

Information Authority Data Service Start- up Projects

Harmonising OFSTED and LSC Qualification Success Reporting

Final Report

May 2008

Contents

Introduction	1
Background and terms of reference	1
Methodology.....	1
Summary.....	3
The methodology	3
Issues causing the differences.....	3
Other concerns raised by providers.....	3
The way forward.....	4
Implementation.....	5
Success rate reports.....	6
Overview	6
Basic methodology.....	6
FE.....	6
WBL.....	7
Current gaps in success rate reporting.....	7
Preparation of the data used to collect success rates.....	7
The underlying data used to calculate success rates.....	7
Timetable for producing the underlying data and success rate reports	8
Level of aggregation of LSC and OfSTED success rates	9
Issues causing real and apparent differences in success rates.....	9
Overview	9
Data differences	9
Differences in grouping the qualifications	10
Other concerns raised by providers.....	11
The way forward	12
Consistency.....	13
Quality	14
Transparency	14
Timeliness	14
Methods of aggregation	16
Governance.....	17
Implementation	19
Appendix 1 detail quotations from providers	20

Introduction

Background and terms of reference

- 1 This report was commissioned by the Information Authority (IA). The need for the report arose as the qualification success rates prepared by the LSC and OfSTED had apparent differences that have been generating burden on learning providers. The report is intended to put the IA in a strong position from which it can significantly reduce the burdens on learning providers caused by these different qualification success rate calculations and reports by delivering agreed recommendations on how they can be harmonised and how the changes can be implemented.
- 2 This report is the main deliverable from the assignment and documents the current situation and an approach to harmonisation that has been agreed with key stakeholders in OfSTED and the LSC.
- 3 The report covers:
 - (1) A clear overview of the existing algorithms, processes and reports that are produced and a summary of the problem. This splits the problem into its component parts – frequency, timing, calculation algorithms, inclusion/exclusion of certain qualifications, unclear reporting and documentation, etc. The findings set out in the report include the views of a sample of learning providers including a number who have recently been inspected as well as the opinions of the central agencies the LSC and OfSTED.
 - (2) A recommendation for harmonisation which has been discussed and prepared jointly with the LSC and OfSTED teams involved in the preparation of success rate reporting
 - (3) An executive summary
 - (4) Implementation plans for LSC, OfSTED and other user stakeholders.

Methodology

- 4 The approach broadly took the following main steps:

Project initiation

Initial meetings

Review of existing materials

Meetings with OFSTED and LSC

Initial briefing and discussions

Review of actual calculations in the LSC

Discussions with OFSTED

Group workshop/brainstorm

Produce interim paper

Focus group with learning providers

Initial test with 4 providers

Focus group

Reporting

- 5 The work was completed in close cooperation with key stakeholders in the LSC and OfSTED. The teams involved in both organisations have spent a considerable amount of time reviewing the findings in this report. We should like to take this opportunity to thank all of those involved who were:
- (1) Information Authority
 - a) Pete Ashton
 - b) Una Bennett
 - (2) LSC
 - a) Ellie Frazier
 - b) Ruth Mahon
 - c) Jonathan Price-Marlow
 - d) Tim Smith
 - e) Stuart Vickers
 - (3) OfSTED
 - a) Mike Davies
 - b) Nicola Edwards
 - (4) Providers/users of the data and provider MIS suppliers
 - a) Jon Collis – IA board and provider representative
 - b) Jo Eales – Software Solutions
 - c) Avril Greave – GetaHead Training
 - d) Paul Phillips – Weston College
 - e) Julie Mills – Milton Keynes College
- 6 In addition we also made use of several existing papers and guidance notes prepared by the LSC and OfSTED including the earlier Information Authority Paper - 03a: Success Rates - Data Collections and Algorithms.
- 7 Following a short summary, this report sets out the key findings and recommendations arising as a result of the review.

Summary

The methodology

- 8 Contrary to the initial expectations as set out in your original invitation to tender we found that the LSC and OfSTED use the same basic methodology to calculate success rates. We have described the methodology in overview in the section headed success rate reports (paragraph 18).

Issues causing the differences

- 9 The differences historically arise from (see paragraph 35):
- (1) OfSTED and the LSC using different underlying datasets to calculate success rates. These changes arise through:
 - a) The LSC altering the datasets after passing them to OfSTED to correct errors.
 - b) Reclassification of the qualifications so that they are grouped in a different way between the time the LSC publish their first reports and the time that OfSTED issue their analysis
 - c) Mergers of the providers necessitating the historic data to be merged to provide trend information for the new organisations.
 - (2) For Work Based Learning – OfSTED grouping qualifications in different ways to the LSC.

Other concerns raised by providers

- 10 In addition we have identified a number of concerns from providers that also require addressing (see paragraph 43). These concerns were identified from a relatively small sample of 5 providers. Whilst the sample is too small to make firm recommendations in all cases, as they were so consistent, the LSC/OfSTED should take their views seriously. However, we recommend that further more in depth consultation with providers is undertaken before acting on these views fully. The concerns raised were:
- (1) Whilst the benchmarking data prepared by the LSC was considered useful, the providers felt that the facilities for the analysis of their individual performance reports on the national LSC website are too limited and do not provide enough facilities for them to check the detail of their reports adequately.
 - (2) The timeliness of the data – success rate reports for the academic year 06/07 are not fully available until August 08 in the following year – the providers we spoke to felt that this generated an ‘uneven playing field’ for OfSTED assessments and is unfair as some providers are compared against expected national averages whilst others, inspected later in the year, are compared with the actual levels. (The actual position is not so clear cut as OfSTED use provider generated data as well as the published reports to inform their inspections.)
 - (3) The weighting methodology for some of the reports used to average and group qualifications success rates tends to favour providers who deliver short, easy to pass programmes. Again the actual position is more complex as the LSC and OfSTED use different averaging methodologies in different reports.
 - (4) Differences in the way OfSTED and LSC treat success rates with OfSTED applying statistical levels of confidence in the figures whilst LSC use the absolute value

means that a provider can have acceptable performance for OfSTED whilst still fail the Minimum Level of Performance (MLP) test for the LSC.

- (5) Technical issues generated by:
- a) Discrepancies in the guidance published by the LSC and the actual formulae used to calculate success rates
 - b) Limited information on the versions of the various datasets being used so that providers could replicate the qualifications with their own software.

The way forward

- 11 The LSC and OfSTED teams responsible for preparing the current reports have already identified a number of immediate improvements to ensure consistency. These are being implemented and should result in fewer, if any, discrepancies in the published reports for this year.
- 12 However, consistency in reporting is only one of the challenges identified within this study. In discussion with OfSTED and some staff within the LSC on earlier drafts of this report, the issue of overall governance of the reporting and quality assurance mechanisms was raised.
- 13 There are already many stakeholders involved and, with the Machinery of Government changes, there will be more including: the strengthened role of local authorities; the need for providers covering pre and post 18 age ranges to report on DCSF and DIUS programmes; and other key stakeholders at local, regional and national levels. With an increase in the range of stakeholders requiring information and performance data on their own policy initiatives, there is a strong likelihood that there could be a proliferation of reporting requirements and even more ways that the same data on qualification success rates is aggregated, averaged and reported. Even now within the LSC alone there is growth in the way qualification success rates are aggregated and averaged to measure performance for new and emerging policy initiatives. Whilst there are very sound business reasons for different views of the data, this proliferation can cause confusion. Left unchecked, and without an organisation or governance structure to provide a robust challenge, there is a danger that the current situation could become even more complex and potentially result in multiple and conflicting interpretations of the same basic performance data.
- 14 The short term improvements detailed in the report and currently being implemented are a 'stop gap' set of measures to ensure consistency in current reporting. These are being achieved through the joint working between the LSC and OfSTED. This is on a 'grace and favour' basis. This is to be applauded and the staff involved congratulated on their good working relationships. However in the longer term, this should be 'systematised' and brought under formal governance arrangements that have the agreement of the major organisations involved – particularly, DIUS, DCSF, OfSTED and the LSC.
- 15 The Information Authority has these organisations and the providers represented on its board and it would therefore seem appropriate for it to take the lead in brokering an agreement between the parties on the overall governance of the reporting and quality assurance mechanisms. The IA also has the opportunity to regularise the production of the reports through the Data Service for all parties according to agreed specifications and service levels.
- 16 Our detailed recommendations are covered in paragraphs 46 onwards. These detailed recommendations were prepared in conjunction with the teams in OfSTED and LSC who are responsible for analysing the data and users of the reports. They cover improvements in:

- (1) Consistency – The LSC and OfSTED teams have already taken steps to improve the consistency of their reports and ensure they use the same datasets for 2008 onwards.
- (2) Quality – Again the LSC/Data Service have already commissioned an investigation into how the data quality of the underlying dataset in the Individualised Learner Record (ILR) can be strengthened
- (3) Transparency – A number of improvements in making sure it is easier for providers to understand the way the success rate reports are produced and the underlying methodology. Some of these need further review and investigation with providers to ensure their needs are adequately represented.
- (4) Timeliness – There are two main recommendations:
 - a) In the short term being firm on the cut off dates and speeding up the reporting process.
 - b) In the longer term considering a different approach where the LSC/OfSTED produce reports only to an overall provider or qualification group level and not try to provide detailed reports for success rates within a provider. These could be prepared by the provider more easily and quickly as they generate the underlying data and then have to wait for the LSC and OfSTED to calculate the reports on their behalf. We understand that both the LSC and OfSTED may have some concerns over this and feel that they might lose control over the accuracy of the reports. However, the reports are only as accurate as the data supplied by the provider originally – irrespective of who processes the data and prepares the reports. In a self regulating environment, providers should be capable of and trusted to prepare their own reports. Clearly however, this approach would need several safeguards and should be considered carefully before implementation.
- (5) Methods of aggregation – OfSTED and the LSC should agree on the methods of aggregation and ideally have one approach across all reports for grouping and averaging the qualifications success rates. If there is a need to have other views of the data for sound ‘business’ reasons the basis of the alternative methods of aggregation should be clear, transparent and fully explained to the sector.
- (6) Governance (as set out in the first part of this section on the way forward)

Implementation

- 17 The final section (paragraph 76) provides an initial implementation timescale for both the short term improvements and the longer term changes.

Success rate reports

Overview

- 18 Contrary to the initial expectations as set out in your original invitation to tender we found that the LSC and OfSTED use the same basic methodology to calculate qualification success rates. The differences arise from a variety of causes which we cover in the next section. This section provides an overview of the methodology used by both the LSC and OfSTED.

Basic methodology

- 19 The basic calculation, in its simplest form, is the ratio:

$$\frac{\text{Number of learners actually achieving their qualification}}{\text{Number of learners expected to achieve a qualification excluding transfers}}$$

- 20 The joint LSC and OfSTED methodology covers the rules setting out which learners go into the nominator and denominator and how the individual success rates are aggregated. These include rules on:
- (1) Which learners and qualifications to include in the calculation – For example, it is well known that 16 year olds progressing from school often change their minds about their programmes shortly after starting in the autumn term. They then change course to another subject that they feel will suit them better. These changes are excluded from the success rate calculations as although the programme has been changed the learner is still participating and may ultimately be successful elsewhere.
 - (2) How the qualifications are to be aggregated into groups – eg automotive training
 - (3) How success rates for qualifications should be averaged within these groups. This is done differently in the various uses of success rates and can be on:
 - a) An individual basis – ie weighted by the numbers of learners taking the programme
 - b) Programme size – ie some form of weighting by the number of guided learning hours in each programme that leads up to the qualification.

The LSC and OfSTED use different approaches for averaging across the various reports produced.

FE

- 21 Reports produced for Further Education are at a qualification aim level and published annually. The following FE reports are produced:
- (1) LSC Statistical First Release (SFR)
 - (2) LSC Benchmarking data – These show overall success rates for each qualification aim and provider and the national and regional averages for comparison. These reports assist in the preparation of self assessment reports.
 - (3) Qualification Success rate reports – As each provider has a different set of subjects, qualifications success rate reports aim to present provider success rate reports based on the curriculum mix of the organisation and compare performance against national averages .

- (4) Minimum Levels of Performance – These are generated to assess how successfully providers meet the minimum acceptable levels of performance set by the Department to ensure the LSC funds support successful provision.
- (5) OfSTED College Performance Reports – These provide information and data for providers and inspectors as part of the annual assessment and full college inspections. The information complements self assessment reports produced by providers.

WBL

- 22 Work based learning success rate reports are produced quarterly at provider level. The reports exclude learners who leave courses within 6 weeks of the commencement of study, Entry to Employment programmes, Train to Gain learners and pilot programmes. The success rates are also adjusted for other matters such as transfers between programmes and for learners who are taking planned breaks in provision.
- 23 Reports are at framework level, e.g. Apprenticeship. An Apprenticeship is made up of a NVQ, a technical certificate, range of Key Skills – e.g. communications, literacy certificates – vocationally specific courses, for example a Care apprenticeship will include a course on people moving and handling.
- 24 LSC WBL qualification success rate reports are published on the LSC website using the Provider Gateway and are at qualification framework level. These reports show timely – achieved by the expected end date - and overall success rates of frameworks. Providers, Inspectors and the LSC can access these reports.
- 25 OfSTED will produce provider performance reports based on data supplied by the LSC. These reports provide data for provider self assessment reports and are used to support inspections.

Current gaps in success rate reporting

- 26 Currently there are no agreements on the appropriateness of a qualification success rate or the methodology for the following types of provision:
 - (1) Adult and Community Learning
 - (2) Ufl provision
 - (3) Adult safeguarded learning
 - (4) Overall skills for life (as opposed to Skills for Life reporting within existing funding streams ie FE and Train to Gain).
 - (5) Individual learner success - the methodology does not handle differences between individual success on qualifications and learner success – For example a learner can fail one qualification but still gain entry to their chosen employment or Higher Education on the others they have passed.

Preparation of the data used to collect success rates

The underlying data used to calculate success rates

- 27 The data in the Individualised Learner Record (ILR) is used as the basis for calculating success rates by both the LSC and OfSTED. Full details of the ILR are available on the LSC's website and we have not repeated them here. In summary, the ILR is prepared each year with interim returns made within the year and contains the details of individual learners within each provider together with their qualification aims and where appropriate

their achievements. Many of the learning programmes span two or in some cases more years. In order to prepare success rates the individual learners have to be tracked across the ILR returns commencing from the time they start the programme to the time they complete or leave the provider.

- 28 The LSC matches the current year ILR with the 4 previous years ILR data to produce success rates. Hence the underlying data used to calculate success rates for the 06/07 academic year is based on the learner records from the 06/07, 05/06, 04/05, 03/04 and 02/03 ILR returns. These returns each cover about 4 million or so learners whose records have to be matched across the different returns.
- 29 The resulting Further Education processed learner aims file, called the “FESB pre-aggregated file” has 5 years of data as the starting point for all reports produced by the LSC and OfSTED. Reports either aggregate this data, or incorporate additional information as required.
- 30 A similar methodology applies to work based learning information. An aggregated file incorporating the history data is produced which is used by the LSC for reporting work based learning measures of success. The LSC also makes available to OfSTED all activity to date of the work based learning data on a monthly basis.

Timetable for producing the underlying data and success rate reports

- 31 For WBL, the data are available throughout the year and success rates are calculated monthly with a final Qualification Success Rate (QSR) report prepared in November.
- 32 In FE, the data required to calculate success rates for a full academic year is only available when the ILR is finally ‘closed off’ in March of the following year. This period is needed to:
 - (1) Ensure the results of the exams taken at the end of the academic year in the summer have been finalised, and any appeals or requests for re marking made – This takes up to the end of November. The timetable is driven by the slowest part of the system which is for Vocational Qualifications. Here a learner can finish formal learning in July and then has six months up to the end of January to complete a portfolio which then needs assessing.
 - (2) Allow the learning providers to record the results and return them to the LSC in the ILR
 - (3) Enable the LSC to match the records of each learner across the data sets so that the progress/success of each learner can be identified.
- 33 One month after the ILR has been ‘closed’, the LSC publish the Statistical First Release (SFR) in April. They also pass the underlying dataset to OfSTED so they can produce the College Performance Reports (CPRs). The timetable is:
 - (1) The Providers send in their final ILR for an academic year by the end of March in the following year.
 - (2) LSC produce Statistical First Release (SFR) at a national level in April
 - (3) OfSTED receive the final matched ILR data (Master and demographic files, provider level data and national benchmarks) in April
 - (4) LSC publish national benchmarks on their website in August
 - (5) OfSTED produce their College Performance Reports (CPRs) in August

Level of aggregation of LSC and OfSTED success rates

- 34 The LSC success rates are initially prepared at an individual provider level as first drafts and returned to each provider as they submit their final ILR. After March all the reports are then combined at a national level to a high level of aggregation for publication in the SFR. Both OfSTED and the LSC then break the figures down into more granular details at qualification level – for OfSTED in the College Performance Reports (CPRs) and for the LSC in the national benchmarks. Both OfSTED and LSC calculate success rates at ‘map-code’ level (where qualifications of the same type and subject are aggregated together regardless of either awarding body or syllabus. The LSC publish these at national, map code and provider level. OfSTED also make them available at the same levels and Local Authority level. (The aggregation method is one of the areas of difference identified in the next section)

Issues causing real and apparent differences in success rates

Overview

- 35 OfSTED use the same underlying data to prepare their success rates as the LSC but:
- (1) The underlying data may change between the time the LSC publish the SFR and OfSTED publish CPR success rates in August. The underlying data may also change between different LSC publications. (This issue has already been identified by the LSC and OfSTED teams involved and is being addressed in this year’s reporting round. We cover these actions in more detail in the way forward section of this report.)
 - (2) The way that OfSTED group qualifications into programme areas for Work Based Learning qualifications is different from the methods used by the LSC
- 36 This section expands upon the way these differences arise. The next section goes on to review more general concerns expressed by providers during this review.

Data differences

Alterations and corrections to the underlying ILR data

- 37 The ILR and the underlying data sets used to calculate success rates are not static. There are three main reasons for this:
- (1) Correction of errors

The LSC undertake a series of checks on the data and raise any apparent errors with the provider. The provider is also given their own QSR report within 3 days of submitting their final ILR file which they can then use to identify errors. Providers should then check and verify their returns. Sometimes small alterations/corrections can make significant changes to their performance measures for subject areas. This is particularly relevant at low levels of aggregation where success rates are virtually at individual programme level and in a small group of learners one or two errors either way makes very significant differences in the performance scores. At higher levels of aggregation at say figures for the provider as a whole, the errors tend to average out and are hence less significant statistically.

Errors should be resolved by the provider by the end of March when the SFR is prepared but, if the provider is able to demonstrate that the LSC has omitted corrections in error, the LSC accept these changes after the cut off point.
 - (2) The way qualifications are classified can be changed and hence the way success rates are aggregated alters.

The LSC use the Learning Aims Database (LAD) to classify and group the qualifications in the ILR. The LAD is currently maintained by the LSC and will be transferred to the Data Service. The LAD contains details of over 90,000 qualifications that the LSC is interested in. A subset of these contains qualifications in the QCA framework where the QCA specify details of the qualification including its type and level. Not all of framework qualifications have been fully classified by the QCA. The QCA have a continuing classification programme and are progressively providing updates to the LSC/Data Service for the LAD.

The LSC use the most current version of the LAD at the time the data is submitted to analyse the ILR data and to aggregate success rates in their reports. Hence the aggregated success rates might change between different reporting dates if qualifications are re-classified. OfSTED work on the LAD codes provided to them in the matched dataset and hence do not necessarily use the same ones as the LSC for reporting purposes. Furthermore, one stakeholder commented to us that there are no formal mechanisms to alert them that the LAD had changed other than regularly checking the LSC website.

- (3) Providers may merge causing the LSC to adjust historic data and reports.

When this occurs the underlying data has to be 'merged' to represent the new organisational structures so that annual trends for the merged organisation(s) can be presented. This means that the historic data as well as the current data on success rates has to be reassessed and the success rates recalculated for the new organisation.

- 38 In theory the SFR should match the national benchmarking website as they are all based on final ILR data. However, the LSC take on these changes or make amendments meaning that the two publications can be based on different underlying data. OfSTED have the challenge of creating the CPR in parallel to the LSC creating the benchmarking data but are not always aware of changes.
- 39 In 2007 OfSTED did not find out about all the changes until they were to publish in August, four months later than originally required. This was mainly due to providers continuing to correct their ILR data after the March close down. This meant that OfSTED had to make a choice between making last minute manual changes and publishing different figures from the LSC. The LSC has taken significant steps this year to try to ensure that they do provide details earlier so 2008 should be much less problematic.
- 40 However, the basic problem remains that the reported success rates for both current and historic data can change between reporting dates due to the changes in the underlying data, its grouping into providers, and the classification of the qualifications.

Differences in grouping the qualifications

- 41 In FE, OfSTED use the same qualification groups as the LSC so when using the same underlying data sets their success rates should tally. In WBL however, whilst OfSTED and LSC figures tally at a national level, OfSTED use a different methodology to group the qualifications into programme areas based on Sector Skills Areas and termed Areas of Learning (AOL). The OfSTED approach was originally developed by the Adult Learning Inspectorate (ALI) before it was absorbed into OfSTED in 2007. LSC use Sector Subject Area Tiers (SSAT) to group qualifications. Whilst the two approaches are very similar, grouping qualifications in different ways inevitably leads to confusion as the success rates for apparently the same groups are often different.
- 42 This issue is exacerbated by terminology in the published figures. For example, the LSC success rate tables on the internet are headed 'Table 8: Success rates by area of learning, gender and programme type, 2006/07 (period 12 data)' even though the column

containing the details is headed 'Sector Subject Area'. For the uninitiated, this would imply that OfSTED and LSC are using the same classifications but in fact they are not.

Other concerns raised by providers

- 43 During the course of this review, we held discussions with staff from a sample of 5 providers and 1 provider of MIS applications to gain further views on the issues being raised. The providers we spoke to have concerns which echo the differences identified above and also provide greater insight into the use of the SRs both internally within their own organisations and during their discussions with the LSC and OfSTED.
- 44 The discussions and views of the people interviewed were fairly consistent and similar. Whilst, in our view, the sample was too small to justify any firm recommendations, given this consistency of views, we believe that their comments should be treated seriously and point to additional areas which the LSC and OfSTED might address following further more in depth examination.
- 45 We have included a list of the 'quotes' we received from them in Appendix 1. In summary their views were as follows:

(1) LSC benchmarking data

They were all generally positive concerning the use of the national and regional data available on provider and qualification aim performance. They found it useful to be able to compare their own performance with that of similar providers and helpful in determining the areas requiring most attention for improvement in their own organisations. In addition, the technical facilities allowing them to look at College Performance Reports in different ways were found particularly useful.

Set against this positive view, the WBL providers interviewed found the benchmarking data useful but difficult to compare with their own data and consequently felt that they were less able to check the data and reconcile it to their own figures.

In addition, the Colleges interviewed had generally more sophisticated internal MIS than the WBL providers probably reflecting their size and scale of operation. They felt that their own independent software provided them with better analysis allowing them to drill down into other kinds of groupings to identify pockets of poor performance needing attention.

They felt that the interactive facilities on the CPR helps but does not go far enough for their needs. In addition, they organised their internal reporting along organisational lines rather than in the LSC or the OfSTED groupings. Hence they use SRs internally to assess performance in their academic departments and have to realign their figures to match the national benchmarks and in preparation for OfSTED who use them by subject grouping. This meant that any cross department provision eg Literacy would be fragmented in the College's internal SR reports and they would have to bring them together for OfSTED comparisons with the national benchmarks which they did not always feel was helpful or useful in their own performance management.

(2) Timeliness of the data

They all highlighted the length of time it took to prepare the national figures by the LSC and OfSTED. They felt that this had a consequential impact on their performance management and inspection discussions. This led to a series of discussions at the start of an inspection where they would find that, once the accuracy was proven, OfSTED would rely on the provider's figures for the inspection but only report officially against out of date nationally prepared SRs. There was also some concern that providers would get different ratings depending

on the time of year their inspection took place as in the early part of the year OfSTED relied on expected national averages whilst towards the end of the academic year they worked on actual. They felt that this provided an 'uneven' playing field and is unfair.

(3) Weighting methodology

Some expressed concern over what is included in the success rates and the weighting methodology. The current approach to grouping qualifications (with the exception of the minimum level of performance (MLP) calculation) is based on generating a simple average of all the success rates for a subject or provider level grouping. Hence a provider delivering many short, easy to pass programmes would potentially have better success rates overall as poor performance on more challenging longer programmes would be hidden by the shorter ones. They felt a fairer and more representative approach would be an average weighted by the length of the programme or its guided learning hours (similar to the MLP).

(4) Request for consistency in measurement

They felt that as far as possible only one set of measures should be used and highlighted that various LSC programmes were asking for subtly different SR measures that meant considerable recalculation of the data and then the ability to reconcile it back to official figures.

They also highlighted the different ways measures might be treated by LSC and OfSTED – for example they perceived that OfSTED use confidence levels on the figures to determine if a provider is delivering acceptable levels of success. The MLP approach is a hard 'pass or fail' and does not take into account the statistical level of confidence in the figures. Hence they could 'pass' an OfSTED inspection but be underperforming on the MLP assessment.

(5) Technical issues:

a) Inconsistencies in the published guidance

They had found inconsistencies between the published guidance and the actual methods used to calculate SR by the LSC. The MIS supplier interviewed highlighted that there were sometimes differences between the LSC published guidance and the sample computer program code issued by the LSC. They were able to point to one current specific example – which we have passed on separately to the LSC. Whilst inconsistencies in different documents are probably inevitable from time to time in a set of calculations of this kind – it highlights the need for the LSC to have adequate feedback loops so that any such issues can be raised and resolved quickly.

b) Version control

The providers and their MI suppliers were not easily able to identify the exact versions of the various data sets to use when comparing their internal SR reports generated by their internal MIS applications with LSC and OfSTED figures. They highlighted that it would help if any performance reports detailed the precise versions of the sample program code, map-codes and LAD that should be used to re-create the SRs in their own software.

The way forward

- 46 We discussed the above findings in a workshop with key stakeholders in the LSC, IA and OfSTED. The workshop verified the general findings and identified a set of actions that would be needed to improve the situation. These are detailed here together with our

recommendations for change and improvement. We have grouped these improvements into: consistency, quality, transparency, timeliness and methods of aggregating the data, success rates and reports and also the overall governance of the processes as follows:

Consistency

- 47 This work had already commenced when we started this review. Taking note of the issues that arose in 2007, the teams of analysts in both the LSC and OfSTED who are responsible for the development of the SR reports identified a series of steps that they could take to improve the consistency of the reports produced. The full work plan is available from either Stuart Vickers of the LSC or Nicola Edwards of OfSTED. In summary they are working to improve:
- (1) Version control
Ensure that they both use the same version of the ILR data in the publication of the SFR and the LSC Benchmarking information (National Averages) - and also to make sure the same version of data is used by OfSTED in the CPRs. This involves close liaison between the two teams and ensuring that the data and changes to it are subject to rigorous version control.
 - (2) Reporting on all the data.
Previously it was common practice to consult with regional LSC colleagues as to whether any of their providers should be excluded from the "national average" reporting due to various issues that would present difficulties with showing their data. For 2007 the LSC will include all providers, and not undertake any consultation.
 - (3) LSC:OfSTED Communications
The teams have committed to maintain regular contact to address any matters arising and to keep up to date with any changes, which can be reflected in all reports.
 - (4) Discipline to ensure the cut-off date is sacrosanct
On previous occasions, there were some instances where providers had been allowed to re-submit their learner data - following which revised reports had been generated. This had effectively made the closure of the ILR in March "worthless". For reporting in 2008, the LSC will not provide any revised reporting if any new data is submitted, and will only use the data available at the beginning of April after the ILR collection has closed.
- 48 These actions should ensure that the two organisations prepare consistent reports that are easier to reconcile and where the underlying data and calculations are clearer and more transparent.
- 49 However, in addition to the current challenges there is also the potential for success rates to diverge further. At present there are 4 teams using success rates:
- (1) LSC – FfE, Quality Assurance and Intervention; Planning and Performance and Demand Led Funding. These teams are looking at refining success rates and using weighting factors to tailor success rates more closely to provision. Hence longer programmes would have greater significance in overall provider performance
 - (2) OfSTED are looking at the possibilities of introducing statistical neighbourhoods, and greater analysis through the use of confidence levels and standard deviations.
- 50 Whilst there is communication between the existing OfSTED and LSC teams this appears to be on a 'grace and favour' basis and there is limited if any overall formal

governance to ensure consistency in approach. At present the teams work together well and there is considerable good will and cooperation. In the longer term, however, we recommend that a more structured approach is taken to ensuring consistency. Without this formality, the likelihood of further divergence of or at best confusion over success rate reporting is therefore high. We return to this issue under the governance heading below.

Quality

- 51 As well as making the reports consistent the quality of the underlying data in the ILR returns could be improved. The LSC has commissioned a study into ILR data quality which is reporting in April 2008. A near final report identifies that the quality of data is not consistent across or within different types of providers and where assurance work has been undertaken, a number of data quality issues have been identified. The report makes a number of recommendations to support implementation of a data quality model.
- 52 We have not reproduced these recommendations here as they are in the final stages of being agreed. However the LSC/Data Service should take these recommendations on board as appropriate to assure improved data quality in the future.

Transparency

- 53 As well as ensuring good communications between the OfSTED and LSC teams, both organisations should respond to the concerns raised by providers in the sector for improved transparency in the figures. As a minimum:
- (1) All published reports should be subject to formal version control. Published reports should include details of the versions of the LAD and Map-codes used to analyse and aggregate the SRs so that providers can replicate the calculations using their own software.
 - (2) There should be greater attention to detail on the reports themselves so that there is no confusion over how the success rates have been analysed. This should include clarity in which analysis method is being used – ie SSAs or AOLs – and detailed explanations of the differences. Mixing the terms on the reports and the current ambiguity should be resolved.
 - (3) There should be greater attention to ensure that the published computer software code and the guidance notes are consistent. However, despite best efforts, errors of this kind are inevitable from time to time and so the LSC/Data Service should provide a central contact point where providers and their MIS suppliers can feedback any problems or discrepancies that need to be addressed. This could include the provision of a helpline for technical issues
- 54 In addition to the above, the smaller WBL providers in our sample expressed difficulty in being able to check the OfSTED/LSC figures by drilling down to the individual learners comprising the success rates for a particular grouping. Given our relatively small sample size, we are not clear if this is a widespread problem but, subject to further verification of the issue with a wider sample of providers, the LSC/OfSTED should consider providing facilities to enable providers to check their data more easily and break it down into the original learner level information.

Timeliness

- 55 Under current reporting rules and demands for confidence in the figures produced, there is probably little that can be done to improve the overall timescale. As detailed in paragraph 32(1), the timetable is driven by the slowest part of the awards system and has to wait for learners to complete a portfolio for their vocational qualifications. These

results are not known until the end of January in the year following the academic year when the learner took the main part of their qualification.

Short term

- 56 In the short term therefore, the timeliness of the reports can only be improved by:
- (1) Being firm on the deadline that is set – currently the end of March for all the returns
 - (2) Examining the possibility of bringing the deadline forward to the end of February
 - (3) Reducing the gap between preparing the SFR and the subsequent LSC benchmark reports and OfSTED's CPRs.

Longer term

- 57 In the longer term there may be another option to improve timescales but this will need careful consideration by relevant stakeholders before implementation.
- 58 The timescale at present is driven by the need for accuracy in the figures. However, on a national data set of this size, it is doubtful if there could ever be 100% confidence in the accuracy of the figures generated. There has to be some trade-off between speed of delivery of the success rates and the statistical confidence levels that can be placed on the accuracy of the data. Furthermore, for performance management purposes, trend information is important to determine if the subject area, provider, region and nation is improving over time. Providing there is relatively the same level of confidence in the figures year on year, the trend information can provide an adequate picture of the performance of the sector.
- 59 The level of confidence in the figures is also proportional to the sample size. Hence at national, regional and provider level, even taking into account errors, the accuracy of the benchmarking information probably has a fairly high level confidence and is reasonably representative of performance. Similarly, national, regional and area level benchmarks on subject groupings (SSAs or AoLs) can be prepared fairly early in the cycle to an adequate level of statistical confidence for overall performance management.
- 60 The challenge comes at lower levels of aggregation within a provider where the sample size comprising the success rates is based on sometimes fairly small cohorts of learners. Here just one error can make the difference between a satisfactory and unsatisfactory level of performance and be highly significant.
- 61 In this situation the LSC/OfSTED could:
- (1) Publish the high level national, regional and area reports on performance and subject groupings earlier together with their levels of confidence. There could then be earlier versions of the SFR culminating in the final version in April. This would provide much earlier indications of performance than currently available.
 - (2) Allow providers to prepare their own performance reports on activity within their institutions. This would have to be completed using the LSC/OfSTED agreed methodologies for grouping and aggregating data. As the providers supply the base data for the reports in any case, there should be no more questions over the authenticity of the results than there are now providing the LSC/OfSTED have some form of quality assurance over the computer software code used by the providers.
- 62 This is a transformational approach but would fit with the current philosophy and move towards self regulation. However, it is a fundamental change to the existing processes and would need careful consideration before implementation.

Methods of aggregation

63 Changing or rationalising the methods of aggregation will also require detailed consideration by the various stakeholders involved in the process. There are three issues:

- (1) The method of grouping qualifications
- (2) The methods used to average qualification success rates within a group.
- (3) Confidence levels

Grouping qualifications

64 As stated in paragraph 34 above, whilst the LSC and OfSTED use the qualification same groupings for FE they use different ones in WBL. It was suggested to us during the review that the approach to grouping qualifications should be rationalised. However, providing there is a sound reasoning for the differences and clarity as to why different groups or views of the success rates are needed we believe that the different approaches can be justified.

65 In addition, providers generally use a third approach for using success rates internally that reflects their academic structure. Hence, for example, where literacy and numeracy skills are taught in various centres, they would probably have numeracy and literacy success rate figures for each centre in order to compare the individual centre teaching staff performance with internal benchmarks as well as national ones. Grouping the success rates in different ways can add to the strength of the success rate reports for performance management and simply provides different 'views;' of the same data.

66 This points to two recommendations:

- (1) The different approaches currently used by OfSTED and the LSC towards WBL groupings should be reviewed and fully justified if they are to continue or be rationalised.
- (2) Providers should have facilities to analyse their success rates in different ways and according to their academic structures as well as to national SSA or AoL levels. This could be achieved either using their own software or potentially through an extension of the facilities provided by the LSC on the Provider portal on the website.

67 The exact approach adopted will need further consideration to test the feasibility of the different methods and whether any joint or shared service approach would be cost beneficial.

Averaging success rates

68 As stated in paragraph 49 there are several teams currently using success rates in different ways to support performance or funding decisions. Ideally all approaches should use the same methods of averaging success rates. This again will need careful consideration and agreement between the various stakeholders involved. In summary the options include (but are not limited to):

- (1) Learner level averaging – this is essentially the approach used in the benchmarks and prepares a simple average of the success rates.
- (2) Averaging weighted by the classification of short medium and long programmes. This is currently used by OfSTED
- (3) Guided learning hour averaging – this would weight the success rate of each qualification by the guided learning hours needed to deliver it.

(4) Standard learner number weighting – this is essentially the same as guided learning hour weighting but adjusted by a factor to try to relate the qualification success rate back to the numbers of learners involved.

69 Each of these has its various advantages and disadvantages. The simple learner averaging, for example, tends to favour providers who run a large number of short easy to pass qualifications. The Standard Learner Number averaging would introduce an element/indication of overall learner success.

70 Making these changes now should also take into account the needs that will be generated by the new Qualifications and Curriculum Framework over the next years. The LSC, OfSTED and other relevant stakeholders should agree and approach to be used and then apply this across all programmes unless there are compelling reasons why other methods should be introduced.

Confidence levels

71 This is to some extent an issue of consistency. However we have included it here. The need arises due the differing practice between OfSTED and the LSC MLP assessment. OfSTED calculate the statistical level of confidence in the success rate figures. So providers who fall below the MLP can still pass an Inspection if their performance is within the lower level of confidence in the figures. They would however fail the MLP assessment by the LSC.

72 We question if this approach is acceptable and feel that one measurement approach – either to use the absolute figure or use one within the statistical confidence limits – should be adopted by both organisations to avoid apparently conflicting judgments.

Governance

Preparation of the reports

73 As covered in paragraph 50, governance of success rate reporting is via OfSTED and the LSC teams and coordination of the work is undertaken on a grace and favour basis. With the formation of the Information Authority and the proposed Data Service there is an opportunity to formalise the governance of the reports and their specifications. Rather than have two organisations preparing different versions of essentially the same reports, the task could be allocated to the Data Service under specifications set out by both OfSTED and the LSC. This would:

- (1) Mean that only one team would prepare the detail of the report acting as a 'shared service' to the LSC and OfSTED.
- (2) Make it much easier to coordinate the publication dates of the various reports from the two organisations.
- (3) Make it much easier to ensure that the reports are prepared using the same versions of the underlying data
- (4) Potentially reduce costs
- (5) Still enable the LSC and OfSTED to undertake some modelling of the data themselves but relieve them of the publication tasks.

Oversight of changes to the methodology and approach to averaging and aggregation

74 Whilst OfSTED have a remit from Parliament and are responsible for setting their own performance measures for Inspection, we do not feel that this would be compromised by agreeing one method of developing the average success rates for the sector. The Information Authority is now composed of all the key stakeholders who have an interest in this subject. It would seem appropriate that they should take the lead in governing

changes to the methodology and to ensuring that stakeholders' requirements are met by the reports and publications. This role could also extend to an oversight of the specifications given to the Data Service by the two organisations.

- 75 At this stage the role of the Local Authorities and the impact the Machinery of Government changes will have is not fully clear. Any governance structure that is created should also be capable of being developed and adjusted as the new structures are clarified.

Implementation

76 The above recommendations will take some time to implement and will need a sub programme in its own right to coordinate and drive progress. The table below provides an initial timeline to progress these recommendations. As can be seen in the table it will require considerable joint working between the LSC and OfSTED.

Recommended improvements	Status	LSC / Data service	OfSTED	Providers	Target completion dates
Consistency OfSTED and LSC teams work closely together to ensure the reports produced in 2008 are consistent across both organisations	In progress	yes	yes		Aug-08 For reports prepared in 2008
Quality LSC/Data service should implement, as appropriate, the recommendations from the current data quality review	Pending report	yes		yes	Sep-08 For reports prepared in 2009
Transparency Consult with wider sample of providers to confirm transparency requirements	To be agreed	yes	yes	yes	Sep-08
Publish details of the versions of the LAD, Map-codes, matched datasets used along with the success rate reports	To be agreed	yes	yes		Aug-08 For reports prepared in 2008
Ensure clarity on the headings of all reports	To be agreed	yes	yes		Aug-08 For reports prepared in 2008
Subject to findings of the wider consultation implement a central contact point for success rate reporting technical issues and queries	To be agreed	yes			Jan-09 For reports prepared in 2009
Subject to the findings of the wider survey assess the need for further facilities to enable providers to drill down into the success rate and identify their individual learners	To be agreed	yes	yes		Jan-09 For reports prepared in 2009
Timeliness Being firm on the March deadline	In progress	yes	yes		Aug-08 For reports prepared in 2008
Examine the possibilities of speeding up current reporting and implement	To be agreed	yes	yes		Jan-09 For reports prepared in 2009
Assess the alternative approach suggested in the report so that national figures can be produced earlier and then revised and use provider generated reports for assessing performance internally	To be agreed	yes	yes		Jan-09 For reports prepared in 2009
Implement alternative approach if appropriate	To be agreed	yes	yes		Jan-09 For reports prepared in 2009
Methods of aggregation Review methods of aggregation and averaging and consult with key stakeholders on one common approach	To be agreed	yes	yes	yes	Nov-08 For reports prepared in 2009
Implement the common approach across all programmes	To be agreed	yes	yes		Jan-09 For reports prepared in 2009
Governance Agree governance structure and the involvement of Local Authorities as well as OfSTED and the LSC	To be agreed	yes	yes		Jan-09 For reports prepared in 2009
Subject to governance structure specify and agree service levels for the production of the reports with the Data Service	To be agreed	yes	yes		Jan-09 For reports prepared in 2009

Appendix 1 detail quotations from providers

Please note – these are comments received from providers – they are not all necessarily ‘well informed’ and some may not be fully accurate. However, they illustrate the providers’ views on the current reporting mechanisms.

“Being able to make comparisons with national performance in ... is great”

“LSC’s breakdown very useful and benchmark very useful”

“Interactive CPR a good model as it lets us look at performance in various ways”

“Provider gateway data analyses SRs in a different way to our internal software”

“LSC Data is too complex to make judgements and difficult to determine how it works”

“We used to be able to marry up learners that made up the success rates but can’t do this with LSC figures. Goes into years and cannot understand which learners the numbers refer to”

“OfSTED/LSC figures can be higher or lower than ours – we can’t predict which way they will be and can’t check them“

“CPR shows different benchmarks than the nationally published ones and can be different from the ones on the website”

“The colleges were given a version of the pseudo code that is different to the calculation used in LSC’s figures”

“We don’t know which version of the pseudo code, hierarchy table or LAD are the ones to use”

“Providers have to have their own independent software so they can drill down into other kinds of groupings to identify pockets of poor performance needing attention - Interactive CPR does help but does not go far enough”

“Can be duplicate benchmarks with very slight differences – which ones should we use?”

“Figures OfSTED use are out of date – At the start of the inspection OfSTED check our figures and use these in the inspection – but don’t report them. They only report the national ones”.

“OfSTED figures are different to ours as they filter out some of the short programmes and do not use partial achievement”

“Colleges that do a lot of short provision can have their results distorted by OfSTED figures”

“Different LSC regions treat SRs differently – Some pick out small pockets of low performance and concentrate discussions on them rather than overall high performance”

“When LSC were in performance discussions with our SMT they focused on a pocket of poor performance in a small cohort studying one qualification – We have 1,000s of students and 1000’s of aims – the SMT do not have that level of detail they monitor and improve performance of their departmental managers. We felt the LSC wasted our time”

“We use SRs internally by department – have to realign them for OfSTED who use them by subject groupings – hence any cross department provision eg Literacy is

fragmented in our figures and bringing them together for OfSTED tends to hide as much as it shows”

“If you have an early inspection date you will probably do better than a late one as you can work off estimated figures in the autumn”

“LSC ask for data in their various programmes in several different ways eg Train to Gain has mixed measures in the required information”