



National Audit Office

MEASURING UP

HOW GOOD ARE THE GOVERNMENT'S
DATA SYSTEMS FOR MONITORING PERFORMANCE
AGAINST PUBLIC SERVICE AGREEMENTS?

JUNE 2010

Comprehensive Spending Review 2007 covering the period 2008-2011

Review of the data systems for Public Service Agreement 11 led by the Department for Education:

‘Narrow the gap in educational attainment between children from low income and disadvantaged backgrounds and their peers respectively’

Our vision is to help the nation spend wisely.

We apply the unique perspective of public audit to help Parliament and government drive lasting improvement in public services.

The National Audit Office scrutinises public spending on behalf of Parliament. The Comptroller and Auditor General, Amyas Morse, is an Officer of the House of Commons. He is the head of the National Audit Office which employs some 900 staff. He and the National Audit Office are totally independent of Government. He certifies the accounts of all Government departments and a wide range of other public sector bodies; and he has statutory authority to report to Parliament on the economy, efficiency and effectiveness with which departments and other bodies have used their resources. Our work leads to savings and other efficiency gains worth many millions of pounds; £890 million in 2009-10.

Contents

Summary	4
Findings and conclusions for individual data systems	9
Indicator 1: Achievement gap at Early Years Foundation Stage	9
Indicator 2: Achievement gap between pupils eligible for Free School Meals and their peers achieving the expected level at Key Stage 2 and 4	11
Indicator 3: Proportion of pupils progressing by 2 levels in English and Mathematics at Key Stages 2, 3 and 4	13
Indicator 4: Proportion of children in care achieving level 4 in English and Mathematics at Key Stage 2	16
Indicator 5: Proportion of children in care achieving 5 A*-C GCSEs (or equivalent) at Key Stage 4	16
Indicator 6: The gap between the initial participation in full time higher education rates for young people aged 18, 19 and 20 from the top three and bottom four socio-economic classes	18

The National Audit Office study team consisted of: Shahryer Hussain, Georgina Light and Duncan Russell under the direction of Sid Sidhu. KPMG completed the detailed fieldwork and initial draft report working to the NAO.

This report can be found on the National Audit Office website at www.nao.org.uk

For further information, please contact:

Sid Sidhu
National Audit Office
157-197 Buckingham Palace Road
Victoria
London
SW1W 9SP
Tel: 020 7798 7281
Email: sid.sidhu@nao.gsi.gov.uk

Summary

Introduction

1. This report summarises the results of our examination of the data systems used by the Government in 2009 to monitor and report on progress against PSA 11 “Narrow the gap in educational achievement between children from low income and disadvantaged backgrounds and their peers respectively”.

The PSA and the departments

2. PSAs are at the centre of Government’s performance measurement system. They are usually three-year agreements, set during the spending review process and negotiated between departments and the Treasury. They set the objectives for the priority areas of Government’s work.
3. This PSA is led by the Department for Education – formerly the Department for Children, Schools and Families – (the Department), with data provided by a range of sources. Each PSA has a Senior Responsible Officer who is responsible for maintaining a sound system of control across Departmental boundaries that supports the achievement of the PSA. The underlying data systems are an important element in this framework of control.
4. The most recent public statement provided by the Department of progress against this PSA was in the 2009 Autumn Performance Report in December 2009.

The purpose and scope of this review

5. The Government invited the Comptroller and Auditor General to validate the data systems used by Government to monitor and report its performance. During the period October to December 2009, the National Audit Office carried out an examination of the data systems for all the indicators used to report performance against this PSA. This involved a detailed review of the processes and controls governing:
 - The match between the indicators selected to measure performance and the PSA: the indicators should address all key elements of performance referred to in the PSA.
 - The match between indicators and their data systems: the data system should produce data that allows the Department to accurately measure the relevant element of performance.
 - For each indicator, the selection, collection, processing and analysis of data: control procedures should mitigate all known significant risks to data reliability. In addition, system processes and controls should be adequately documented to support consistent application over time.

- The reporting of results: outturn data should be presented fairly for all key aspects of performance referred to in the target. Any significant limitations should be disclosed and the implications for interpreting progress explained.
6. Our conclusions are summarised in the form of traffic lights (see figure 1). The ratings are based on the extent to which Departments have:
 - (i) put in place and operated internal controls over the data systems that are effective and proportionate to the risks involved; and
 - (ii) explained clearly any limitations in the quality of its data systems to Parliament and the public.
 7. The remaining sections of this report provide an overview of the results of our assessment, followed by a brief description of the findings and conclusions for each individual data system. Our assessment does not provide a conclusion on the accuracy of the outturn figures included in the Department’s public performance statements. This is because the existence of sound data systems reduces but does not eliminate the possibility of error in reported data.

Figure 1: Key to traffic light ratings

Rating	Meaning ...
GREEN (Fit for purpose)	The data system is fit for the purpose of measuring and reporting performance against the indicator.
GREEN (Disclosure)	The data system is appropriate for the indicator and the Department has explained fully the implications of limitations that cannot be cost-effectively controlled.
AMBER (Systems)	Broadly appropriate, but needs strengthening to ensure that remaining risks are adequately controlled.
AMBER (Disclosure)	Broadly appropriate, but includes limitations that cannot be cost-effectively controlled; the Department should explain the implications of these.
RED (Systems)	The data system does not permit reliable measurement and reporting of performance against the indicator.
RED (Not established)	The Department has not yet put in place a system to measure performance against the indicator.

Overview

8. The aim of this PSA is to narrow the gap in educational achievement between children from low income and disadvantaged backgrounds and their peers

respectively. This PSA is supported by six indicators. There is a named officer within the Department responsible for each of these indicators. This officer is supported by a lead analyst. Performance against the indicators is monitored quarterly within the Department as part of its internal PSA performance reporting. The six indicators underpinning this PSA represent six mutually reinforcing priorities to monitor developments that impact on all children and young people's educational achievement. The progression targets are a major factor in helping all pupils, regardless of their background and circumstances, to succeed in English and mathematics in particular.

9. Figure 2 summarises our assessment of the data systems.

Figure 2: Summary of assessments for indicator data systems

No	Indicator	Rating
1	Achievement gap at Early Years Foundation Stage	AMBER (Systems)
2	Achievement gap between pupils eligible for free school meals and their peers achieving the expected level at Key Stage 2 & 4	AMBER (Systems)
3	+ Proportion of pupils progressing by 2 levels in English and Mathematics at Key Stages 2, 3 and 4	AMBER (Systems)
4	Proportion of children in care achieving level 4 in English and Mathematics at Key Stage 2	AMBER (Systems)
5	Proportion of children in care achieving 5 A*-C GCSEs (or equivalent) at Stage 4	AMBER (Systems)
6	The gap between the initial participation in full time higher education rates for young people aged 18, 19 and 20 from the top three and bottom four socio-economic classes	GREEN (Disclosure)

+ This indicator has been revised following the ending of Key Stage 3 testing. It now measures progression by Key Stage 2 and expected progression between Key Stage 2 and Key Stage 4 in English and Mathematics. The first reporting of progress against the revised indicator will be in the 2010 Departmental Annual Report.

10. The Department has worked to integrate the indicators within this PSA into its operational and performance management activities, for instance by integrating them into its business plan and performance reports.

11. The Department has a Data Services Group chaired by the Head of Profession for Statistics, this acts as a central point within the Department for the review of data systems underpinning the majority of the Department's PSAs.

12. The Head of Profession for Statistics has day to day responsibility for data quality issues, with direct access and accountability to the Department's Accounting Officer as required.

-
13. Director Generals are responsible for data quality in their respective areas of activity and take a proactive role in promoting high quality performance information, for example through the review of indicator definitions and involvement in the design of data systems. Furthermore, the Department's relevant members of staff receive training within this area appropriate to their roles, with regular reviews of their training needs.
 14. The Department has formal mechanisms for identifying and assessing areas of risk and reporting these to the Board. The Department's risk management processes include consideration of issues related to its PSAs.
 15. The Department undertakes internal monitoring and analysis in respect of its performance against its PSAs and the underlying indicators which support them, including the preparation of detailed reports which set out (per indicator): current performance, significant risks to performance and further action to be taken in order to mitigate the risks identified and to further achieve the Department's objectives. The Department reports performance against its PSAs to the Board on a monthly basis.
 16. Full performance is reported externally twice a year in the Autumn Performance Report and the Departmental Annual Report.
 17. Our main conclusions on the Department's overall arrangements with respect to the PSA and the indicators that it encompasses are as follows.
 - The Department is currently in the process of developing a Data Quality Strategy. This document will be used to codify the Department's overall approach to data quality, the roles and responsibilities of officers involved in data collection, data analysis and reporting. This document will then be used as the basis for ensuring data quality is embedded throughout the Department.
 - Quality control processes are undertaken either by individual Data Owners (officers responsible for data compilation), who complete these checks on their respective indicator, or through the Data Services Group. However the Department does not have a standardised quality control methodology which can guide and inform Data Owners on the processes which they must follow to ensure that data is of the required quality prior to it being used for the calculation of indicators. For example some Data Owners undertake reconciliation checks to ensure data which is transferred across IT systems is consistent; however this process may not be undertaken by another Data Owner for a data system which has a similar IT element.
 - Performance against the Department's PSAs reported within the published 2008 Autumn Performance Report contained performance reporting errors. These errors were identified after publication and corrected in subsequent versions. These were primarily due to performance data not being cleared for publication by the Data Owner. We were informed by Data Owners that they were not

aware that the data which they were producing would be featured within the Autumn Performance Report. A revised process has been implemented for the publication of performance data for the 2009 Autumn Performance Report to ensure that data reported is accurate and has been authorised for publication by the Data Owner.

- The Department has agreed measurement annexes for all of its PSA indicators, setting out the definition of the indicator and the data sources to be used. The current National Indicator Set (NIS) was introduced following the Government's Comprehensive Spending Review 2007. In the majority of cases in respect of indicators defined through the NIS, a target which measures performance has not been set. However we noted that in some cases, internal targets have been set and performance reported to the Department's Board.
- The Department does not in all cases have detailed written procedure notes in place explaining how each indicator is to be calculated and how any outliers or missing data are to be addressed. While the Department's current procedures are in most cases robust, the fact that they are not all recorded formally may make it difficult for the Department to ensure the comparability of data over time, particularly if responsibility for the calculation of performance against a given indicator is passed to a different member of staff. Where this finding has implications for individual indicators, we explore it in the next section of this report. We recommend that the Department develops for each indicator formal procedure notes setting out how the indicator is to be calculated and reported, so that this can be undertaken consistently over time and by different members of staff.
- The Department's Data Services Group has a remit to ensure robust processes are in place over the Department's data collection processes. However we noted that in some instances there are data streams which are used to compile indicators which are not reviewed by the Data Services Group. This occurs in some cases where data is provided directly to a Data Owner by another government body or an external contractor. This means that data which is used to compile indicators has not undergone an independent review to ensure it is of the required quality to support the indicator calculation. The Data Services Group is not fully aware of all the data systems within the Department which are used to compile indicators supporting its DSOs.

Assessment of indicator set

18. In undertaking the validation we reviewed the documentation associated with the PSA and considered whether the indicators selected to measure progress are consistent with the scope of this PSA.
19. There are two issues which are currently affecting the Department's measurement of this PSA. Firstly the Government's decision to end Key Stage 3 testing has a

direct impact on the reporting of indicator 3 within this PSA. The Department has revised the elements underpinning the indicator and will report progress using the new methodology in the 2010 Departmental Annual Report. Secondly, the PSA does not measure, assess or report on all children from disadvantaged groups as part of its indicator set.

Findings and conclusions for individual data systems

20. The following sections summarise the results of the NAO's examination of each data system.

Indicator 1: Achievement gap at Early Years Foundation Stage

Conclusion: AMBER (Systems)

21. We have concluded that the data system underlying this indicator is broadly appropriate, but needs strengthening to ensure that remaining risks are adequately controlled. While nearly all local authorities are quality assuring EYSP data, some are at an early stage of this process so that there remain risks around the consistency of teacher assessments.

Characteristics of the data system

22. This indicator is defined under the National Indicator Set. NI 92 measures the gap between the 20 per cent lowest achieving children and the rest at the Early Years Foundation Stage Profile (EYFSP). The Department has a target to improve the average (mean) score of children with the lowest 20 per cent of the EYFSP results, so that the gap between that average score and the median score is reduced to 33 per cent by 2011.

23. The Early Years Foundation Stage Profile is a continuous observational assessment during the academic year in which a child becomes 5. The assessment monitors development against 6 areas containing 13 scales devised by the Qualifications and Curriculum Development Authority. These are as follows:

- Personal, Social and Emotional Development (PSED) (3 Scales)
- Communications, Language and Literacy (CLL) (4 Scales)
- Problem-solving, reasoning and numeracy (3 Scales)
- Knowledge and understanding of the world (1 Scale)
- Physical Development (1 Scale)
- Creative Development (1 Scale)

24. Each scale is measured between 0 and 9, with 9 being the greatest level of achievement. The maximum score for the 13 scales is 117 with the target being each child should achieve a minimum score of 78, with scores of at least 6 in Communications, Language and Literacy, and in Personal, Social and Emotional Development.

24. The indicator is reported on the basis of individual child data produced by schools and settings which is collated by the relevant local authorities. Schools and settings provide data to local authorities through their management information system. Local authorities are responsible for validating the data and sending through data files to the departmental data collection system. Data is collated and validated within the Department and then matched into the National Pupil Database. The National Pupil Database is a longitudinal database that holds individual pupil level attainment data for all children in maintained schools in England and in non-maintained and independent schools which take part in the assessments. Early Years Foundation Stage profile data is created when pupil teacher assessment records are matched to school census records and prior achievement records. The matching process identifies and eliminates duplicate records.

Findings

25. The data system has been operational since 2003 with responsibility for calculation of the indicator being allocated to a Service Manager within the Department. The Service Manager maintains a Risks, Actions, Issues and Dependencies log for the data stream. This log records mitigating controls and required actions to ensure data is of the required quality to produce performance against the indicator. These controls and actions are monitored by a departmental working group, which holds a series of meetings to assess and manage data collection risks.
26. Consistency of pupil scoring is ensured through a moderation process undertaken by local authorities. The Department has an agreement with Ofqual to oversee this moderation process and it relies on Ofqual to ensure that data submitted by local authorities is valid and robust. The Qualifications and Curriculum Development Authority develops and modernises the national curriculum, assessments and examinations and Ofqual accredits and monitors qualifications in schools and colleges and within workplaces in the UK. Up to 31 March 2010 both of these functions were undertaken by the Qualifications and Curriculum Authority (QCA).
27. The QCA's report on implementation and moderation of the early years foundation stage profile (published November 2009) suggested that local authority mechanisms for ensuring robustness of EYFSP data had improved, including headteacher sign off of data and benchmarking against national comparatives of summary data for schools and settings. Some 98 per cent of local authorities are quality assuring EYSP data although some are at an early stage of this process.
28. The Department's data collection system validates data entered by reviewing it against expected parameters, and highlights instances of erroneous data which the Department then investigates and corrects as appropriate.

Indicator 2: Achievement gap between pupils eligible for Free School Meals and their peers achieving the expected level at Key Stage 2 and 4

Conclusion: AMBER (Systems)

29. We have concluded that the data system underlying this indicator is broadly appropriate, but could be strengthened by assessing the accuracy of data back to source documentation held by local authorities and schools.

Characteristics of the data system

30. This indicator is defined under the National Indicator Set. NIS102 is made up of two measures:

- the percentage point gap between pupils eligible for free schools meals achieving at least Level 4 in English and Maths at Key Stage 2 and pupils ineligible for free school meals achieving the same outcome. Pupils' attainment is assessed in relation to the National Curriculum and pupils are awarded levels on the National Curriculum scale to reflect their attainment (Level 4 is the target level of attainment for a pupil at the end of Key Stage 2); and
- the percentage point gap between pupils eligible for free school meals achieving 5 A*-C grades at GCSE (and equivalent), including GCSE English and Maths, at Key Stage 4 and pupils ineligible for free school meals achieving the same outcome.

31. Key Stage 2 is the stage of the National Curriculum between ages 8 and 11 years. The KS2 examinations are administered and processed by the Qualifications and Curriculum Development Authority (QCDA). Ofqual now has responsibility for ensuring that standards in the tests remain consistent from year to year.

32. Key Stage 4 is the stage of the National Curriculum between the ages of 14 and 16 years. The General Certificate of Secondary Education (GCSE) is the principal qualification at the end of KS4. Other qualifications include, National Vocational Qualifications, GCSEs in a vocational subject and Vocationally Related Qualifications. These qualifications are administered and processed by the relevant awarding (examination) bodies.

33. The QCDA and awarding bodies provide the test results for each pupil to the Department and an external contractor, appointed by the Department. The contractor processes this data and produces the Key Stage 2 and Key Stage 4 Achievement and Attainment Tables (AATs). The AATs set out the proportion of pupils achieving the relevant exam results and are used by the Department to produce the indicator.

34. Free school meal data is collected by the Department as part of the Pupil Level School Census, which contains information on personal characteristics of pupils in

schools. Pupil data within this system is then matched with the data from the AATs to produce both aspects of the indicator.

35. The Department has an agreement in place with the contractor setting out the process by which data should be assessed for robustness and the quality control procedures which should be applied to ensure that the data is accurate.
36. As an additional quality control check the Department recalculates the AATs from the data provided by the QCDA and awarding bodies to ensure that information reported within the tables is accurate. This process is documented and overseen by the Data Services Group. Any issues and inconsistencies raised through this process are discussed with the contractor prior to the publication of the AATs and the indicator. The Department monitors the contractor's risk register and has also identified the risks to data quality on its own risk register.

Findings

37. The Pupil Level School Census data is passed to an external contractor for validation. This process involves ensuring the data is as expected, within relevant parameters, consistent with previously submitted data and fit for purpose. This process of validation is documented and produces the data for identifying pupils eligible and ineligible for free school meals. We noted that the Department does not assess the accuracy of data back to source documentation held by local authorities and schools.
38. The Department has established robust arrangements to ensure the examination results data for the AATs are appropriately collected, processed, analysed and reported. In particular, the Department recalculates the AATs from the raw data and compares results with the contractor's, investigating and resolving any discrepancies prior to the calculation of the indicator. Schools are also given the opportunity to challenge the results reported in the provisional AATs prior to both the AATs and indicator being published
39. The Department has stipulated data quality and assurance requirements in its agreement with the contractor and has a programme governance model in place.
40. The Department has recently revised its governance structure for Achievement and Attainment Tables. A Steering Group monitors progress against the planned annual reporting cycle and is responsible for identifying and mitigating data quality risks and issues.

Indicator 3: Proportion of pupils progressing by 2 levels in English and Mathematics at Key Stages 2, 3 and 4

Conclusion: AMBER (Disclosure)

41. We have concluded that the data system underlying this indicator is broadly appropriate, but includes limitations over the consistency of Key Stage 1 teacher assessment data that cannot be cost-effectively controlled; the Department should explain the implications of these limitations.

Characteristics of the data system

42. The elements underpinning this indicator are defined under the National Indicator (NI) Set as follows:

- progression by 2 levels between Key Stage 1 and Key Stage 2 in English (NI 93) and Mathematics (NI 94);
- progression by 2 levels between Key Stage 2 and Key Stage 3 in English (NI 95) and Mathematics (NI 96); and
- progression by 2 levels between Key Stage 3 and Key Stage 4 in English (NI 97) and Mathematics (NI 98).

43. Key Stage 1 reflects the achievement at age 7 years; Key Stage 2 at age 11; Key Stage 3 at age 14 and Key Stage 4 reflects achievement at the end of compulsory school age.

44. In October 2008 the Secretary of State announce significant changes to the school assessment and accountability system, including the end of compulsory Key Stage 3 tests. This led to a revision of the National Indicators methodology and the development of new indicators to measure progression between Key Stage 2 and Key Stage 4. At the time of the NAO fieldwork, the revised indicators were being developed. Consequently, this report focuses solely on the systems underpinning the measurement of progression at Key Stage 2. The Department has recently published a Statistical First Release providing a time series of the proportion of pupils making expected progress in English and Mathematics between Key Stage 2 and Key Stage 4 and will report this in the 2010 Departmental Annual Report.

45. The data systems underpinning the measurement of progression at Key Stage 2 relate to teacher assessments made at the end of Key Stage 1 and tests taken by pupils at the end of Key Stage 2. The Department collects end of Key Stage 1 teacher assessment data from local authorities. The Key Stage 2 examinations are administered and processed by the Qualifications and Curriculum Development Authority (QCDA). Ofqual now has responsibility for ensuring that standards in the tests remain consistent from year to year.

-
46. The QCDA provides the Key Stage 2 test results for each pupil to the Department and an external contractor, who has been appointed by the Department. The contractor processes this data which is then matched with Key Stage 1 attainment data within the National Pupil Database. This is a longitudinal database that holds individual pupil level attainment data.
 47. The Department has an agreement with the contractor setting out the process by which data should be assessed for robustness and the quality control procedures which should be applied to ensure that the data is accurate.
 48. Separately, the contractor uses the Key Stage 2 test results to produce Achievement and Attainment Tables (AATs) which underpin other indicators. As an additional quality control check the Department recalculates the AATs from the data provided by the QCDA and awarding bodies to ensure that information reported within the tables is accurate. This process is documented and overseen by the Data Services Group. Any issues and inconsistencies are discussed with the contractor prior to the publication of the AATs and the indicator. The Department monitors the contractor's risk register and has also identified the risks to data quality on its own risk register although this is not specific to the Key Stage results used for this indicator.

Findings

49. The Department has established robust arrangements to ensure the Key Stage 2 test results are appropriately collected, processed, analysed and reported. In particular, the Department recalculates the AATs from the raw data and compares results with the contractor's, investigating and resolving any discrepancies prior to the calculation of the indicator. Schools are also given the opportunity to challenge the results reported in the provisional AATs, prior to both the AATs and indicator being published.
50. The Department has stipulated data quality and assurance requirements in its agreement with the contractor and has a programme governance model in place.
51. The Department has recently revised its governance structure for Achievement and Attainment Tables. A Steering Group monitors progress against the planned annual reporting cycle and is responsible for identifying and mitigating data quality risks and issues.
52. The Department relies on each local authority to ensure that this teacher assessment process is consistent across the authority. Local authorities are required verify the end of Key Stage 1 teacher assessment results of at least 25 per cent of their settings each year. The local authority should ensure that this verification has taken place in all of its settings within a four year period. Where it has identified anomalies in the consistency of the data during the initial verification, the local authority should undertake further monitoring. Ofqual conducts monitoring visits with local authorities to check systems they have put in place to ensure teacher

assessment data is accurate and robust. Some 25 per cent of local authorities were visited in 2008/09. This limitation is not disclosed in the reporting of this indicator.

53. The Department has recently revised the methodology for calculating this indicator due to changes in the way assessments at Key Stage 1 were carried out and in the way data is now collected. Technical notes to the methodology have been published in a recent Statistical First Release.

Indicator 4: Proportion of children in care achieving level 4 in English and Mathematics at Key Stage 2

Indicator 5: Proportion of children in care achieving 5 A*-C GCSEs (or equivalent) at Key Stage 4

Conclusion: AMBER (Systems)

54. We have concluded that the data system underlying these indicators is broadly appropriate, but could be strengthened by assessing the accuracy of data back to source documentation held by local authorities. However the Department is revising its data collection process from 2010 to use a matched administrative data source from established databases holding data at individual pupil level.

Characteristics of the data system

55. In September 1999 the Government established 11 objectives for children's social services in England. One of these was 'to ensure that looked after children gain the maximum life chance benefits from educational opportunities, health care and social care'.

56. The data system for these indicators uses data from the statistical return on outcome indicators for looked after children (OC2 return). This is an electronic data return completed by local authorities and submitted to the Department via an online data collection portal.

57. The OC2 return provides data on a number of social and educational outcomes for children in care within each local authority, including educational achievement at Key Stage 2 and Key Stage 4. Key Stage 2 is the stage of the National Curriculum between ages 8 and 11 years. Key Stage 4 is the stage of the National Curriculum between the ages of 14 and 16 years.

58. Data for the OC2 return is based on the 12 months to 30 September, with a submission deadline of 28 November each year. For the purposes of reporting educational achievement the relevant school year is that which largely overlaps with the reporting period.

59. The OC2 return reports the total number of children in care at 30 September who have been in care for at least 12 months and a number of subsets. The relevant subsets for these indicators are:

- 'number of children in Year 6 at school who were eligible for the end of Key Stage 2 tests' and 'of these; number of children who attained at least level 4 in English and Mathematics'; and
- 'number of children in Year 11 at school attempting any qualification at GCSE or equivalent (including entry level qualifications)' and 'of these; the number who obtained at least 5 A* to C grades including English and mathematics at GCSE or equivalent'.

-
60. Detailed guidance on completion of the return is available on the Department's website. The guidance gives details on the relevant inclusions and exclusions for each subset. In addition local authorities can contact the Department for advice.
 61. The Department compiles the aggregated data set from the individual returns and uses a spreadsheet to calculate totals and the relevant percentage for the indicator.
 62. It is planned that 2009 will be the last year that the OC2 return is used to collect the data and that from 2010 the Department will use data from the National Pupil Database and the Children Looked After database to compile the indicators.

Findings

63. There are inbuilt validation checks within the collection system to check that the data is submitted in accordance with the guidance notes. Once the data has been submitted the Department performs credibility checks, and do year on year comparisons.
64. The Department is revising how it collects data for this indicator from 2010. To aid this process the Department has released an "Experimental Statistical Release", which contains a comparison of outcomes for children in care, collected via the current OC2 data collection system and a matched administrative data source from the National Pupil Database and Children Looked After database. The Department has stated within the release that these are "Experimental Statistics" and the figures should be treated with caution.
65. Once the change of data stream has been confirmed the Department should ensure that policies and procedures are established which will ensure that data used is valid and robust.
66. The Department should also ensure that within any external reports it discloses that the data collection process has changed, together with a narrative explaining how this change affects the baseline and previously reported data.

Indicator 6: The gap between the initial participation in full time higher education rates for young people aged 18, 19 and 20 from the top three and bottom four socio-economic classes

Conclusion: GREEN (Disclosure)

67. The data system involved is relevant and appropriate to measure progress against this indicator; however there are limitations that cannot be cost-effectively controlled.
68. The limitations of the system have been explained in detail in the Annual Update on this statistic published in July 2009 but is not reported in the 2009 Departmental Annual Report or 2009 Autumn Performance Report.

Characteristics of the data system

69. Participation is measured using the Higher Education Initial Participation Rate (HEIPR) which is released as a National Statistic in April each year. It takes account of data from the Higher Education Statistics Agency (HESA), the Learning and Skills Council (LSC) Individualised learner Record and mid-year population estimates from the Office for National Statistics (ONS).
70. The participation rate (HEIPR) is the sum of the participation rates for each single year of age from 17 to 30. These rates are calculated as the number of first-time England domiciled entrants to higher education divided by the number of England domiciled people.
71. The socio-economic class of each student is determined using information provided from the HESA Student Record which includes information on parental occupation. Where this data is not available the Higher Education Funding Council for England's Participation of Local Areas indicators are used which uses postcode data as a proxy for socio-economic class.
72. The population of young people in each socio-economic class is estimated using data provided by the Labour Force Survey on the socio-economic class of the Household Representative Person for 13-15 year olds. The proportions are then applied to the population estimates for 18-20 year olds provided by ONS to estimate the population of young people in each class.
73. The data for the top three socio-economic classes is aggregated as is the data for the bottom four classes. The gap is the difference between the two proportions.

Findings

74. The data system is reliable and previous weaknesses identified have been addressed including:

-
- under-counting due to the exclusion of English students studying higher education in further education colleges in Wales and Scotland, as both these cohorts are included;
 - over and under-counting due to prior experience of higher education not being taken into account for LSC data but excluded from HESA data even if less than six months experience - resolved by the introduction of a new matching process;
 - over-counting due to duplicate HESA and LSC records is corrected for by using the new matching process; and
 - clarification of movement required to measure progress (one percentage point every two years).
75. A weaknesses with the system previously identified by the National Audit Office that has not been resolved concerns under-counting in higher education overseas, or private institutions, and those studying in Northern Ireland and overseas. Research into quantifying those in overseas and private institutions has been conducted and no reliable method for their inclusion found. The Scottish and Welsh components together make up less than a tenth of a percentage point, and excluding those studying in Northern Ireland is unlikely to impact on the quality of data reported.
76. The Higher Education Funding Council for England's matching files which go back to 1994 are used to identify prior higher education experience. Due to data protection a "fuzzy matching" process is used where details excluding personal information is used to identify individuals. The matching process is well designed and an effective method to identify those who have previously studied at higher education institutions.
77. Methodological changes to the HEIPR means that from 2007/08 onwards rates are not comparable to previously published rates. The Department published revised comparatives for 2006/07 academic year in the March 2009 Statistical First Release under the new methodology.
78. An additional key risk to the quality of data identified in previous validation work was the estimation of socio-economic classes, using students' known postcodes. An improved data set to estimate the socio-economic class of students (POLAR 2 - Participation of Local Areas) which mitigates the risk of an incorrect socio-economic class being allocated to students is now being used. Improvements include:
- national cohort estimates now consistent with school-year aligned Office for National Statistics (ONS) population estimates;

-
- a broadening of the definition of entrants to include a wider range of qualification aims;
 - improved methods for dealing with unrepresentative postcodes for both the population and entrant counts; and
 - moving to the 2001 Census geographical framework (the previous work used the 1991 Census geography).

79. There are still weaknesses in the data system which includes:

- volatility due to the use of several data sources;
- information on parental occupation supplied by students subject to error on both the part of the student (entering an incorrect profession) and errors coding ambiguous descriptions;
- proxy used for a proportion of students with unknown socio-economic class;
- different derivation of class for those ages over 21 which is based on their own occupation on applying; and
- proxy data used for the underlying population.

80. The implication of the above weaknesses has been fully explained in the annual update for this statistic published by BIS in July 2009, but is not reported in the 2009 Departmental Annual Report or 2009 Autumn Performance Report.