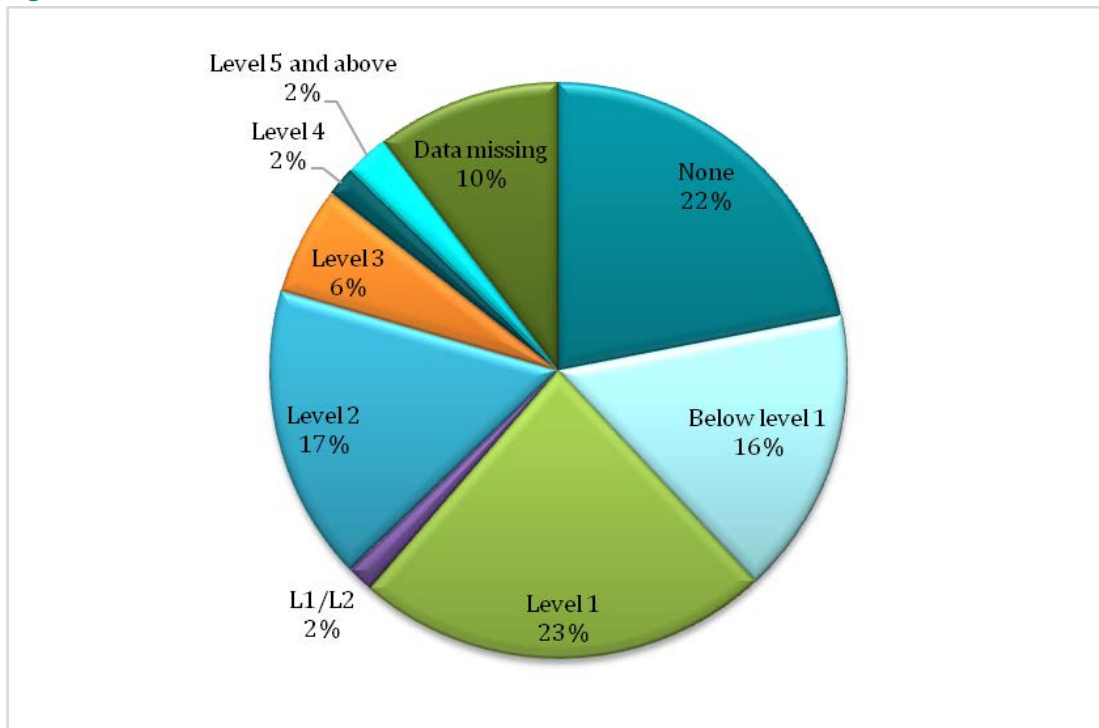
Qualification level of cohorts

5.10 Many of the learners in the study lacked either a prior qualification (at any level) in maths (33%) or a formal qualification in English (36%), though only 10% lacked *any* prior qualification in maths, English or any other subject. At the other end of the attainment spectrum, however, were those who had already achieved a qualification above Level 3 in at least one subject(see Figure 5-3); two per cent were reported to be qualified at Level 5 or above, though none were qualified above Level 2 in maths or above Level 3 in English.

5.11 This suggests that the cohort studying adult literacy and numeracy in the Pathfinders will be doing so for a variety of different reasons. In addition to those who lacked any formal qualifications, there will be those whose English fluency may be low (not all of

whom will be designated as ESOL learners)¹⁸ or learners who, while being successful in other academic subjects, nonetheless failed previously to attain well in maths, for example (none of the cohort was qualified above Level 2). In monitoring distance travelled therefore, it is likely that some may make rapid progress between and across the levels (as English fluency or confidence in math's skills improve), while others (including those with no or low prior qualifications) may make limited progress, requiring a great deal of granularity in the tools used to assess their skills gain.

Figure 5-3: Qualification level of cohort



Source: Pathfinder pilot data

Relationship between assessments and courses?

- 5.12 Although we had pre- and post-test assessment data for all 1,154 learners, information on the level of the course to which they were working was missing for seven per cent of them; the following analysis is based on the 1,077 learners for whom full data on both the courses being followed and the level to which they were working (missing for 7% of learners.) was available (see Figure 5 4). A high proportion of those studying maths or English had no prior qualification in those areas; nearly half (47%) of the maths learners had no maths qualification, while over two fifths of the English learners (42%) were similarly unqualified. A higher proportion of English learners (47%) than maths learners (38%) were working

¹⁸ Indeed, those who spoke a first language other than English were more likely to have a higher prior qualification in at least some subject(s) (at Level 3 and above), even though they had low prior attainment in maths and English.

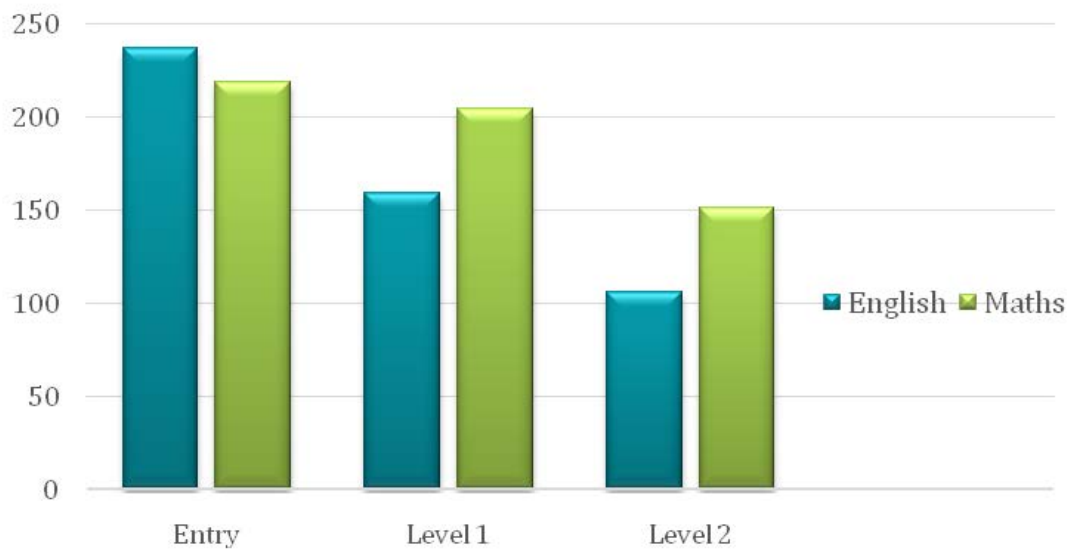
towards an Entry level qualification. Equally a higher proportion of maths learners (26%) than English learners (21%) were studying for a Level 2 qualification.

5.13 When looking more closely at those who are enrolled at different levels, some further differences are highlighted. Most telling is the difference, for some learners, between the qualification level previously achieved and the level of course on which they were placed following initial assessments. The data suggests that:

- Of those 402 learners for whom prior formal qualifications in maths had been provided *and* who were also studying maths:
 - 44% were following Entry Level courses (E1, E2 or E3). Of these, 13% had previously achieved a Level 1 or Level 2 qualification in maths.
 - 21% of those who had no prior formal maths qualification were placed on a Level 2 course and over half (51%) followed courses at Level 1
 - 15.1% of those with a previous Level 2 maths qualification were assessed as performing at Level 1 or lower and placed on a course at that level.

- The story was similar for English:
 - 56% of the 378 learners for whom course and prior qualification data were available were following Entry Level courses, although 10% of them had previously attained English qualifications at Level 1 or above.
 - 16.5% of those who had no prior formal English qualification were placed on a Level 2 course and 28% followed courses at Level 1
 - 46% of those with a previous Level 2 English qualification were assessed as performing at Level 1 (33%) or lower (12.6%) and placed on courses at those levels.

Figure 5-4: Enrolments by subject and qualification level (based on 1077 learners)



Source: Pathfinder pilot data

5.14 These findings raises three significant sets of questions:

- To what extent are learners' prior qualifications (which may have been gained in different educational systems, or at a different time point) a good indicator of current skill levels? Does the difference between previous qualification levels and current assessment levels suggest that, in the intervening period, there has been skills atrophy for some, but skills gain for others?
- How reliable and credible are the assessment tools being used to assess current skills (particularly given the outcome of the review summarised in Chapter 3)? Do they under- or over-estimate learners' skills? Or do the initial assessment tools/diagnostic tools focus on different skills from those assessed in formal qualifications?
- How much credence or weighting is given to these two different measurements as an indicator of the course type and level to which to assign a learner?

Analysis of validity and distance travelled

5.15 As indicated in Chapter 2, the study sought to measure the distance travelled of the learners in the study and to provide information on the ability of the tools selected by the Pathfinders to measure that distance travelled accurately and reliably. In addition to the critical review of the tools (presented in Chapter 3), the statistical analysis looked at:

- the relationship of learners' scores in the pre-test tools and their previous attainment (the convergent and discriminant validity of the tools)
- the extent of learner improvement between pre-test and post-test scores (the measurement of distance travelled)
- the measurement of pre-test and post-test scores and distance travelled for learners with different skill levels (concurrent validity)
- the relationship of learners' scores in the post-test tools and their subsequent attainment (the predictive validity of the tools)

Discriminant validity

- 5.16 The assessment of discriminant validity is important in checking the extent to which the tools that were used in the initial assessment of learners were effective in discriminating between tests of different skills. In other words, we would not expect to see a strong linear relationship between an assessment of numeracy and a learner's prior qualifications in English, nor of an assessment of literacy and a learner's prior qualifications in maths. In order for the analyses to take place, data on each learner's prior qualifications (in English and/or maths) and initial assessment scores (in literacy or numeracy) needed to be available.
- 5.17 Data on prior qualifications, in general, was available for the majority (82%) of the 2,599 learners for whom pre-test data was available. However, the proportion of learners for whom prior qualifications in maths and initial assessment data in literacy (497) and for whom prior qualifications in English and initial assessment data in numeracy (611) was available was smaller, and it is on the basis of these **1,108** learners that the analysis was conducted. That demonstrated that there was no linear relationship between, for instance, learners' prior attainment in English and the scores they attained in initial assessments of numeracy, suggesting that the tools used by the pilot providers **had strong discriminant validity** and, to that extent, were fit for purpose.
- 5.18 Good discriminant validity, however, simply ensures that a test that purports to assess numeracy skills tests that skillset and not some other skillset (such as literacy). It does not indicate that the tool measures those skills accurately – or in a way that indicates that the course a learner might follow is appropriate for their skill levels. To do that we need to look at concurrent validity.

Concurrent validity

- 5.19 In assessing concurrent validity, we wanted to examine any apparent relationship between learners' prior attainment and their initial assessment scores, and also between the initial assessment scores and the level of courses onto which learners had been placed following those assessments. Since assessment and placement

practices amongst the providers varied, as did the types of tools they used, we also explored the relationship between learners' prior qualifications and the distance-travelled,

5.20 The discussion in paragraphs 5.12 to 5.14 outlined some of the anomalies between prior qualifications, initial assessments and course placements (Table 5-2 and

5.21 Table 5-3), and further analysis suggests that the concurrent validity of prior qualifications and initial assessments was generally weak. A detailed analysis was hampered not only by the extent of missing data on prior qualifications and initial assessments (this was available for only 780 learners), but also by the fact that some learners were on courses for which the level was not fixed (mixed-level courses) and by the spiky profile of many learners. This last point meant that it was not possible for us always to be certain as to where the variations lay between qualifications, initial assessment and course allocation. Qualification data might be unreliable (with learners providing incorrect prior attainment information), the prior qualifications may have measured different skills from the initial assessments, or the initial assessments may have had a low level of convergent validity.

Table 5-2: Maths/numeracy: Qualification level (working towards) by baseline assessment results

Level of the course		Highest qualification in Maths				Total
		none	Entry	L1	L2	
Entry	Count	93	60	18	5	176
	% within Level/Qual	52.8%	34.1%	10.2%	2.8%	100.0%
L1	Count	63	10	34	15	122
	% within Level/Qual	51.6%	8.2%	27.9%	12.3%	100.0%
L2	Count	22	2	55	25	104
	% within Level/Qual	21.2%	1.9%	52.9%	24.0%	100.0%

Source: Pathfinder pilot data

Table 5-3: English/literacy: Qualification level (working towards) by baseline assessment results

Level of the course		Highest qualification in English				Total
		none	Entry	L1	L2	
Entry	Count	87	67	10	8	172
	% within Level/Qual	50.3%	38.6%	5.6%	4.3%	100.0%

Level of the course		Highest qualification in English				
		Count				
L1	Count	44	10	31	21	106
	% within Level/Qual	41.5%	9.4%	29.2%	19.8%	100.0%
L2	Count	26	0	40	34	100
	% within Level/Qual	26.0%	0.0%	40.0%	34.0%	100.0%

Source: Pathfinder pilot data

5.22 From the data to hand, it appears that there was a clearer linear relationship between initial assessments in numeracy and prior qualifications in maths (379 learners) than there were in English (430 learners), where the indications are that concurrent validity was weaker.

5.23 This assessment of the extent of concurrent validity remained when we explored the relationship between initial assessment data and distance travelled, where the linear relationship between initial assessments in numeracy and distance-travelled was clearer for maths/numeracy learners than for English/literacy learners (see paragraph 5.30).

Distance travelled

5.24 Each of the tests used by the providers led to the production of a learner level and/or raw test scores. Since learner levels are too broad to identify small changes in skills gain (see Chapter 1), raw score data and information on the number of possible points at each level (and for each test) were essential in order to calculate the percentage of correctly answered questions and, once the post-test scores were available, assess distance travelled. Data at this level of disaggregation is available for **1,149** learners.

5.25 Ideally, in order to examine the distance travelled, we would have looked at the distance between individual learners' scores on a test and the arithmetic mean for that test (for all learners at the same level who were assessed using the same tool). This would have involved the identification of raw scores, their conversion into percentages and the calculation of the mean and standard deviation (the average distance of scores away from the mean) for each tool. Assuming that the scores fall into a normal distribution, 95% of the students would have scores that were within two standard deviations above or below the mean. Knowing the mean and standard deviation for each test would make it possible to interpret raw scores and to compare different individuals' performances with each other, or an individual's performance on one test with his or her performance on another test (or over time).

5.26 The data we obtained from providers was not always amenable to such detailed analysis, however, primarily because the tools themselves were not standardised. In

many cases, the data from the initial assessments and the post-tests simply provided raw scores and an indication of the maximum score possible, but not the level of assessment used. For many learners, the only indication of assessment level was the course upon which they were subsequently enrolled. Using the obtained score and the maximum score possible, we calculated percentages to make comparisons (within tools) possible, but the data itself could not be standardised sufficiently to allow a robust and meaningful analysis. Variability in the tools themselves means that the analysis we conducted should be seen as *indicative* of distance travelled and not a robust comparative analysis across all tools.

- 5.27 The distance travelled was calculated simply as a difference between percentages in pre and post assessments (including the calculation of standard deviations). This simple approach enabled us to minimise the effect of the different baseline scores (and their range) of the different tools used for the assessments. Nonetheless, the tools were still not directly comparable and the findings should be seen as illustrative only.
- 5.28 That analysis found that, *on average*, there appeared to be a significant positive improvement in learners' scores that could be conceptualised (potentially) as distance travelled. In the maths (or numeracy) scores, for example, there was a significant mean improvement (across 647 learners) of 15 percentage points between the initial assessment and the post-test, and a smaller (but nonetheless significant) improvement in percentage points in English (or literacy) test scores for the 502 English learners (Table 5-4).

Table 5-4: Distance travelled. Statistical test of difference between pre and post assessments

Subject of qualification		Mean %	N	Std. Deviation	P-value from paired t-test
Maths	Pre-test	63	647	20.4	0.000
	Post-test	78	647	16.9	
English	Pre-test	60	502	19.5	0.000
	Post-test	70	502	18.7	

Source: Pathfinder data

- 5.29 This analysis, however, overlooks the variation, by learner, by subject and by tool, which was identified during the study. In exploring the data further, we looked at the profile of individual learners and found that not only did some learners appear to make no progress, others had *lower* post-test than pre-test scores. This accords with the views expressed by some providers, some of whom felt that learners were de-motivated by the use of a post-test (particularly where it was not related to the assessment process for their qualification). None of the tools in use had true parallel forms and it is possible that, in the case of some of the adaptive tools, the post-test tested different skills from those assessed in the initial assessment.

5.30 Indeed, despite the apparent average point score gain (Table 5-4), there was in fact a strong negative correlation (-0.6 for maths and -0.4 for English) between pre-assessment scores and distance travelled for individual learners (Table 5-5). While the mean point score gain for maths learners was 15.11 percentage points, the standard deviation was 16.8; the point score gain for English was 10.2 percentage points, but the standard deviation was even higher, at 14.4.

Table 5-5: Distance travelled. Correlation between pre-assessment score and distance travelled

Subject	Mean	Std. Deviation	Correlation
Maths: Change in point score	15.11	16.8	-0.6
English: Change in point score	10.2	14.4	-0.4

Source: Pathfinder data

5.31 An analysis of initial assessment data and distance travelled suggested that the amount of *measured* progress made by Entry Level learners was generally greater than that for higher level learners. Entry Level maths learners, for instance, scored a mean additional 21.5 percentage points (though with a standard deviation of 19.6), compared to those initially assessed at Level 2, where progress was more in the order of 10.7 percentage points (with a standard deviation of 13.6). The differences were lower in English (where concurrent validity was weaker – see paragraph 5.20), but followed the same pattern. This does not necessarily mean that Entry Level learners made more progress than higher level learners (who may, of course, have had less distance to travel), but that the tools used picked up at least some changes at this level.

5.32 Nonetheless, given the conclusion drawn from the review of tools (that most of the instruments used in this project were poorly suited for measuring learners' skills gain, especially small amounts of distance travelled) and the statistical analyses conducted with the data obtained from providers, we found no evidence that the assessments that were used during the pilot would be either valid or reliable as a measure of skills gain or distance-travelled at any level.

Predictive validity

5.33 We explored predictive validity in terms of the relationship between:

- initial assessment and likelihood of course completion
- initial assessment and likelihood of achievement of a qualification (note that numbers were too small to examine the predictive validity by qualification type, such as Functional Skills or GCSEs, for example)

- initial assessment, distance travelled and likelihood of achievement of a qualification

5.34 In each case, the analyses took account of learners' socio-economic characteristics, and the subject focus and level of their course.

- In the first case, course completion appeared more likely amongst those learners (whatever their background) whose mean score (at whatever level) was higher in their pre-course assessment (62 %+).
- In the second case, the predictive validity of the initial-assessment tools were less apparent; there was no significant difference between those who achieved their qualification and those who did not in relation to their scores at initial assessment.
- However, those learners who achieved their qualification had a statistically significantly higher distance travelled score compared to those who did not complete their course. While those who did not complete their course showed a mean of seven percentage points distance travelled for maths (compared to a sample mean of 15.11), and six for English (compared to a sample mean of 10.2), those who achieved their qualification recorded average distance-travelled scores of 19 percentage points for maths and 16 percentage points for English.
- Incorporating the socioeconomic background of learners and their initial assessment results into the analysis (exploring the predictive validity of distance-travelled on qualification achievement) meant that the statistical association with qualification achievement was no longer evident.

5.35 In effect, the analysis indicates that the initial assessment tools used in the pilot had some predictive validity in relation to course completion, but no evident predictive validity in relation either to course completion or to qualification levels achieved.

5.36 We also explored the predictive validity of the tools in relation to **post-course progression**, though that was hampered by limited data (which was available for only **455** learners across maths and English) and the fact that there was no clear hierarchy of outcomes in terms of progression. Those who experienced the greatest distance-travelled from their initial assessment appeared more likely to progress to further learning than to employment, but those who gained employment (whatever their socio-economic background or course characteristics) generally scored higher on pre-course assessments.

In conclusion

5.37 In our analyses, we sought to establish whether or not the tools used in the pilot could be used to measure distance travelled or whether the scores they produced

could be used to predict likely course completion, attainment or progression outcomes. While the range of tools used by the pilot providers appeared to have good discriminant validity (in terms of measuring the type skills that they were set up to assess), the data provided only limited evidence for their weak concurrent validity and even weaker predictive validity. Their ability to measure distance travelled was very limited and, while it was possible to see distance travelled in the aggregated figures, the standard deviations were high. We conclude that while the tools used in the pilot may be indicative of progress at an individual level, they should not be used to monitor or assess skills gain between courses or providers as the basis of a funding model (or any other comparative assessment between or within providers).

6. Conclusions and implications

6.1 It should be noted that, prior to the Pilot, none of the participating sites had specifically administered a post-test, relying in general on external qualifications to measure learner progression. Moreover, as some providers noted, the primary outcome many were seeking to achieve for their learners was entry into employment, rather than measurable progression in maths or English *per se*. The study has, therefore, highlighted a number of issues that have implications for the longer-term use of a distance-travelled approach to funding.

- Providers' experiences, the critical review of the assessment tools they deployed and the statistical assessment suggest that none of the tools used in the pilot is suitable for use as the metric by which distance travelled should be assessed. While four of the 36 tools that were reviewed had statistically equated parallel forms (or by virtue of being online and adaptive, could be used in this way), most had not and so could not be used dependably to measure distance travelled.
- The statistical analysis indicated that none of the tools had good concurrent or predictive validity and so could not be used in any reliable way as a pre-test alongside learners' subsequent formal qualifications for measuring distance-travelled on a larger scale.
- Online and adaptive tools would, arguably, be the most cost-effective way of conducting pre-test/post-test assessments, but raise significant logistical (and other) challenges in cases where access to ICT facilities is limited. In these locations paper-based tools would still be needed; none of the existing paper-based tools used by the providers met the needs of a distance-travelled funding approach.
- The level of preparedness for data collection, recording and collating varied widely across the pilot sites, all of whom were volunteers and eager to make a contribution to this process. This highlighted a potential concern for the wider roll-out of any such approach. Differences in provider practice and recording and sharing of data means that guidance and support will need to be given to adult learning providers. It should be noted that not all of the detailed data that was needed for the pilot will be needed for any national roll-out.
- If distance-travelled is to be used as the basis for a future funding mechanism, pre and post assessment results will need to be recorded on the ILR as compulsory fields alongside data on course completion and qualification achievement. Moreover, additional data fields (such as language used in the

home) may be needed on the ILR, while existing data fields (including employment status) may need to be more rigorously completed.

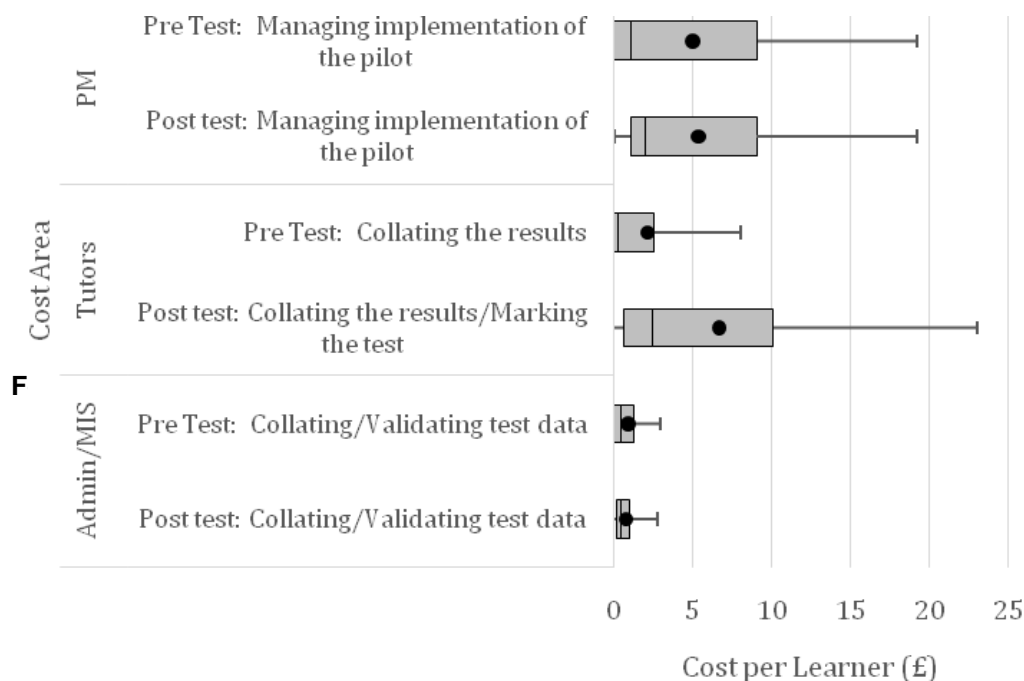
Annex A: Calculating the financial costs of rolling out a distance travelled approach

- A.1 The following section details the method used to calculate the financial costs of rolling out a distance travelled approach using the data collected during round two of the fieldwork. Where appropriate, justification is given for the approach taken and the potential implications of this for the analysis.
- A.2 Through round one of the fieldwork providers were asked to identify any additional activities that they had undertaken as a result of the pilot. While there was a degree of variation in the precise activities undertaken, and the type of staff required to action these. A number of cost lines were identifiable. These were: marking test scripts, collating/validating test data, and managing the implementation of the pilot.
- A.3 Through round two, providers were asked to estimate how much additional time they had spent on these activities and what this equated to in financial terms. While all but one of the providers proved able to provide an assessment of the additional time commitment made by staff as a result of participation in the pilot, sites had much greater difficulty commenting on the financial costs of participation. As such the following analysis is based on the time commitment of staff multiplied by an hourly rate, based on an approximation of the salary levels of different types of staff¹⁹. These were estimated as follows:
- Admin/Management Information System staff - £10.26 per hour
 - Tutors – £15.38 per hour
 - Project Managers - £17.95 per hour.
- A.4 Providers were also asked to confirm how many learners they expected would complete a post-test. This was found to differ markedly between the sites (from 14-437). As a result, all costs are calculated on the basis of cost per learner. Where possible, providers were asked to reflect on whether this figure would reduce if the number of learners increased, but very few were able to do so.
- A.5 Table A-1 shows each of the cost lines, indicating the arithmetic mean (shown as a dot), the median (or mid-point value of the data) and the quartile values (showing the spread of the data). The dataset included a number of outliers; indeed, all but one of the cost lines contains values more than one and a half times larger than the value of the third quartile. This is particularly pronounced in the case of data pertaining to the collection of pre-test data. Here the maximum value (£8.00 and £12.15 per learner respectively) was over 50% greater than that of the third quartile value (£2.50 and

¹⁹ Salary levels were estimated on the basis of information submitted by the providers.

£3.75 per learner respectively) for sites using both paper and web-based/online tools. There is also evidence that the majority of the cost lines are not normally distributed²⁰, and, in most cases, are skewed towards zero. For seven of the twelve cost lines, the (arithmetic) mean value is 50% greater than the median. This tendency is clearly evident for the cost of managing the implementation of the pre-test in those sites using a web-based tool where the mean value (£1.87 per learner) is 100% greater than the median (£0.00). Given these tendencies, it would be inappropriate to use the (arithmetic) mean as an average and, instead, we used the median value (the mid-point of the data) when reporting findings in the text.

**Table A-1: Distribution of cost per learner data by cost area
For sites using paper tools**



²⁰ By normally distributed it is meant that the population is symmetrically distributed around the arithmetic mean.

- A.6 Table A-2 presents the estimated cost of involvement in the Pathfinder by cost centre. The analysis is also differentiated in terms of the type of tool used by the provider be that a paper or web-based/online tool.
- A.7 The median cost of implementing the pilot for all sites (regardless of the tool used) was £4.72 per learner. The median cost of implementing a distance travelled approach was 11% greater for those sites using paper based opposed to web-based/online tools (£6.62 opposed to £5.92 per learner).
- A.8 Were the approach to be adopted, sites felt that they would be unlikely to require the services of a project manager. Discounting time spent on project management time is likely to give us a more realistic understanding of the overall cost of delivering a distance travelled approach. Once project management time is disregarded, however, a much more complex picture emerges. In this case, the median cost of implementing the pilot for all sites (regardless of the type of tool used) was £2.64 per learner. The median cost for sites using paper based tools was £3.56 per learner, but the median cost for sites using web-based/online tools was greater at £5.56.
- A.9 Consideration of the relative costs of undertaking pilot related activities gives a stronger indication of why this might be the case. As might be expected, given that prior to the pilot none of the sites had administered a post-test, administration of the post-test was found to be more costly than the pre-test (a median cost of £4.54 opposed to £0.18 per learner, for all sites). Collating the results and marking the post-test was the most expensive cost group (a median cost of £2.40 per learner out of a total median cost of £2.64 per learner), accounting for 91% of the total cost of administering a distance travelled approach (discounting PM time) regardless of the tool used. Taking account of the type of tool used, the costs differed markedly. While the median cost of collating the results and marking the post-test was £2.40 per learner for sites using a paper-based test, for those sites using web-based/online tools was £3.85 per learner (just under one third more -31%). When considering the relative cost of using paper-based opposed to web-bases/online tools, it will be important to take the associated costs (such as marking and uploading data) into account.

Table A-2: Estimated cost of involvement in the Pathfinder by cost centre

Cost Centre	Test	Additional Task	All Sites ²¹						For sites using paper tools						For sites using web-based/online tools					
			Cost per Learner (£)						Cost per Learner (£) ²²						Cost per Learner (£)					
			Median	Mean	Min	Max	Q1	Q3	Median	Mean	Min	Max	Q1	Q3	Median	Mean	Min	Max	Q1	Q3
Admin/MIS Staff	Pre Test	Collating/ Validating test data	0.10	1.37	0.00	7.69	0.00	2.36	0.41	0.88	0.00	2.97	0.00	1.28	0.21	1.74	0.00	7.69	0.00	2.56
	Post Test	Collating /Validating test data	0.46	1.37	0.00	7.69	0.21	2.15	0.46	0.76	0.00	2.77	0.15	1.00	0.82	2.10	0.00	7.69	0.21	2.56
For tutors/trainers	Pre Test	Collating the results	0.08	2.02	0.00	12.15	0.00	3.85	0.23	2.12	0.00	8.00	0.00	2.50	0.15	3.03	0.00	12.15	0.00	3.85
	Post Test	Collating the results/ Marking the test	2.46	5.66	0.00	23.08	0.15	8.00	2.46	6.67	0.00	23.08	0.62	10.12	3.85	5.03	0.00	16.92	0.00	8.00
For PM	Pre Test	Managing implementation of the pilot	0.00	3.23	0.00	19.21	0.00	5.56	1.08	4.98	0.00	19.21	0.00	9.11	0.00	1.87	0.00	6.28	0.00	5.03

²¹Data is based on 16 sites, one site felt unable to quantify the costs of participation.

²² One site had used a mixed approach. As such data from this site was used in the analysis of both paper and web-based sites.

Cost Centre	Test	Additional Task	All Sites ²¹ Cost per Learner (£)						For sites using paper tools Cost per Learner (£) ²²						For sites using web-based/online tools Cost per Learner (£)					
			Median	Mean	Min	Max	Q1	Q3	Median	Mean	Min	Max	Q1	Q3	Median	Mean	Min	Max	Q1	Q3
	Post Test	Managing implementation of the pilot	1.62	4.32	0.00	19.21	0.00	6.28	1.97	5.36	0.00	19.21	1.08	9.06	0.90	3.47	0.00	13.46	0.00	5.56
Total	Pre Test		0.18	6.62					1.72	0.88					0.36	1.74				
	Post Test		4.54	11.36					4.90	0.76					5.56	2.10				
	Combined		4.72	17.98					6.62	2.12					5.92	3.03				
Total (Discounting PM Time)	Pre Test		0.18	3.39					0.64	6.67					0.36	5.03				
	Post Test		2.46	5.66					2.92	4.98					4.67	1.87				
	Combined		2.64	9.06					3.56	5.36					5.03	3.47				

© Crown copyright 2014

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. Visit www.nationalarchives.gov.uk/doc/open-government-licence, write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

This publication available from www.gov.uk/bis

Any enquiries regarding this publication should be sent to:

Department for Business, Innovation and Skills
1 Victoria Street
London SW1H 0ET
Tel: 020 7215 5000

If you require this publication in an alternative format, email enquiries@bis.gsi.gov.uk, or call 020 7215 5000.

BIS/14/1286