Education Outcomes Linkage

Guidance Document

Contents

Introduction

This document provides guidance to researchers on the Education Outcomes Linkage (EOL) database. The EOL is a longitudinal education database of post primary schools' data for pupils aged 14 and over in Northern Ireland. It is the first stage of the development of a Longitudinal Education Outcomes (LEO) database for Northern Ireland. The EOL links together a number of datasets held by the Department of Education (DE) in Northern Ireland (NI); the Schools Census (including Attendance data), School Leavers Survey and Exams Database.

Included in this document are background information on the data sources that contribute to the EOL, a description of the EOL Core and how researchers can access the data.

Also included is a description of variables and caveats that are essential for the understanding of the EOL structure and EOL Core membership. When applying to use the EOL data this document should be used in conjunction with the EOL variable list.

The development of the EOL database was completed in 2022. The database will be subject to various structure changes as time progresses which can mostly be attributed to the addition of various data over time. These changes will be reported in future versions of this guidance document.

The owner of this document is NISRA-RSU and any changes to the document should be suggested to NISRA-RSU who will make amendments if required.

EOL datasets

School Census and Attendance data

The School Census is an annual snapshot of pupil and school level data. It is collected in October of each year. As well as data relating to school enrolments and numbers, information on attendance is collected as part of the overall school return.

There are some pupils that are excluded from the School Census or where caveats apply.

- Aggregate data only is supplied from independent schools. As such, independent schools will be excluded from the EOL as it is pupil level data;
- Pupils who have only ever been home schooled are excluded. Pupils educated via 'elective home education' and whose parents have requested that the pupil be removed from the school roll are also excluded. There are a small number of pupils that are home schooled/school refusers that the Department/Education Authority are aware of as they were enrolled in school at some point and so are included. These pupils are treated as 'ETA' exceptional teaching arrangements and are dual registered, i.e. they are recorded as being in their original mainstream school but also as being ETA. DE receive this list of ETA pupils as a manual return through a C2k contact and then attempt to match these pupils with mainstream pupils. ETA can also include other different types of teaching arrangements i.e. attendance at an EOTAS (Education Other Than at School) centre part of the week. DE do not know the degree of home schooling, i.e. they could attend their mainstream school 1 or 2 days a week and then are home schooled the rest of the week or they could be home schooled 5 days a week;

- Most pupils who are registered at an EOTAS Centre will also be registered at a mainstream-school. For those pupils who are dual registered, the school will have included them in their return. It is not possible to identify how much time pupils who are either home schooled or attend an EOTAS Centre spend in a school setting. Pupils who are only registered at an EOTAS Centre are also included;
- Pupils who register in a NI school after the School Census are excluded; and
- If a school closed at end of last academic year the data on attendance will be missing.

It should be noted that, as with EOTAS centres, pupils can also be dual registered between mainstream and special schools/learning support centres at other mainstream schools. These pupils will be recorded at their mainstream school.

School Census data is available for pupils aged 14 years and over for the following years:

- 2015/16
- 2016/17
- 2017/18
- 2018/19

To note the School Census data for the years 2015/16-2017/18 is only available for the 2018/19 EOL Core.

Attendance data is available for the following academic years for pupils aged 14 years and over:

- 2015/16
- 2016/17
- 2017/18
- 2018/19

To note the Attendance data for the years 2015/16-2017/18 is only available for the 2018/19 EOL Core.

Blanks And Missing

Not all pupils will have full attendance data. Where this occurs, the missing values are recorded as NULL.

The full list of School Census and Attendance variables are included in Annex A and Annex B

School Leavers Survey

The School Leavers Survey (SLS) as an annual exercise to collect and validate the qualifications and destinations of Northern Ireland school leavers. As such it includes pupils who have left the mainstream school system in the preceding academic year with home schooled and EOTAS pupils not captured within the data. The data are at individual pupil level.

There are a number of caveats to be considered when using the School Leavers Survey:

- Data excludes school leavers from special schools, independent schools, and schools that closed at the end of the academic year.
- Destinations relating to Further and Higher Education are defined by institution. Institutions may provide courses at both Further and Higher Education levels.
- Higher Education Institutions include universities and teacher training colleges.
- The number of leavers entering training include those entering the Training for Success programme, operated by the Department for the Economy. The Training for Success programme is delivered by a range of training providers, including Further Education Colleges. Training for Success trainees based at Further Education Colleges are recorded as being in training and not in Further Education. This convention avoids double counting of Training for Success trainees.
- Destination data is reported by the pupils and recorded by the school. It is not therefore a record of enrolments at Higher/ Further Education Institutions, etc. and no further validation of destinations is carried out.
- The recorded leaving date may not always reflect the actual date a pupil left the school system. The end of the school year can at times be reported by a school when the actual date of leaving is unclear.
- Data relates to pupils recorded as leaving the school system from the October census date in the previous academic year up to the day before the October census date in the current academic year.
- Data relates to pupils aged 14 and over on the first day of the July prior to the academic year and who were subsequently recorded as leaving from curriculum years 12 and above.

Blanks And Missing
Not all pupils will have a SLS record in each academic year, the main reason for this is that they have not reached the age to leave compulsory education. Where this occurs the missing values are recorded as NULL

The full list of School Leavers Survey variables is included in Annex C.

Exams Database

The Exams Database is created using examinations and pupil data extracted from the school IT system (SIMS) as part of the annual School Census process that takes place each October. It includes information for all current pupils (at the time of Census) as well as for all school leavers during the previous academic year. Information for pupils in schools that closed at the end of the previous academic year is extracted in a separate process and merged with the Census exams data. The data covers pupils in all post-primary schools in Northern Ireland, including attainment by school pupils who took exams at, or in collaboration with Further Education colleges. The Exams Database includes exams taken in the current year and previous exams by the pupil.

There are a number of caveats to be considered when using the Exams Database:

- Data excludes special schools and independent schools.
- Data excludes attainment by home schooled pupils.
- EOTAS pupils are not captured within the data.
- The majority of the data is taken from Northern Ireland Entitlement Framework Qualifications Accreditation Number (NIEFQAN) files maintained by the Council for the Curriculum, Examinations and Assessment (CCEA) Regulation. The NIEFQAN collates information, relating to qualifications approved for teaching in Northern Ireland, from the OFQUAL Register of Regulated Qualifications. As an administrative dataset, created to assist schools with planning their curriculum offer, a level of error within the data is expected.
- The "QUAL_LEVEL" variable (e.g. GCSE/FC*, BTEC/CE3, NVQ/L2 etc.), however, was created to enable the Schools Information Management System (SIMS) software to link individual qualifications to the appropriate grade sets (e.g. A*-E, 9-1, Distinction/Merit/Pass etc.). As more than one "level" will link to the same grade set, and as an administrative variable simply required to link a qualification to a specific grade set, a level of error in the "QUAL_LEVEL" variable is to be expected.
- The "EXAM_MTH" variable should relate to the month in which an examination was sat. On occasion, however, it may erroneously relate to the month in which the qualification was entered/applied for. The results data within SIMS can be received electronically (EDI – Electronic Data Interchange) directly from Awarding Organisations or added manually (Non-EDI) by schools. Any results added manually by schools are allocated a "EXAM_MTH" of 6 (i.e. June). Results need to be added manually where entries were not applied for electronically, by schools, through the C2k School System.
- General validation work is undertaken by DE statisticians on this administrative database in order to, for example, improve the internal consistency of the data, remove duplicated entries etc. No further validation is undertaken directly with schools.

Blanks And Missing

It should be noted that when qualifications are no longer offered, they drop out of the NIEFQAN file, but results can still be attained a few years after this. New and old versions of the NIEFQAN files are combined in order to include as many of the results as possible and these may go back a number of years. The "GLH", "SEC_SUB_AREA" and "STEMFLAG" variables are recent additions and so for older qualifications this data does not appear in the older NIEFQAN files and appear in the EOL as Null.

The full list of Exams Database variables is included in Annex D.

Collection of data

The School Census is collected in the period October to December each year and captures all those enrolled in the school for that academic year. Therefore, the School Census data included in the EOL 2018/19 was collected in the autumn of 2018.

Data on attendance is collected each year as part of the School Census exercise but relates to attendance in the previous school academic year. Therefore, the Attendance data included in the EOL 2018/19 was collected in the autumn of 2019. Similarly, data contained within the Exams Database and School Leavers Survey are collected in the period October to December each year and relates to exams taken/leavers in the previous school academic year. Therefore, the exams and leavers data included in the EOL 2018/19 were collected in the autumn of 2019.

EOL Core

The population in the EOL is comprised of all pupils aged 14 years and over; returned in the 2018 School Census with an exam record in the 2018/19 Exams Database. There will be no unlinked records provided to researchers. The EOL Core for 2018/19 is comprised of 52,806 pupils (please see Table 1 below for a summary of the number of pupils included). Future School Census and exam records will allow us to increase the sample size of the EOL. In subsequent years, additional pupils appearing in the Census and fitting the criteria will become live and be added to the EOL Core table.

Core EOL variables

The core variables for the EOL are listed below. These will be stored in a separate long file with people appearing only once. The variables in this file are mostly considered to be static and should not change from one census to the next. The exception would be age, which will increase by 1 each year. Researchers will receive an extract from this table as a bespoke standalone table so that they can link to the other tables once they have identified their research core population.

- EOL_PID

- GENDER

- RLGN_5WAY

- ETH_5WAY

- JUL_AGE_2018

- HOME_LANG_AGG

Table 1: Summary of EOL 2018/19 – number of pupils

	School Census	Exam Database	Attendance	School Leavers Survey
2018/19	52,806	52,806	52,661	20,602
2017/18	46,578	N/A	46,086	N/A
2016/17	25,582	N/A	25,288	N/A
2015/16	13,176	N/A	13,060	N/A

Data access

The EOL data may be accessed by accredited researchers who have successfully submitted a project application to RSU. RSU will assist researchers from the beginning of their application proposal to the close of the project. Included in the proposal will be a variable list and researchers are required to justify the inclusion of each variable within the scope of their projects aims and objectives. In addition, all projects must be accredited by the UK Statistics Authority Research Accreditation Panel (RAP) and have ethical approval. DE will give final approval to all projects.

Upon approval by RAP and DE, researchers will be provided with a suite of linkable tables supplied as flat files. The data will be made available to researchers in NISRA's secure environment or via the Office for National Statistics (ONS) Secure Research Service (SRS).

Acknowledgements

When publishing or presenting findings from the EOL the following acknowledgement should be used:

The Administrative Data Research Network takes privacy protection very seriously. All information that directly identifies individuals will be removed from the datasets by trusted third parties, before researchers get to see it.

All researchers using the Network are trained and accredited to use sensitive data safely and ethically, they will only access the data via a secure environment, and all of their findings will be vetted to ensure they adhere to the strictest confidentiality standards.

The help provided by the staff of the Administrative Data Research Network Northern Ireland (ADR-NI) and the Northern Ireland Statistics and Research Agency (NISRA) Research Support Unit is acknowledged. The ADR-NI is funded by the Economic and Social Research Council (ESRC). The Department of Education data have been supplied for the sole purpose of this project.

EOL_PID

This is a unique number that will allow Pupils to be followed over time. If a pupil moves school, the old school must send their common transfer file (CTF) to the new school. There are some cases where a school believes a pupil to be new to NI and they will issue a new UPN not knowing that they have attended another school previously. This variable will facilitate linkage with other tables.

YR_CEN

Academic Year that the School Census was Held.

GENDER

Gender of the EOL Member (Will be held in the CORE table).

JUL_AGE_2018

Calculated Age of the EOL Member in the July Preceding Collection of the School Census, 2018 (Will be held in the CORE table).

YOUNG_IN_YEAR

This is a derived variable created by DE . It identifies any pupils who have a 'late' birthday in April, May, June or the 1st of July. These pupils are the youngest in the year group. Any pupils who had a 'late' birthday but were a year (or more) older that the expected age have been omitted as they will actually be among the oldest in the year group.

RLGN_5WAY

Religion of the EOL Member (Will be held in the CORE table).

ETH_5WAY

Ethnicity of the EOL Member (Will be held in the CORE table).

FSME

Entitlement Status of the EOL Member for Free School Meals.

ACTUAL_YR_GRP

This is a derived variable created by DE to reflect the year of study, and adjusted to reflect pupils repeating years. For example, a pupil who is the expected age for year 13 but is repeating the final year of their GCSE's will be recorded as a year 12 and not as a year 13.

SEN_STAGE_5

This is the 1998 SEN Code of Practice for children who have Special Educational Needs (SEN). The code has five stages (see data dictionary). The Department of Education is currently finalising a new Code of Practice to replace the 1998 Code. This new Code, will record children on the SEN register under three stages. This will introduce a new variable to the EOL dataset SEN_STAGE_3 in future School Census (2021/22 onward). For more information on these variables see the below link.

https://www.education-ni.gov.uk/publications/de-circular-202106-three-stages-special-educational-provision

COG_TOT

SEN Cognitive and Learning Needs.

SEB_TOT

SEN Social, Emotional and Behavioural Needs.

COM_TOT

SEN Communication and Interaction Needs.

SEN_TOT

SEN Sensory Needs.

PHY_TOT

SEN Physical Needs.

MED_TOT

SEN Medical Conditions/Syndromes Needs.

OTH_TOT

SEN Other Needs.

PRIM_NEED_TYPE
SEN that school has placed in the primary position as most predominant.

DOA_YR

Date that Pupil Arrives at the School (Year).

DOA_QTR

Date that Pupil Arrives at the School (Quarter).

REG_STAT

Registration Status. Dual Registration applies when a pupil enrolled in one school attends another school, as a result of a SEN or EOTAS referral.

LSC_ATT

Pupil is in a learning support centre.

LSC_TYPE

Type of learning support centre. This variable is not available prior to the 2017/18 School Census.

SETT_15_P

Settlement Band for 2015 NISRA Urban/Rural Classification for Pupil's Home Address (Aggregate). This variable is missing for a number of pupils due to invalid or missing address information.

MDM_QUIN_P

Northern Ireland Overall Multiple Deprivation Measure for Pupil's Home Address (In Quintiles with 1 most deprived and 5 least deprived). This variable is missing for a number of pupils due to invalid or missing address information.

MDM_EDU_QUIN_P

Northern Ireland Education, Skills and Training Deprivation Domain for Pupil's Home Address (In Quintiles 1 most deprived and 5 least deprived). This variable is missing for a number of pupils due to invalid or missing address information.

TRAV_DIST_R

This is the distance measured in miles between the pupil's home address and the school address, as the crow flies. This variable is rounded to whole miles and bottom coded at one mile and top coded at twenty miles.

LGD_14_P

Local Government District (2014) for Pupil's Home Address. This variable is missing for a number of pupils due to invalid or missing address information.

HOME_LANG_AGG

First Language Spoken at Home. Aggregated to 13 levels. (Will be held in the CORE table).

NEWCOMER

A Newcomer pupil is one who has enrolled in a school but who does not have the satisfactory language skills to fully participate in the school curriculum, and the wider environment, and does not have a language in common with the teacher, whether that is English or Irish.

TIME_AS_NEW

Length of Time the Pupil is Recorded as A Newcomer (years). Pupils are normally registered for three years, with extensions granted in some circumstances. It is very normal for pupils to be recorded as NEWCOMERS for longer than 3 years. Those that are 0 instead of blank, are newcomers who have a start date that was less than a year previous. An assessment against the Common European Framework of Reference for Languages (CERF) must be carried out to provide evidence that the child still meets at least one of the criteria to be classified as a NEWCOMER.

EVER_NEW

Has Pupil Ever Been a Newcomer.

CH_LKED_AFT

Child is Looked After (Child looked after is defined as being 'in the care of the authority; or provided with accommodation by the authority. This includes children who are fostered'.

PREV_REG_SS

Pupil previously registered with a Surestart developmental programme for 2 and 3 year olds.

PREV_ATT_SS

Pupil previously attended a Surestart developmental programme for 2 and 3 year olds.

LEVEL_STUDY

This variable was created by DE to validate those studying at level 3 for funding purposes. Data is only completed for pupils in Years 11 – 14. It is not available prior to the 2016/17 School Census.

FSME_6

Pupil has an entitlement status for Free School Meals on Census day in any of the Last 6 years**.** Not available prior to the 2018/19 School Census.

EOL_SCID

System generated Unique Reference Number for School. This variable is **restricted.** To be used for modelling work only and no analysis will identify a school.

SC_TYPE

Type of school.

SC_MAN_TYPE

Management type of the school.

SC_MAN_TYPE1

This variable is derived from SC_MAN_TYPE and shows a more detailed breakdown of management type of voluntary schools.

IME

Irish Medium Education. This variable is a combination of 'In Irish medium unit' and 'In Irish school'. Because of the small number of schools involved this is a **restricted** variable and will only be given out with pupil level data. No school identifier variables will be given out with this variable.

SETT_15_SC

Settlement Band for 2015 NISRA Urban/Rural Classification for School's Address (Aggregate).

MDM_QUIN_SC

Northern Ireland Overall Multiple Deprivation Measure for School's Address (In Quintiles with 1 most deprived and 5 least deprived).

MDM_EDU_QUIN_SC

Northern Ireland Education, Skills and Training Deprivation Domain for School's Address (In Quintiles with 1 most deprived and 5 least deprived).

LGD_14_SC

Local Government District (2014) for School's Address.

YR8_12_SIZE

This variable provides information on the size of the school in increments of 50. The total will include all pupils who are enrolled at the school as per census. This variable is **restricted** with limits applied in relation to combining this with school LGD. Any request for this variable in combination with pupil geography will be subject to further SDC assessment.

YR13_14_SIZE

As above. Also included are pupils repeating year 14. This variable is **restricted** with limits applied in relation to combining this with school LGD. Any request for this variable in combination with pupil geography will be subject to further SDC assessment.

Annex B Attendance variables

EOL_PID

System generated Unique Reference Number for EOL Member.

RATE_ATT

Proportion of School Days Attended by Pupil (%).

RATE_ABS

Proportion of School Days Not Attended by Pupil (%).

RATE_AUTH

Proportion of Authorised Absences (%).

RATE_UNAUTH

Proportion of Unauthorised Absences (%).

ATT_YR

Academic Year to which attendance relates.

EOL_PID

System generated Unique Reference Number for EOL Member.

SLS_YR

Academic Year to which School Leavers Survey relates.

DATE_LEFT_MTH_YR

Date Pupil Left School (Month and Year).

LEAVE_CODE

Reported Destination and Level of Entry of School Leavers.

DEST

Reported Destination of School Leavers.

EOL_PID

System generated Unique Reference Number for EOL Member.

QUAL_LEVEL_SHORT

Qualification level Short.

EXAM_YR

Examination Year (calendar).

EXAM_MTH

Examination Month.

EOL_GRADE

Grade Achieved

Notes:

- '5-5' is a double award qualification that uses the 9-1 English grades instead of the A*-E grades. It would be equivalent to a C*C*.

- The variable has both 'F ' and 'Fail' . For the EOL 'F' represents F grades (i.e. between an E and a G grade) rather than "Fail" grades. However, it should be noted that in some instances an F may represent a 'Fail'.

- The variable has 'Merit 2' and 'Merit Merit' The "Merit 2" appears in the database for Level 3 Pre-University Certificate courses that are the equivalent of one A-Level, so could not be a Merit Merit (which would infer a double award). The available grades for these courses are D1, D2, D3, M1, M2, M3, P1, P2, P3 (i.e. Distinction, Merit and Pass with 1, 2 or 3) and are valued in that order (i.e. a D1 is better than a D3, and a D3 is better than an M1 etc.). A "Merit 2" appears to simply be a "Merit 2".

- The variable has 'P1' and 'Pass'. A "P1" relates to a "Pass at Level 1".

- The variable has 'P2' and 'Pass Pass'. A "P2" relates to a "Pass at Level 2". However (and the following would be relevant to "P1" as well), if a qualification is only at Level 1 or at Level 2 (and not both) then it may just be a "Pass". So a "Pass" could actually be a "P1" (Pass at Level 1) or a "P2" (Pass at Level 2).

QUAL_LEVEL

Qualification level

EF_DESIG

Entitlement Framework Designation. 'The Entitlement Framework (EF) is the Post-14 curriculum which puts the needs of pupils first'.

AWARD_ORG

Awarding organisation (anonymised).

KS4_ENT

Key Stage 4 Size Equivalence.

KS4_AC

Key Stage 4 A* - C Size Equivalence.

KS4_AE

Key Stage 4 A* - E Size Equivalence.

KS4_AG

Key Stage 4 A* - G Size Equivalence.

LEV3_ENT

Level 3 Size Equivalence.

LEV3_AC

Level 3 A* - C Size Equivalence.

LEV3_AE

Level 3 A* - E Size Equivalence.

GLH

Guided Learning Hours.

SEC_SUB_AREA

Sector Subject Area.

STEMFLAG

STEM Qualification Flag.

STAGE

Qualification Stage.

SUBJECT

Qualification Subject.