

Administrative Data Research Centre Wales
Data Science Building
Swansea University
Singleton Park
Swansea
SA2 8PP

For general enquiries please call +44 (0) 1792 606918
or to speak to a member of the Research Support Team call +44 (0) 2920 870928

ADRCWales@Swansea.ac.uk www.adrn.ac.uk @ADRCWales



An ESRC Data
Investment

Better Knowledge Better Society

Data Brief: The Welsh
National Pupil Database

Funded by



September 2015

Available in Welsh on request



The Welsh National Pupil Database (NPD) contains linked individual pupil records for all children in Wales in the state school system. It includes some key information relating to the characteristics of children in schools and contains other details relating to their educational progress, principally teacher assessments, public examination results and attendance data.

This Data Brief provides an outline of the NPD in Wales, summarising the data items that are available and particular issues surrounding the use of the Welsh NPD.

Mae Cronfa Ddata Genedlaethol Cymru am Ddisgyblion yn cynnwys cofnodion wedi'u cysylltu ar gyfer pob plentyn yn system ysgolion y wladwriaeth yng Nghymru. Mae'n cynnwys rhywfaint o wybodaeth allweddol yn ymwneud â nodweddion plant mewn ysgolion a manylion eraill yn ymwneud â'u cynnydd addysgol, gan bennaf asesiadau eu hathrawon, canlyniadau arholiadau cyhoeddus, a gwybodaeth am bresenoldeb.

Mae'r Daflen Wybodaeth hon yn rhoi amlinelliad o Gronfa Ddata Genedlaethol Cymru am Ddisgyblion, gyda chrynodeb o'r eitemau data sydd ar gael ac yn nodi problemau penodol o ran ei defnyddio.

Introduction

Within England, the National Pupil Database (NPD) first came into existence in 2002, containing linked individual pupil records for all children in the state school system. It has included a range of variables such as pupil age, gender, ethnicity, special educational needs, free school meal entitlement, key stage assessments and public examination results, details of residence and school attended. A majority of this information is available from the 2001/2 academic year, although some assessment data are available from the mid-1990s. The Department for Education and Skills provides access to a series of linked pupil data sets (including bespoke extracts) at varying levels of sensitivity for the purposes of research use. These

data sets include links to both Further Education (FE) and Higher Education (HE) records¹. Research use of these datasets has also been supported by the Pupil Leave Annual School Census (PLASC)/NPD User's Group (PLUG) based at the Centre for Management of Public Organisations (CMPO) at Bristol University². PLUG aims to promote and support the use of the NPD in educational research through the hosting of workshops and the development of metadata.

In Wales, researchers who have sought to access the NPD have applied to the School Statistics team at the Welsh Government (school.stats@wales.gsi.gov.uk). The decision to release data is made on a case-by-case basis and all data releases will be subject to a data access agreement. Supporting

documentation and metadata for the Welsh NPD has been relatively limited, although the guidance notes issued to schools to aid their completion of statutory returns have been made available from the Welsh Government website³. The aforementioned range of linked databases that are available in England (i.e. to FE and HE records) are not routinely available in Wales. However linking exercises have been undertaken by the Welsh Government and these databases have been made available to support research. Supporting research use of educational data in Wales will therefore form an important part of the work of ADRC Wales, which will include the development of new linked databases along with metadata and user-guides that support the needs of educational researchers.

Structure and Contents of the National Pupil Database in Wales

The NPD combines information on pupil and school characteristics data from the PLASC dataset with information on pupil attainment, as well as incorporating reference data on schools and Local Education Authorities (LEAs). It is generally accepted that the quality of the NPD data in Wales improved from 2004/5 onwards (for all cohorts) and therefore data prior to this period is not generally made available for research use.

A unique pupil reference number allows information held on the same pupils to be linked longitudinally. At the individual level, the NPD provides information related to three broad areas; demographic information, absenteeism and attainment. The availability of school identifiers enables researchers to aggregate the individual level data in order to derive school level variables. However, other separately derived school level variables are held by the Welsh Government and can be linked to the individual level pupil data via unique school level identifiers, such as adult pupil ratios (sometimes recorded at Year Group level or School level), whether the school is English or Welsh Medium, the governance structure of the school (Community, Foundation or Voluntary Controlled/Aided) and the religious character of the school. The Welsh Government also retains records of school closures and mergers (including records of the unique school identifiers before and after a merger).

Demographic Data

The PLASC is a census of pupils taken in January each year. The data provides individual level information on the demographic characteristics of pupils (age, ethnicity, gender) and information on special educational needs (distinguishing SEN status according to whether pupils are School Action, School Action Plus or Statemented⁴). In terms of socio-economic characteristics, information held on the NPD is limited to whether or not pupils are eligible for Free School Meals (FSM). Currently, pupils are eligible for FSM if their parents are in receipt of a range of benefits that are primarily aimed at those out of work; primarily Income Support or income based Jobseekers Allowance⁵. It must also be noted, that not all pupils who are eligible for FSM actually take them up⁶.

Attendance Data

All children receiving an education at school during the normal school day must be placed on the attendance register, which records the attendance of all pupils during each half day session (morning and afternoon) during every day that the school is open to pupils. Where a pupil is recorded as absent, the register records whether the absence was authorised or unauthorised. Definitions of authorised and unauthorised absences, as provided by the Welsh Government are as follows⁷:

- ▶ Authorised absence - an absence with permission from a teacher or other authorised representative of the school. This includes instances of absence for which

a satisfactory explanation has been provided (e.g. illness, family bereavement or religious observance).

- ▶ Unauthorised absence - an absence without permission from a teacher or other authorised representative of the school. This includes all unexplained or unjustified absences.

Absenteeism data only relates to children of compulsory school age (those aged five and above) and so children in reception classes are not covered. Individual level data shows the number of sessions that a pupil attended school in a given academic year. The total number of sessions that a pupil could have attended school is also provided.

The analysis of absenteeism derived from the NPD generally focusses on three measures:

- ▶ Days present: this relates to the time that pupils are present in school, measured in terms of the proportion of half-day sessions that pupils were in attendance.
- ▶ Persistent absenteeism: this refers to pupils who have been absent for at least 20% of half-day sessions during the school year and is a measure used by the Welsh Government in the presentation of data on pupil absenteeism⁸.
- ▶ Unauthorised absence: this relates to the proportion of pupils who have had at least one unauthorised absence during the school year.

¹ See <https://www.gov.uk/national-pupil-database-apply-for-a-data-extract>

² See <http://www.bristol.ac.uk/cmppo/plug/>

³ See <http://wales.gov.uk/topics/educationandskills/schoolhome/schooldata/ims/datacollections/pupillevelannualschoolcensus/?lang=en>

⁴ See <http://learning.wales.gov.uk/docs/learningwales/publications/131016-sen-code-of-practice-for-wales-en.pdf>

⁵ See <https://www.gov.uk/apply-free-school-meals>

⁶ See <https://stats.wales.gov.uk/Catalogue/Education-and-Skills/Schools-and-Teachers/Schools-Census/Pupil-Level-Annual-School-Census/Provision-of-Meals-and-Milk/PupilsTakingFreeSchoolMealsOnCensusDay-by-LocalAuthorityRegion-Year>

⁷ See <http://wales.gov.uk/docs/statistics/2011/111214sdr2312011en.pdf>

⁸ See <http://new.wales.gov.uk/docs/statistics/2013/130515-absenteeism-pupil-characteristics-2011-12-en.pdf>

Assessment Data

Educational outcomes data included within the NPD is captured via a combination of teacher assessments and public examinations at the end of Key Stage 4. Teacher assessments are conducted at the end of the Foundation Phase (Year 2, age 7) Key Stage 2 (Year 6, age 11) and Key Stage 3 (Year 9, age 14). Public examinations are conducted at the end of Key Stage 4 (Year 11, age 16) and will typically comprise of GCSE results or other equivalent qualifications.

Key Stage 1/Foundation Phase

The introduction of the Foundation Phase in Wales was accompanied by new methods of assessment which limit the extent to which comparisons in educational outcomes can be made before and after its introduction. The last year for which Year 2 outcomes were assessed via Key Stage 1 was 2010/11 (2011/12 being the first year in which children from the final roll-out schools were assessed via the Foundation Phase). Those who were previously assessed according to the Key Stage 1 National Curriculum were graded to one of six levels. These grades were awarded for Maths, Science, English and Welsh. A majority of pupils achieved Level 2 in each of these subject areas, Level 2 being the expected level of attainment of Year 2 pupils. Under the Foundation Phase, pupils are graded to one of seven outcomes for Personal and Social Development, Well-being and Cultural Diversity Outcomes (PSDWC), Language, Literacy and Communication Skills (LLC) in either English or Welsh first language and Mathematical Development (MD). In English medium schools, pupils are also assessed against the Welsh Development (WD) area of learning. It is only a statutory requirement for schools to compile and

report Foundation Phase assessments in areas of learning, LLC, PSDWC and MD. Under the Foundation Phase, the majority of Year 2 pupils are expected to achieve Outcome 5. There is some evidence to suggest that the introduction of the Foundation Phase was associated with a reduction in the proportion of pupils who achieved the expected level at Foundation Phase (Level 5) compared to those who achieved the expected level at Key Stage 1 of the National Curriculum (Level 2), possibly as a result of the new levels to which pupils were being assessed⁹.

Key Stage 2

Key Stage 2 assessments are undertaken at Year 6 when children are aged 10 or 11. The benefit of examining Key Stage 2 outcomes is that all children are assessed on a consistent basis, irrespective of whether or not they were previously assessed via the Foundation Phase or Key Stage 1 at Year 2. Prior to the introduction of the Foundation Phase in Wales (and as remains the case in England), National Curriculum standards were designed so that most pupils will progress by approximately one level every two years. Therefore, with most pupils having achieved Level 2 by the end of Key Stage 1, by the end of Key Stage 2 (age 11), Year 6 pupils are expected to achieve Level 4.

Key Stage 3

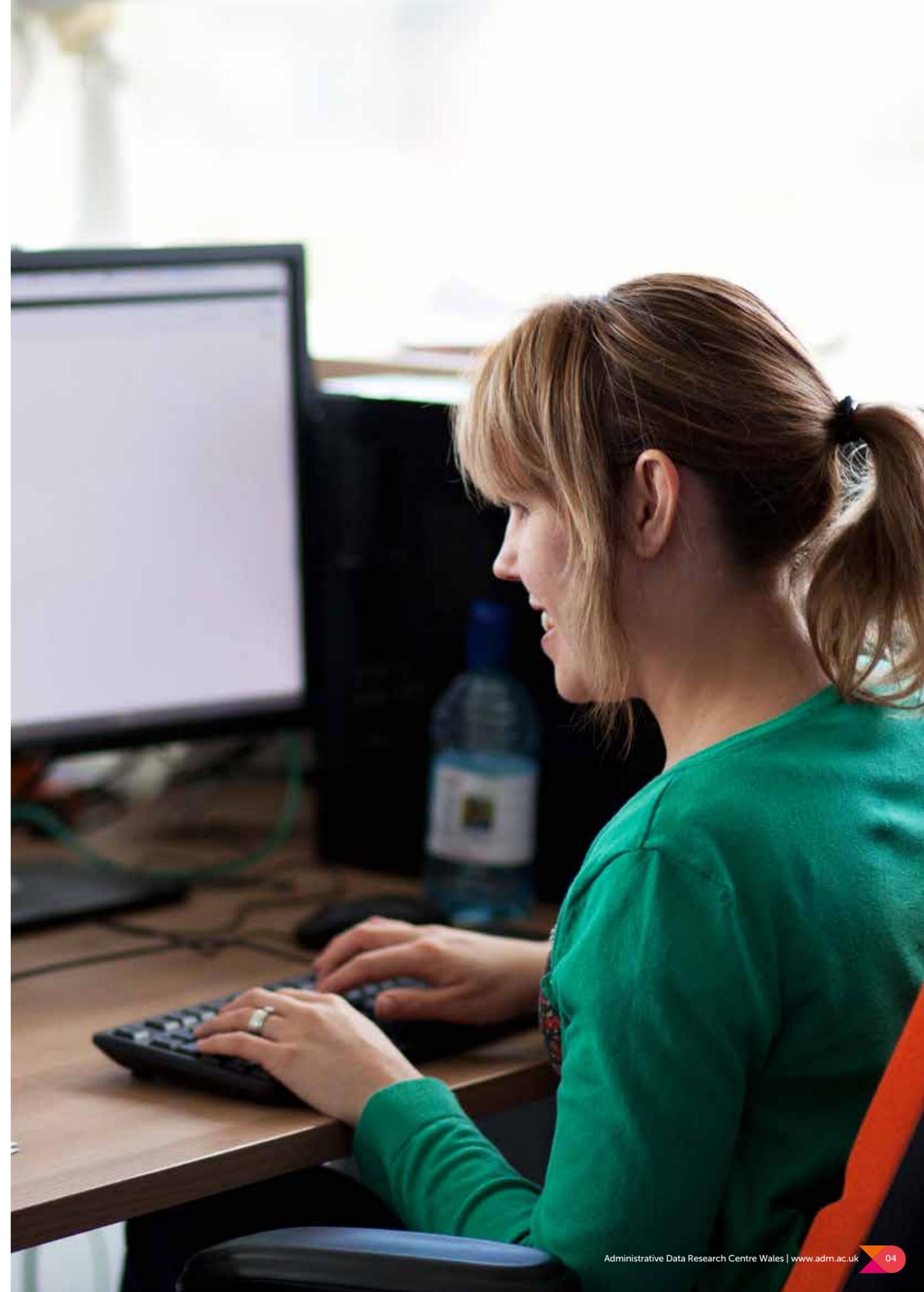
Key Stage 3 refers to three years of schooling when children are aged between 11 and 14. This is covered in years 7, 8 and 9 in England and Wales. Key Stage 3 teacher assessments are taken in Year 9 with pupils awarded Level 1-8 or exceptional performance for each subject area, with Level 5 representing the expected level of attainment.

Key Stage 4

There are a variety of ways in which attainment at Key Stage 4 is captured within the NPD. The derived core subject indicator identifies those pupils who achieved grade A*-C in each of the core subjects of English or Welsh first language, mathematics and science. Examination level data also allows researchers to construct various measures of GCSE achievement based upon standard threshold measures, such as 5 or more GCSEs A*-C; 5 or more GCSEs A*-G including English/Welsh and Maths and 3 or more GCSEs A/A*. A selection of more complex measures of attainment derived at the level of the individual and used in the publication of official statistics are also available within the NPD. These include:

- ▶ Level 1 – a volume of qualifications at Level 1 equivalent to the volume of 5 GCSEs at grade A*-G; Level 2 – a volume of qualifications at Level 2 equivalent to the volume of 5 GCSEs at grade A*-C;
- ▶ Level 2 EWM – a volume of qualifications at Level 2 equivalent to the volume of 5 GCSEs at grade A*-C but which also includes passes at English or Welsh first language and mathematics.
- ▶ Average (capped) wider points score – the summation of standardised points allocated to all qualifications approved for pre-16 use in Wales. The capped wider points score is calculated using the best 8 results.

⁹ See <http://wales.gov.uk/statistics-and-research/examination-results/?lang=en>



Observations Within the National Pupil Database in Wales

Table 1 (below) provides an overview of the number of pupils included within the NPD in Wales during 2012. The data is restricted to those pupils up to and including those in Year 11, the final year of compulsory education.

Information on approximately 435 thousand pupils is held for each year and includes data for those children based in state funded nursery settings. Table 1 highlights the availability of attainment data for those pupils in year groups which relate to the last year of a particular Key Stage. It can also be seen that attendance data is available

for a large majority of pupils. Finally, it is noted that the NPD does cover pupils staying on at secondary school beyond the final year of compulsory education. Referred to as Key Stage 5, pupils typically study for two years within such Sixth Form settings, referred to as Years 12 (Lower Sixth) and 13 (Upper Sixth). However, a comprehensive picture of participation in post-compulsory education can only be gained by combining information held on the NPD with further sources of data related to attendance in FE, referred to as the Wales Lifelong Learning Record (LLWR). An overview of data held for those attending post-compulsory education is covered in a separate data brief.



approximately
435,000
pupils' information
is held for each year

Examples of Recent Research Using the NPD in Wales

As noted earlier, the NPD is a well-established research resource in England which has supported a wider range of academic and evaluation studies¹⁰. Examples of research projects that have utilised the NPD in Wales have included the recent evaluation of the Foundation Phase (the statutory curriculum for all 3 to 7 years old children in Wales that replaced Key Stage 1 in September 2008)¹¹, the construction of the Welsh Electronic Children's Cohort; a research platform based upon routinely collected health and non-health data that can be used to answer specific questions about child health and well-being¹² and the evaluation of the Pupil Deprivation

Grant in Wales¹³. As mentioned, the Welsh Government have previously supported the construction of a linked database of school and student records that includes data from the NPD with individual learner records from the LLWR for young people who are registered at post-compulsory educational institutions and individual student records from the Higher Education Statistics Agency (HESA). At its core, the database follows the educational trajectories of three cohorts of young people who were in Year 11 (the final year of compulsory schooling) during 2004/5, 2005/6 and 2006/7. This dataset has been used to support research into widening access to HE¹⁴ and the effectiveness of financial support provided to students in post-compulsory education¹⁵.

Table 1: Observations within the National Pupil Database, 2012

Year Group	Observations	Attendance	Outcomes			
			FP	KS2	KS3	KS4
N1	5,594	0	0	0	0	0
N2	31,517	0	0	0	0	0
Reception - KS1/FP	34,423	52	0	0	0	0
Yr1-KS1/FP	33,387	33,196	24	0	0	0
Yr2-KS1/FP	33,003	32,860	32,301	0	0	0
Yr3-KS2	32,250	32,053	5	0	0	0
Yr4-KS2	31,723	31,523	0	0	0	0
Yr5-KS2	31,059	30,852	0	11	0	0
Yr6-KS2	31,982	31,757	0	31,576	0	0
Yr7-KS3	32,347	31,970	0	0	0	<10
Yr8-KS3	33,489	33,079	0	0	<10	0
Yr9-KS3	34,506	34,064	0	0	33,999	0
Yr10-KS4	35,956	35,445	0	0	<10	161
Yr11-KS4	34,214	33,537	0	0	0	33,569

Total 435,450

The Administrative Data Research Centre Wales (ADRC Wales) is one of four UK centres that make up the Administrative Data Research Network (ADRN). Together these Centres provide a safe, secure and transparent data linkage service for accredited, approved research using de-identified UK administrative data.

The Administrative Data Research Centre Wales does not hold datasets. It works closely with government departments to make them available to researchers, but this is negotiated on a case-by-case basis.

¹⁰ The PLUG workshop series provides a good overview of the range of studies that have utilised the NPD in England. See <http://www.bristol.ac.uk/cmpo/plug/workshops/>

¹¹ See <http://wales.gov.uk/statistics-and-research/evaluation-foundation-phase/?lang=en>

¹² See <http://decipher.uk.net/research-page/wales-electronic-cohort-children/>

¹³ See <http://wales.gov.uk/statistics-and-research/evaluation-pupil-deprivation-grant/?lang=en>

¹⁴ See <http://www.wiserd.ac.uk/research/education/current-projects/impact-and-effectiveness-widening-access-he-wales/>

¹⁵ See <http://wales.gov.uk/statistics-and-research/evaluation-education-maintenance-allowance-assembly-learning-grant/?lang=en>