PIONEERED

Pioneering policies and practices tackling educational inequalities in Europe

Deliverable No. 4.1

Data Harmonisation Guidelines

The PIONEERED project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004392.

Grant agreement no.	101004392
Project full title	PIONEERED – Pioneering policies and practices tackling educational inequalities in Europe

Deliverable number	**D4.1**
Deliverable title	**Data Harmonisation Guidelines**
Type[1]	R
Dissemination level[2]	PU
Work package number	WP4
Work package leader	LISER
Author(s)	Taylor KROEZEN, Aigul ALIEVA
Keywords	Data; survey; harmonisation; education; inequality

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101004392.

Acknowledgments

We thank the project partners for their contributions in identifying new and relevant surveys and studies in their respective countries. We also thank UBERN for the report on harmonisation and standardisation prepared as a part of Task 2.2b in Work Package 2, and thank UCM for their review and feedback of the deliverable. The content of the current deliverable, including opinions expressed and any remaining errors, is the responsibility of the authors.

[1] **Type**: Use one of the following codes (in consistence with the Description of the Action):
- R: Document, report (excluding the periodic and final reports)
- DEM: Demonstrator, pilot, prototype, plan designs
- DEC: Websites, patents filing, press & media actions, videos, etc.
- OTHER: Software, technical diagram, etc.

[2] **Dissemination level**: Use one of the following codes (in consistence with the Description of the Action)
- PU: Public, fully open, e. g. web
- CO: Confidential, restricted under conditions set out in the Model Grant Agreement
- CI: Classified, information as referred to in Commission Decision 2001/844/EC

Content

1 Introduction

The main objective of deliverable D4.1 is the preparation of data harmonisation guidelines together with the accompanying coding files (written for Stata programme). These elements will assist the research teams in harmonising country-specific and international datasets and prepare them for the analysis. The deliverable consists of two main parts. The first part in section 2 starts with an overview of country-specific and international data sources, identified both during the PIONEERED project preparation stage, as well as during the preparation of the deliverable with the assistance of national partners. The goal was to provide an as comprehensive as possible list of data sources that address education-related themes throughout a life course, from ECEC to post-secondary/tertiary, as well as adult education in each partner country. Additionally, some comparative surveys of adult population and households that can contribute to a better understanding of the reproduction of educational inequalities and impact on society in large are included. An overview of each data source provides brief information regarding the focus of the study or survey, its main target group(-s), as well as general access conditions.

The second part (in section 3) of deliverable D4.1 focuses on the harmonisation of key variables across the surveys and studies that were retained for the empirical analysis. The first group of variables include those pertinent to the main axes of educational inequalities, such as socio-economic background, migration & language background, gender and home resources. The second group consists of a set of variables that measure educational outcomes either directly, such as achievement results in reading, mathematics, but also related to education, such as extra-curricular activities, sense of belonging and educational aspiration. Each of these variables was reviewed with respect to its definition, measurement, and availability in each dataset and wave in case multiple waves of data were available. Following this step, the harmonisation solution was proposed for each variable and coding syntax was written for each variable and respective survey. The coding syntax (in Stata) is not included directly in this guideline, but can be provided upon request.

The original status of the D4.1 was set to confidential and is, therefore, shared with the project members only. This is the revised and condensed version of D4.1 which is an additional public version of the deliverable.

2 Overview of the Datasets

In this section, we outline the national education data available in each of the PIONEERED participating countries, as well as international education databases. We provide a brief overview of the data and its accessibility.[3] This information serves to offer a clearer picture of the data possibilities and assists in narrowing down the inventory of data that will be used in the quantitative research of PIONEERED.

2.1 Finland

KAPPAS!

The **Assessment of Undergraduate Students' Learning Outcomes in Finland (Kappas!)** is a cross-sectional study of the generic skills of undergraduate students attending university or university of applied science in Finland. The project is a joint initiative of the Finnish Institute for Educational Research (KTL), the University of Jyväskylä and the Centre for University Teaching and Learning of the University of Helsinki (HYPE). The testing was administered online between 2019-2020, to students at the beginning and end of their undergraduate studies. Its aim was to collect information on the development of generic skills over the course of higher education, it included eleven universities and seven applied science universities and a final sample of 2,402 students (Ursin *et al.* 2021). The generic skills in the study were measured using the Collegiate Learning Assessment CLA+ instrument, and included problem solving skills, analytical reasoning and evaluation, writing mechanics and argumentative writing (Ursin *et al.* 2021). The Kappas! data is not publicly available but the results have been published by the Ministry of Education and Culture in a final report (Ursin *et al.* 2021).

2.2 Germany

NEPS

The **National Education Panel Survey (NEPS)** is a multi-cohort longitudinal study conducted by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg. The first wave of sampling was done between 2009-2012, with yearly waves in the following years. The aim of NEPS is to collect data across the lifespan of German students, focusing on the development of competencies, educational processes, educational decisions, and returns to education in formal, nonformal, and informal contexts. The six different cohorts each have their own target populations; starting cohort 1 (SC1): new-borns, starting cohort 2 (SC2): kindergarteners, starting cohort 3 (SC3): grade 5s, starting cohort 4 (SC4): grade 9s, starting cohort 5 (SC5): first year students and starting cohort 6 (SC6): adults. The NEPS includes data on competence measures, learning environments, educational decisions, migration background, socio-demographic information, self-concept and motivation, and transitions between major education stages (Skopek *et al.* 2012a, 2012b, 2013, FDZ-LIfBi 2022a, 2022b, 2022c). Additional cohort specific topics are included and discussed in more detail on their webpage (NEPS Website). The data is available in three different formats, the download data is the most accessible and anonymized followed by the RemoteNEPS and On-site versions. The NEPS download data is available to individuals with university degrees and employed by a scientific institution, and access is

[3] Due to the lack of availability of technical documentation or language barriers for some surveys and studies, some information might be missing.

granted to registered users with a signed data use agreement, see the data access section of their webpage.

CILS4EU-DE

The **Children of Immigrants Longitudinal Survey in Four European Countries - Germany (CILS4EU-DE)** builds upon the original CILS4EU which is a longitudinal study of immigrant and native students from the age of 14 in four countries: England, Sweden, the Netherlands, and Germany. The CILS4EU consists of 3 waves of data, starting in 2010 and followed up yearly the next two years. CILS4EU-DE follows up on the German sample and comprises an additional six waves of data collected annually from 2014 to 2016 and biannually thereafter. The specific topics covered in the study depend on the wave, but in general the study collected data on leisure time, family and friendships, feelings and beliefs, school and educational attainment, occupation and vocational training, with a special wave of the survey dedicated to Covid-19 and the consequences of the pandemic in Germany (Kalter *et al.* 2021). Both the CILS4EU and CILS4EU-DE data can be applied for from GESIS Data Archive for the Social Sciences (GESIS Website).

2.3 Hungary

NABC

The **National Assessment of Basic Competencies (NABC)** is a yearly assessment of students in grade six, eight and ten, which follows the student cohorts every two years. The assessments cover both mathematical skills and reading literacy. The NABC also includes background data on the students and the institutions. The institutional data covers school equipment and school features, while the student level data collects information of the family's social, cultural and economic background (KRTK Adatbank 2020). The target population of the NABC includes every student in grades six, eight and ten except for those who were absent or refused to take the test and excluding schools that solely educate special education needs students. Most of the documentation including the assessment tests, student questionnaires and school questionnaires are only available in Hungarian. The NABC data is available upon request and generally requests are accepted by the chief of the Databank in every case (KRTK Adatbank 2020).

HLCS

The **Hungarian Life Course Survey (HLCS)** was a single cohort panel survey completed by TARKI Social Research Institute. The aim of the survey was to study the secondary and tertiary opportunities of students in Hungary while analysing students' disadvantages. The initial sample was constructed from the eighth grade students both Roma and non-Roma, that participated in the NABC in 2005-2006, oversampling students with special needs and in the lower third of competence scores. The HLCS had a total of six yearly waves and an initial sample size of over 10,000 students. The content of the survey differed over the waves, beginning with a focus on family environment and early schooling, followed by school completion, further studies and the eventual transition to the job market (TARKI Foundation 2006, 2007, 2008, 2009, 2011, 2012). Other topics of the survey include housing, wealth, dropout rates, leisure time, drugs and alcohol consumption (TARKI Foundation 2006, 2007, 2008, 2009, 2011, 2012).

The HLCS dataset is available in Hungarian and the documentation in both Hungarian and English, a data request sent to TARKI is required to access this dataset.

2.4 Ireland

Growing up in Ireland

Growing up in Ireland (GUI) is a government-funded longitudinal study conducted jointly by the Economic and Social Research Institute (ESRI) and Trinity College Dublin. The survey began in 2006 and follows two different cohorts, the '08 Cohort also known as the Infant Cohort and the '98 Cohort also known as the Child Cohort. The main objectives of the GUI survey are to describe the lives of children in Ireland, chart their development over time and identify factors that hinder children's development or even further lead to social disadvantage and educational difficulties as well as look at the effect on later life of early childhood experiences (Economic and Social Research Institute 2021). Additionally, the most recent wave of the survey conducted in December 2020 was a special COVID edition. The target population depends on the wave and cohort of the survey, in the different waves information is collected from the students, parents, caregivers, teachers and principals. Growing up in Ireland additionally includes some assessment tests. At nine years old, the child cohort is tested in reading and math, while the infant cohort is tested on vocabulary, and at thirteen years old, the child cohort is tested in vocabulary and numerical reasoning (Thornton *et al.* 2010, Quail *et al.* 2011, 2014, Murray *et al.* 2013, 2015). There are two types of datasets available for the GUI, the anonymized micro data (AMF) files are the basic set of variables approved for distribution, where the research micro data files (RMF) contain more detailed information but require a data training course and can only be accessed within Ireland. The AMF files are available to researchers and a data request form can be submitted to the Irish Social Sciences Data Archive (ISSDA).

2.5 Lithuania

NMPP

The **National assessments of student achievement (NMPP)** by the National Education Agency (NSA) began in 2012 collecting assessment and background data of Lithuanian students in the fourth and eighth grades. Since 2016, students in the second and sixth grades were included as well. Since 2020 the student background questionnaires are no longer administered, thus only assessment data is available. The national assessments focus on the following subjects in the second and fourth grades: reading, writing and math. In the sixth and eighth grades, literacy, writing, and math are assessed, with the addition of social science in the eighth grade. The aggregated student assessment data is released for use, but the student background data and questionnaires are only available upon special request via the National Education Agency. The tests and questionnaires were administered in Lithuanian, Russian and Polish, other languages and translations are most likely unavailable.

2.6 Luxembourg

ÉpStan

The **Épreuves Standardisées (ÉpStan)** is a national school monitoring program evaluating student competencies in key school subjects to determine whether educational goals from the previous learning cycle were met. The national assessments take place each year and monitor students who are starting a new learning cycle, more precisely this includes students in grades 1, 3, 5, 7 and 10. Data collection began in 2010 with assessments in grades 3 and 9, grade 1 assessments were added in 2014, grade 5 added in 2017 and grade 7 added in 2018. The entire population of students in compulsory schooling and in the respective grades take part in the standardised tests, excluding the grade 7 tests which are only administered to students with access to tablets (currently about half of grade 7 students) thus the sample of seventh graders is not a fair representation of the full population (Luxembourg Center for Educational Testing 2021). To date the sample size currently amounts to approximately 28,000 students. The content of the assessments includes a language portion (which differs in each learning cycle) and a mathematical abilities portion. In grade 1 (cycle 2.1) the language assessment includes competencies in Luxembourgish listening comprehension, in grade 3 (cycle 3.1) they include reading and listening comprehension in German and from cycle 4.1-9e the assessments additionally include reading comprehension in French. The school monitoring program also collects additional data on learning motivation and school/class climate which are reported back at the class and school levels, as well as family background data which is collected in order to ensure a fair comparison of scores. The ÉpStan data is available upon request from the Luxembourg Centre for Educational Testing (LUCET).

2.7 Norway

UngData

The **Municipal Youth Survey** (UngData known as YoungData in English) is a cross sectional survey on adolescent health and well-being conducted by the Welfare Research Institute NOVA at OsloMet. UngData began in 2010, and since then has collected information on the well-being and time use of over 665,000 students. The surveys are completed entirely online within the school setting, and are targeted towards students in the lower and upper secondary levels with an adjusted version of the survey known as UngData Junior introduced in 2017 for students in grades 5 to 7. The questionnaires include themes such as relationships with parents and peers, school, local environment, health, leisure activities, well-being and at the secondary level there are additional themes of risky behaviour (such as alcohol and drug use), violence and sexuality (Frøyland 2017). Outside actors who wish to access the data should contact the Norwegian social science data service (Norsk samfunnsvitenskapelig datatjeneste).

Ung I Norge

The **Young in Norway Study (Ung I Norge)** is a longitudinal study following the life course from adolescence to adulthood of students in Norway by researchers from Norwegian Social Research (NOVA) at Oslo Metropolitan University. The first wave of the survey was administered in 1992 to students between the ages of thirteen and nineteen with over 12,000 students participating. To date there have been a total of 5 waves of data - in 1992, 1994, 1999, 2005 and 2020. The most recent wave

was carried out digitally and interviewed the now 40+ year old adults. Young in Norway examines the life course of those who participated including but not limited to their living conditions, interpersonal relationships, identity, mental health, education and leisure activities (Strand and von Soest 2008). An application to access the data can be made to OsloMet but access is limited and potentially difficult to obtain.

MoBa

The **Norwegian Mother, Father and Child Cohort Study (MoBa)** is a longitudinal study of the cause of disease in mothers and children by the Norwegian Institute of Public Health (NIPH). MoBa recruited over 90,000 pregnant women to participate in the study between 1998 and 2008 and after pregnancy the mother continued to provide information about the child at months 6 and 18, and at age 3, 5, 7, 8 and 14. The questionnaires administered after the birth of the child include information about the child's physical and mental health, language development, social development and about developmental disorders such as autism and ADHD (Norwegian Mother and Child Cohort Study 2020). An additional survey was administered to the children's teachers at the age of 8 and 11 for the Language and Learning Study (SOL) and currently this data may only be used in collaboration with SOL. The generic data is accessible, the conditions for application and pricelist can be found online (MoBa Website).

2.8 Spain

EGD

The **National Diagnostic Assessments (EDG)** make up two cross-sectional studies of the key competencies of primary and secondary students, and was carried out by the National Institute for Educational Evaluation (INEE) in collaboration with the Spanish autonomous communities. The fourth year primary students were assessed in 2009, and the second year secondary students in 2010. Some of the key competencies evaluated include, communication linguistics, mathematics, knowledge of the physical world and social-civic skills, following the conceptual framework of the European commission. The EGD data is made publically available in excel format (INEE Website) and the results of these assessments have been published by the Ministry of Education (Ministry of Education 2010, 2011).

PROA

The Evaluation of PROA Plan (Programmes of Reinforcement, Orientation and Support) is a cross-sectional study by the National Institute for Educational Evaluation (INEE) in collaboration with the Spanish autonomous communities. This project studies Spanish student performance in PISA 2012 in relation to their participation in PROA between 2005 and 2012. The dataset includes the matching of PISA data with information concerning student's past participation in the PROA program. Basic background and contextual information is also available. The data for this evaluation is publically available and a working paper containing the results of this study was published by the Ministry of Education (García-Pérez and Hidalgo 2014).

2.9 Switzerland

COCON

The Swiss Survey on Children and Youth (COCON) is a project aimed at examining the psychosocial development and life experiences of children and adolescents from a life course perspective. The study provides both cross-sectional data on specific topics as well as longitudinal data on children's trajectories and development. In addition to the main study, several intensive studies provide more in-depth information on specific groups. Its target population consists of children and adolescents in the French and German speaking regions of Switzerland, specifically three age cohorts: middle childhood (6 years old), middle adolescence (15 years old) and early adulthood (21 years old). The goal of COCON is to follow the first two cohorts until the eventual age of 21, with interviews every three years and additional information gathered before and after crucial transitions. COCON collects data on transition periods within the lives of students (including but not limited to transitions to high school, apprenticeship and employment) as well as differences in competence development as a result of differing home, school and peer group environments (Buchmann and Igel 2019). The survey additionally includes data on responsibility and decision making, focusing on values such as empathy, tolerance, solidarity, cooperation, respect, solidarity, and the moral power of judgement (Buchmann and Igel 2019). The data from each cohort is available on SWISSUbase and requires login for download, the datasets are available in German and the documentation is available in both French and German (Buchmann *et al.* 2013, 2018, 2019).

ÜKG/ COFO/ VeCoF

The **Verification of the Attainment of Basic Competencies (ÜKG/ COFO/ VeCoF)** are used to evaluate the attainment of competencies of students in mandatory schooling at three phases of the school cycle in Switzerland. There currently are two waves of data available, the 2016 wave which assessed the mathematical competencies of 11th grade students and the 2017 wave which assessed the language skills of 8th grade students. The 2017 wave focused both on the language of instruction including reading and writing skills and on foreign languages including oral and written comprehension skills. In each of the assessments over 20,000 students from all 26 cantons in Switzerland participated (Konsortium ÜGK 2019a, 2019b, Nidegger 2019, 2021). The students were asked additional background information about their social origin, home language, immigration status, home and school context, learning experiences and individual characteristics (Hupka-Brunner *et al.* 2016, Erzinger *et al.* 2019, Hascher *et al.* 2019, Pham, Helbling, Verner, and Ambrosetti 2019, Pham, Helbling, Verner, Petrucci, *et al.* 2019). The next waves of the competency tests will take place in 2023 assessing the language skills of 11th grade students and in 2024 assessing the language and mathematical skills of fourth grade students. The ÜKG/ COFO/ VeCoF data is available online on SWISSUbase.

TREE

The **Transitions from Education to Employment (TREE)** study is a longitudinal study that follows school leavers in Switzerland, collecting data on their educational and occupational paths after completing compulsory education. The first TREE cohort began in 2001 with a sample that consists of over 6000 school leavers who had participated in the PISA 2000 survey. The first cohort was surveyed yearly from 2001 to 2007, and then completed two additional surveys in the years 2010 and 2014. There are a total of 9 waves of data available from the first cohort with a tenth wave that is currently ongoing. The

second cohort of school leavers in Switzerland includes with over 8,000 individuals who had participated in the 2016 ÜGK/COFO/VeCoF study. This wave is commonly known as TREE2, and only the first three waves of data are currently available for this cohort (waves 0, 1 and 2). The TREE study focuses on educational and employment pathways, including data on attainment of first post-compulsory certificates, job search activities (duration and conditions), the presence of unemployment spells and job-skills mismatch (TREE 2016a, Hupka-Brunner *et al.* 2021). The data is available to all interested scholars on SWISSUbase (TREE 2019, 2021).

DAB

The **DAB panel study** (Determinanten der Ausbildungswahl und der Berufsbildungschancen) collects longitudinal data on the educational and vocational situation of adolescents in German-speaking Switzerland (DAB Panel Study (DAB): W1 - W8 2022). The target population includes all students in 8th grade of public schools in German-speaking cantons (or German-speaking parts of cantons) of the school year 2011/2012. At present, the sampled individuals have been surveyed eight times about their current training situation, their entry into the labour market, and their educational aspirations (DAB Panel Study 2021). Thus, DAB provides data from a life-course perspective, crosscutting various topics in the fields of educational, labour market and occupational research. The documentation and codebooks are available in German and English. The data is available for all interested researchers on SWISSUBase (DAB Website).

2.10 International Datasets

In this section we summarize and outline international datasets relevant for education research. The included datasets provide contents corresponding to the research questions and objectives of our empirical work program. We acknowledge that the following list of international datasets is not exhaustive.

PIRLS

The **Progress in International Reading and Literacy Study (PIRLS)** by the International Association for the Evaluation of Educational Achievement (IEA) is a cross sectional study that measures the reading achievements of fourth grade students internationally. It is globally known as a standard for assessing reading trends in the fourth grade. PIRLS began in 2001 and takes place every five years in around 50 countries worldwide. The study collects information on the student, parent, teacher and school levels including topics such as basic demographic information, home environment, school climate for learning, and attitudes toward reading (Gonzalez and Kennedy 2003, Foy and Kennedy 2008, Foy and Drucker 2013, Foy 2018). Additionally, the reading assessment portion of PIRLS focuses on four comprehension processes for reading: making straightforward inferences, focusing on and retrieving explicitly stated information, interpreting and integrating ideas and information, and evaluating and critiquing content and textual elements (Gonzalez and Kennedy 2003, Foy and Kennedy 2008, Foy and Drucker 2013, Foy 2018). PIRLS data and documentation is publicly available for download online from the PIRLS and TIMSS website (TIMSS&PIRLS Website).

TIMSS

The **Trends in International Mathematics and Science Study (TIMSS)** is a study by the IEA. The aim of TIMSS is to assess mathematics and science achievements of fourth and eighth grade students internationally. TIMSS began in 1995 and continues to measure math and science trends every 4 years, meaning the fourth grade cohort is reassessed in the eighth grade resulting in the quasi-longitudinal design. The assessments are created through collaboration of the participant countries and National Research Coordinators play an important role in the development of the assessment questions. Additionally, the school, teacher, student and home questionnaires collect information on school curriculum, the educational system, as well as background information on the student's home lives, school lives and attitudes towards learning (Martin 2005, Foy and Olson 2007, Foy *et al.* 2013, Foy 2017). The TIMSS data is made publicly available on the PIRLS and TIMSS website, where additional information can be found on the participating countries in each wave as well as the technical documents of the study (TIMSS&PIRLS Website).

PISA

The **Programme of International Student Assessment (PISA)** by the Organisation for Economic Co-operation and Development (OECD) is a repeated cross-sectional study that measures the reading, mathematics and science knowledge and skills of fifteen-year-old students internationally. The choice of fifteen-year-olds is deliberate since at 15 years old these students are nearing the end of their compulsory education in most OECD countries. PISA started in 2000, with a new wave of the study done every following three years. Since 2000, over 90 countries have been involved. PISA assessments

use the content of national curricula in order to look at student's abilities to analyse, solve problems and communicate effectively and background information is also collected about the students and the schools via principal and student level questionnaires (OECD 1999a, 2003, 2006, 2009a, 2013a, 2017a, 2019a). They include information about student attitudes, home environments, learning experiences, school management and school environment (OECD 2002, 2005, 2009b, 2012, 2014a, 2017b). Some countries include additional questionnaires at the home and teacher level. PISA data and documentation (including technical reports, assessment and analytical framework, and survey implementation tools) are publicly available on the OECD webpage (PISA Website).

TALIS

The **Teaching and Learning International Survey (TALIS)** by the OECD is a cross-sectional study that surveys teachers and school leaders about their schools' learning environment and teaching conditions. In the most recent wave (2018) of the survey, 48 countries participated with a sample size of over 260,000 teachers across 15,000 schools (OECD 2019b). The core population of the study is ISCED level 2 (lower secondary) school teachers, with the option for countries to additionally interview ISCED level 1 and 3 teachers. In the 2013 and 2018 waves, an additional option of the survey was available to countries who participated in the preceding PISA evaluation. For countries who opted for this option known as the PISA TALIS link, the TALIS survey was administered to the same schools as the PISA evaluations creating a link between the two surveys. In 2018 there were a total of 9 countries that opted for the PISA TALIS link. TALIS includes both a teacher and principal questionnaire and focuses on topics such as school climate, school leadership, teachers' instructional activities, teachers' professional practices, teachers' education and teachers' job satisfaction (OECD 2010a, 2014b, 2019c). The TALIS data is publicly available on the OECD website (TALIS Website).

PIAAC

The **Programme for the International Assessment of Adult Competencies (PIAAC)** is the OECD program designed to assess key information processing skills of adults. Specifically, PIAAC measures proficiency in literacy, numeracy and problem-solving skills amongst adults and gathers information on where and how adults use these skills. The target population of this study is 16 to 65-year-old adults. The PIAAC cycles take place every 10 years. In the first cycle there were a total of three rounds of data collection beginning in 2011 and ending 2018, the second cycle of PIAAC is currently in the preparation stages and results are expected to be published in 2024. Focusing on the first cycle of PIAAC, the study includes background information on demographic characteristics, linguistic status, employment status and income, it also includes a skills use module which looks into social skills, cognitive skills, physical skills and learning skills (OECD 2019d). The PIAAC data is available online in the form of public use files and can be downloaded from the OECD website (PIAAC Website).

EUROSTUDENT

EUROSTUDENT began in 1997 with seven rounds of the survey completed to date. EUROSTUDENT samples students attending higher education and focused on the socio-economic conditions of student life in Europe. It includes topics such as transitions within and into higher education, student housing, student expenses and resources, student characteristics and socio-economic background, student

employment, students' type of study and future aspirations and assessment (Cuppen, Muja, Hauschildt, Buck, *et al.* 2021, Cuppen, Muja, Hauschildt, Daniel, *et al.* 2021). The target population is all students who are at the time of observation (semester/term) enrolled in any national study programme regarded as higher education in the country (excluding doctoral programmes). The EUROSTUDENT macro data is available for download online from the EUROSTUDENT database and for the first time the micro data for round VII is available for more in-depth studies as of 2021. The EUROSTUDENT VII data is available upon request from FDZ-DZHW, a data use application form must be submitted in order to access the data (EUROSTUDENT VII Website).

AES

The **Adult Education Survey (AES)** by Eurostat collects data on adult participation in education and training and is one of the main EU data sources for lifelong learning statistics. The target population of AES is adults aged 25-64, although some national samples include data outside this age range. The AES includes information on adult education (formal, informal and non-formal), hours of instruction, costs and employer financing, reasons and obstacles for participating as well as background characteristics of the participant (Eurostat 2013, 2017). The first wave/the pilot of the survey was administered between 2005 and 2008, it covered 29 EU candidate countries or European free trade area countries. Since the first wave, a second wave was released in 2011, a third wave in 2016 and a fourth wave is set to be released in 2022-2023. In the most recent wave (2016) of the survey, there were 35 participating countries of which 33 countries have accessible data in the AES scientific use files. Access to the micro data is granted for scientific purposes online and in order to apply for access the organisation must be an accepted research entity by Eurostat and a research proposal must be submitted. For joint projects, a joint research proposal may be submitted and all project members must be recognized research entities. More information about accessability can be found online (Eurostat Website).

ISCWeB

The **International Survey of Children's Well-Being (ISCWeB)** is an international survey on children's subjective wellbeing supported by the Jacobs Foundation. The survey includes children aged 8, 10 and 12 and all information is collected from the child's perspective. There have been three waves of the survey to date, with the most recent wave covering 35 different countries. The main aim of the survey is to collect information on the daily lives of children, their time use, daily activities and perceptions of their own well-being. The children questionnaires focus on eight aspects of life including, the home and people the children live with, money and things the children have, relationships with friends and others, the area where they live, school, health and time management (Children's World 2013). The data is freely available upon request for academic research purposes and is released within 12-18 months after completion of the database. A data request form can be filled in on the Children's Worlds website (ISCWeB Website).

HBSC

The **Health Behaviour in School-Aged Children (HBSC)** survey is a cross national survey by the World Health Organization Regional Office of Europe. Beginning in 1984 with only five participating countries,

the survey now includes around 50 countries and regions in Europe and America. The survey is administered every 4 years to adolescents aged 11, 13 and 15 years old, focusing on their key health behaviours, well-being and social environments. The survey includes information on the following topic areas: body image, eating behaviours, health, life satisfaction, relationships (family and friends), bullying and fighting, school environment, sexual behaviour, substance use (alcohol, tobacco and cannabis) and socio-economic background (Health Behaviour in School-aged Children 2016a, 2016b, 2016c, 2018). The international data file has restricted access for HBSC member countries for the first three years and becomes open access thereafter. Data access enquiries can be sent to the HBSC Data Management Centre at the Department of Health Promotion and Development in the University of Bergen.

EU-SILC

The **European Union Statistics on Income and Living Conditions (EU-SILC)** by Eurostat is a two-part dataset, with cross sectional data covering the years 2004-2019 and longitudinal data covering the years 2005-2019. The EU-SILC contains data on all EU states as well as Switzerland, Norway, and Iceland. The target population includes all private households and individuals aged 16 and over. The data includes information on living conditions, income, poverty, social exclusion, housing, labour, education and health at both the household and individual levels (Eurostat 2020). Eurostat grants access to EU-SILC data for scientific purposes and an application for the micro data is required (EU-SILC Website).

LIS

The **Luxembourg Income Study (LIS)** is the largest harmonised income microdata database. It covers around 50 countries from Europe, America, Latin America, Asia, Africa, and Australasia. Beginning in 1980 the database currently has 11 waves of data. The harmonisation project is ongoing with many forthcoming datasets in the works. LIS includes household and individual level data on labour income, capital income, pensions, social benefits, private transfers, taxes, demography (immigration, health, and education), employment and expenditures (Luxembourg Income Study 2019). The data can be accessed by registered researchers in LISSY (a remote-execution system) for scholarly, research or educational purposes (LIS Website).

Table 1: Sample of the datasets in the inventory

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
INTERNATIONAL SURVEYS						
PIRLS: Programme of International Student Assessment	**Start:** 2001 **End**: Ongoing **Wave**: (every 5 years) **Size:** N = 346,852 (in 2016) N = 150,000 (in 2001) **Design:** Cross-section	50 countries in the last wave (2016).	Includes aspects of students' home and school lives, including basic demographic information, their home environment, school climate for learning, and attitudes toward reading. Assessment of reading achievement. **Level:** Primary	Students (fourth grade), parents, teacher, school **Language of documentation:** English	Publicly available on IEA website. https://timssandpirls.bc.edu/	Yes **Language:** English
TIMSS: Teaching and Learning International Survey	**Start:** 1995 **End:** Ongoing **Waves**: (every four years) **Size:** N = 724,356 (in 2019) Size depends on the country **Design:** Cross-section	64 countries in the last wave (2019).	Includes background questionnaires about the students' home and school lives, including basic demographic information, their home environment, school climate for learning, and attitudes toward learning mathematics and science. Assessments in mathematics and science achievement. **Level:** Primary, Secondary	Students (fourth and eighth grade), parents, teacher, school **Language of documentation:** English	Publicly available on IEA website. https://timssandpirls.bc.edu/	Yes **Language:** English
PISA: Programme of International Student Assessment	**Start:** 2000 **End:** Ongoing **Waves:** (every 3 years) **Size:** Depends on the technical standards that define the minimum number of sampled	79 countries in the last wave (2018). Participating countries by year: PISA participants	Includes background questionnaires on student's home and school environments, their learning experiences and attitude. Assessments in reading, mathematics and science knowledge and skill. **Level:** Secondary	Students (15 years old), school principals and parents **Language of documentation:** English	Publicly available on OECD website. https://www.oecd.org/pisa/data/	Yes **Language:** English

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
	schools and students provided by PISA. **Design:** Cross-section					
TALIS: Teaching and Learning International Survey	**Start**: 2008 **End:** Ongoing **Waves**: (every 5 years) **Size:** N = 260,000 (2018) N = 170,000 (2013) N = 70,000 (2008) **Design:** Cross sectional	48 countries in the last wave (2018).	Includes topics such as school climate, school leadership, teachers' instructional activities, teachers' professional practices, teachers' education and teachers' job satisfaction. **Level:** Primary and secondary (core focus on ISCED level 2, with additional option to survey ISCED levels 1 and 3)	Teachers and principals **Language of documentation:** English French	Publicly available on OECD website. https://www.oecd.org/education/talis/	Yes **Language:** English
EUROSTUDENT	**Start:** 1997 **End:** Ongoing **Waves:** 2019-VII (2016-2018)-VI (2012-2015)-V **Size:** N = 134,707 (VII) **Design:** Cross-section	13 countries in the last wave (2019).	Includes current study situation, study background, study conditions, living conditions, international mobility, personal situation, family background. **Level:** Tertiary	All students who are at the time of observation (semester/term) enrolled in any national study programme regarded as higher education in the country **Language of documentation:** English	Full international microdata available for wave VII. Requires a data usage application via FDZ-DZHW. https://metadata.fdz.dzhw.eu/en/start	Yes (Data and methods report) **Language:** English
AES: Adult Education Survey	**Start:** 2007 **End:** Ongoing **Waves:** (2007, 2011, 2016)	35 countries in the last wave (2016). Participating countries by	Includes participation in education and training (over the period of 12 months prior to the interview), training in informal and non-formal education, characteristics of the learning, hours spent, reasons and obstacles for participating, costs and employer	Adults aged 25-64 **Language of documentation:** English (questionnaire	Two step application required for data access, 1. Work for a recognised	Yes (quality reports for national participants) **Language:**

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
	Size: N = 183, 000 (in 2007) N = 225, 000 (in 2011) N = 240, 000 (in 2016) **Design:** Cross-section	wave: AES participants	financing as well as background information on the participant. **Level**: Tertiary	language depends on country) Documentation can be found here AES documentation	research entity by Eurostat Recognised research entities 2. Submit research proposal via Eurostat https://webgate.ec.europa.eu/multisite/microdata/	English
PIAAC: Programme for the International Assessment of Adult Competencies	**Start:** 2011-2012 (1st cycle) **End:** 2017 (1st cycle) Ongoing (2nd cycle) **Waves**: (every 3 years) **Size:** N = 166,000 (round 1) N = 50,250 (round 2) N = 35, 100 (round 3) **Design:** Cross-section	Round 1: 24 countries Round 2: 9 countries Round 3: 6 countries See participating countries at PIAAC participants	Includes background information on gender, migration, job status, highest education. Assessment of skills in three domains: literacy, numeracy and problem solving. **Level:** Tertiary	Adults aged 16-65 **Language of documentation:** English	Publicly available https://webfs.oecd.org/piaac/puf-data/	Yes **Language:** English
ISCWeB: International Survey of Children's Well-Being	**Start**: 2011-2012 **End:** Ongoing **Waves:** 3 Waves **Size:** N = 34,000 (2011-2012) N = 61,000 (2013-2014)	35 countries in the last wave (2016-2019).	Includes information on the daily lives of children, their time use, daily activities and perceptions of their own well-being. **Level:** Primary, Secondary	Children at 8 years old, 10 years old and 12 years old. **Language of documentation:** English Other languages available upon	All 3 waves of data are freely available upon request. Data requests forms can be completed via	No **Language:**

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
	N = 128,000 (2016-2019) **Design:** Cross sectional			request (for questionnaires)	https://isciweb.org/the-data/access-our-dataset/	
HBSC: Health Behaviour in School-Aged Children Survey	**Start**: 1984 **End:** Ongoing **Waves:** (every 4 years) **Size:** N = 219,810 (in 2013, the most recent open access data file) **Design:** Cross sectional	50 countries and regions across Europe and America.	Includes information on the following topic areas: body image, eating behaviours, health, life satisfaction, relationships (family and friends), bullying and fighting, school environment, sexual behaviour, substance use (alcohol, tobacco and cannabis) and socio-economic background. **Level:** Secondary	Children aged 11, 13 and 15yo, **Language of documentation:** English	Data access enquiries can be sent to the HBSC Data Management Centre. The international data file becomes open access three years from its completion. https://www.uib.no/en/hbscdata	No **Language:** N/A
EU-SILC: European Union Statistics on Income and Living Conditions	**Start**: 2004 **End:** Ongoing **Waves:** Yearly (individuals are observed periodically over 4 years for longitudinal data) **Size:** N = 292 150 (cross sectional) N =218 300 (longitudinal)	All EU states plus Iceland, Norway and Switzerland.	The data includes information on income, poverty, social exclusion, education, health, labour and other living conditions. **Level:** Tertiary	Households and individuals ages 16 years and over. **Language of documentation:** English	Two step application required for data access, 1. Work for a recognised research entity by Eurostat Recognised research entities 2. Submit research proposal via Eurostat https://webgate.e	Yes (quality reports for EU and member states) **Language:** English

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
	Design: Cross sectional and longitudinal				c.europa.eu/multisite/microdata/	
LIS: Luxembourg Income Study	**Start**: 1980 **End:** Ongoing **Waves:** 11 waves (every 3 to 5 years) **Size:** Depends on the country. **Design:** Cross national harmonised dataset	50 countries in Europe, America, Latin America, Africa, Asia and Australasia.	Includes topics such as labour income, capital income, pensions, social benefits, private transfers, taxes, demography (immigration, health, and education), employment and expenditures. **Level:** Tertiary	Households and individuals. **Language of documentation:** English	Data is available only to registered researchers via LISSY (a remote-execution system) https://www.lisdatacenter.org/data-access/lissy/eligibility/	No **Language:**
NATIONAL SURVEYS						
Kappas!: The Assessment of Undergraduate Students' Learning Outcomes in Finland	**Start**: 2019 **End:** 2020 **Waves:** 1 **Size:** N = 2,402 **Design:** Cross sectional	Finland	The main focus of this study is on generic skills and the development of said skills over the course of undergraduate studies. Generic skills included for assessment were, analytical reasoning and evaluation, problem solving, writing effectiveness and writing mechanics. **Level:** Tertiary	Undergraduate students currently attending a Finnish University or University of applied science. **Language of documentation:** Finnish	Not publicly available.	No (technical report) Yes (final report) **Language:** English Finnish
NEPS: The German National Educational Panel Study	**Start:** 2009-2012 **End:** Ongoing **Waves**: Yearly (follows 6 different cohorts) **Size:** N = 3,111 (SC1 Newborns)	Germany	Main topics of the survey include, developmental competencies, educational decisions, educational processes and returns in formal, non-formal and informal education. **Level:** ECEC, Primary, Secondary, Tertiary	Parents, Children, Childcare workers, Teachers (depending on the wave) **Language of documentation:** German English	Available for download from NEPS (must be registered). https://www.neps-data.de/Data-Center/Data-and-Documentation/N	Yes **Language:** English German

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
	N = 2,996 (SC2 Kindergarten) N = 6,112 (SC3 Grade 5) N = 16,425 (SC4 Grade 9) N = 17,909 (SC5 Students) N = 6,778 (SC6 Adults) **Design:** Longitudinal study				EPS-Data-Portfolio	
CILS4EU-DE: The Children of Immigrants Survey in Four European Countries - Germany	**Start:** 2010 (CILS4EU) 2014 (CILS4EU-DE) **End:** Ongoing **Waves**: Yearly, and biannually **Size:** N = approx. 3,000 **Design:** Longitudinal study	Germany	Includes topics on leisure behaviour, family and friendships, identity, feelings and beliefs, current situation (including completed certificates, class level, and grades), vocational training and employment. **Level:** Secondary and tertiary	Students Parents and teachers (only in wave 1 of CILS4EU)	Applications for both CILS4EU and CILS4EU-DE from GESIS Data Archive for Social Sciences. https://search.gesis.org/research_data/ZA6656	Yes **Language:** English
NABC: National Assessment of Basic Competencies	**Start**: 2006 **End:** Ongoing **Waves:** Yearly (follows student cohorts every 2 years) **Size:** N = ca. 100,000 **Design:** Longitudinal	Hungary	Includes institution and student background questionnaires. The main topics on the institution are school equipment, features and settlement. The student data includes information on the family's social, economic and cultural background. Assessments in mathematics and reading literacy. **Level:** Secondary (lower and upper)	Students in 6th, 8th and 10th grades and their schools (excludes schools that only educate SEN students) **Language of documentation:** Hungarian	Data requests can be made via email: adatkeres@krtk.hu **For full explanation see,** https://adatbank.krtk.mta.hu/en/adatbazisok/adatkeresek-menete/-	Yes **Language:** Hungarian English

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
HLCS: Hungarian Life Course Survey	**Start**: 2006 **End:** 2012 **Waves:** Yearly **Size:** N = 10,022 (initial sample) **Design:** Longitudinal	Hungary	Topics include, leisure time, housing, wealth, family environment, educational attainment, dropout rates and transitions to the job market. **Level:** Secondary and tertiary	Students, parents (depends on wave) **Language of documentation:** Hungarian English (translation)	Data is available upon request and can be requested from TARKI Social Science Data Archive.	Yes **Language:** Hungarian
GUI: Growing Up in Ireland	**Start**: 2006 **End:** Ongoing **Waves:** '08 Cohort: 9m, 3y, 5y, 7/8y and 9y. '98 Cohort: 9y, 13y and 17/18y **Size:** N = 8,000 (cohort '98) N = 10,000 (cohort '08) **Design:** Longitudinal study	Ireland	Main topics of the survey include, emotional development, child behaviour, health (pregnancy, medical care, diet and nutrition, exercise), educational environment, informal education, teacher-student relationship, child day care, mental development. **Level:** ECEC, Primary, Secondary	Parents, Carers, Children, Teachers and Principles (depending on the wave) **Language of documentation:** English	Application for data is required. There are two data files available, AMF (anonymized) RMF (research – only accessible within Ireland). Apply with ISSDA Data request form. https://www.ucd.ie/issda/data/guichild/	Yes **Language:** English https://www.growingup.ie/data-documentation/
NMPP: National Assessment of Student Achievement	**Start**: 2012 **End:** Ongoing **Waves:** Yearly **Size:** N = Unknown	Lithuania	Includes background questionnaires completed by the students as well as assessments in reading, writing and mathematics. **Level:** Primary and secondary	Students in grades 2, 4, 6 and 8 **Language of documentation:** Lithuanian Russian	Data request required via the National Education Agency.	No **Language:**

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
	Design: Cross sectional			Polish		
ÉpStan: Épreuves Standardisées	**Start**: 2010 (grades 3 and 10), 2014 (grade 1), 2015 (grade 5), 2018 (grade 7) **End:** Ongoing **Waves:** Yearly (follows student cohorts every 2 years) **Size:** N = ca. 28,000 **Design:** Longitudinal	Luxembourg	Includes key topics such as student's socio-economic and socio-cultural backgrounds as well as school and class climate, motivation to learn and teaching quality. Contents of the assessments differ depending on the cycle of the student. Cycle 2.1 includes listening comprehension (Luxembourgish) and math skills. Cycle 3.1 includes reading and listening comprehension (German) and math skills. From cycle 4.1-9e include reading comprehension (French and German) and math skills. **Level:** Primary and secondary	Students in grades 1, 3, 5, 7 and 9, Parents (of students in primary school). **Language of documentation:** German French	Applications for data can be sent to LUCET, who will determine eligibility.	No **Language:**
UngData / YoungData: The Munipical Youth Surveys	**Start**: 2012 **End:** Ongoing **Waves:** Yearly **Size:** N = 665,000 **Design:** Cross sectional	Norway	Includes themes such as relationships with parents and peers, school, local environment, health, leisure activities, well-being and at the secondary level there are additional themes of risky behavior (such as alcohol- and drug use) **Level:** Secondary and primary (in Ungdata junior)	Students aged 10-19 **Language of documentation:** Norwegian	Data is available upon request from National Database through a standard agreement with NOVA. https://www.ungdata.no/den-nasjonale-databasen/	Yes **Language:** Norwegian
Ung i Norge (Young in Norway)	**Start**: 1992 **End:** 2020 **Waves:** (1992, 1994, 1999, 2005, 2020)	Norway	Includes information about the life course of students such as their living conditions, interpersonal relationships, identity, mental health, education and leisure activities.	Students in junior high (in 1990s) until adulthood (40+ years old)	Applications for data can be made via OslMet, strict conditions apply and access may	No **Language:**

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
	Size: N = 12,0000 **Design:** Longitudinal		**Level**: Secondary and tertiary	**Language of documentation:** Norwegian	be difficult to gain.	
MoBa: Norwegian Mother, Father and Child Cohort Study	**Start**: 1998 **End:** Ongoing **Waves:** (pregnancy, 6 and 18 months, 3, 5, 7, 8, and 14 years) **Size:** N = 90,000 **Design:** Longitudinal	Norway	Includes information about pregnancy and the mother's health, as well as the child's physical and mental health, language development, social development and about developmental disorders such as autism and ADHD. **Level:** ECEC, primary and secondary	Mothers, Fathers and youth (one wave that also interviews the schools) **Language of documentation:** Norwegian English	Data is available. Specific price lists and accessibility can be found on the MoBa webpage. https://www.fhi.no/en/studies/moba/for-forskere-artikler/research-and-data-access/	No **Language:**
E.G.D. : Evaluaciones Generales de Diagnóstico	**Start:** 2009 (Primary), 2010 (Compulsory Secondary) **End:** 2009/2010 **Wave:** 1 **Size:** N = 28708 (2009) N = 29152 (2010) **Design:** Cross-sectional	Spain	Includes data on key competences (linguistic communication, mathematics, knowledge of the physical world, social-civic skills) and also on relevant sociodemographic and contextual variables such as sex, birth year, social status, migrant background, region, school resources, etc.. **Level:** Primary and Compulsory Secondary	Students at year 4 of primary education, teachers, school heads and families (2009) Students at year 2 of compulsory secondary education, teachers and school heads (2010) **Language of documentation:** Spanish	Data is accessible in Excel formats from the website for the Spanish Ministry of Education https://www.educacionyfp.gob.es/inee/bases-datos/evaluaciones-nacionales.html	Yes **Language:** Spanish
Evaluación del Plan PROA (Programas de	**Start:** 2005 **End:** 2012	Spain	Includes data on students' results in the PISA 2012 test and relates them to the participation in PROA Plan between 2005 and	Students who participated in the	Data is accessible in Excel (not working), SPSS	Yes **Language:**

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
Refuerzo, Orientación y Apoyo)	**Wave:** 1 **Size:** N = 25313 **Design:** Cross-sectional but adding retrospective variables about participation in PROA.		2012. Sociodemographic and contextual variables are also available. **Level:** Compulsory Secondary	PISA 2012 test in Spain. **Language of documentation:** Spanish and most variables (from PISA) in English	and STATA formats from the website for the Spanish Ministry of Education https://www.educacionyfp.gob.es/inee/bases-datos/evaluaciones-nacionales.html	Spanish
COCON: The Swiss Survey on Children and Youth	**Start**: 2006 **End:** Ongoing **Waves:** (every 3 years) until age of 21 **Size:** N = 1,273 (Cohort 1: 6yo) N = 1,258 (Cohort 2: 15yo) N = 584 (Cohort 3: 21yo= **Design:** Cross sectional and Longitudinal	Switzerland (French and German speaking areas)	Includes data on competence development, transition periods, family environments, school environments and peer groups. As well as the acquisition of values (such a responsibility, empathy, tolerance, cooperation, respect and judgement). **Level:** Primary, secondary and tertiary	Students (ages 6 to 21), principal caregivers and principal teachers **Language of documentation:** German French Dataset in German	Data is available on SWISSUbase and requires login for download. https://www.swissubase.ch/en/catalogue/studies/9202/14066/overview	No **Language:**
ÜKG/ COFO/ VeCoF: Verification of the Attainment of Basic Competencies	**Start**: 2016 **End:** Ongoing **Waves:** Irregular (2016, 2017), forthcoming (2023, 2024) **Size:**	Switzerland	Includes background information on students' social status, home language, immigration status, home and school context, learning experiences, and individual characteristics. Assessments in mathematics and language skills.	Students at three defined phases of the school cycle, namely the 4th (forthcoming), the 8th, and the 11th grade	Available on SWISSUbase and requires login for download. (2017) https://www.swissubase.ch/en/cat	Yes **Language:** English

Dataset	Data Collection	Countries	Contents	Respondents	Availability of microdata	Technical Report
	N = 20,177 (2017) N = 22,423 (2016) **Design:** Cross sectional		**Level:** Primary and secondary	**Language of documentation:** English	alogue/studies/13413/15817/overview (2016) https://www.swissubase.ch/en/catalogue/studies/13413/15817/overview	
TREE: Transitions from Education to Employment	**Start**: 2000 **End:** Ongoing **Waves:** Yearly **Size:** N = 6,343 (cohort 1) N = 8429 (cohort 2) **Design:** Longitudinal	Switzerland	Includes data on attainment of first post-compulsory certificates, job search activities (duration and conditions), the presence of unemployment spells and job-skills mismatch. **Level:** Secondary and tertiary	Individuals (15 years and older) **Language of documentation:** English French German	Data is available on SWISSUbase and requires login for download. https://www.swissubase.ch/en/catalogue/studies/12476/18017/overview	Yes **Language:** English French German
DAB: Determinanten der Ausbildungswahl und Berufsbildungschancen (Determinants of educational choices and chances in vocational education)	**Start:** 2012 **End:** Ongoing **Waves:** Irregular **Size:** N= 3,680 **Design:** Longitudinal	Switzerland	Includes data on educational transitions to and within upper secondary education, vocational training, post-compulsory education, and labour market entry. **Level:** Secondary and post-secondary	Individuals (starting from 8th grade) **Language of documentation:** English German	Data is available on SWISSUbase and requires login for download. https://www.swissubase.ch/en/catalogue/studies/10773/14065/overview	Yes **Language:** English German

3 Harmonisation Guidelines for the Key Variables

In this section, we shift our focus to harmonising the key variables to be used in the quantitative research of PIONEERED. The variables are harmonised across two dimensions: **by conceptual content** (based on variable concepts/definitions that are as similar as possible across all datasets) and **by coding structure** (meaning continuous standardized variables use the same units of measurement across all datasets and categorical standardized variables report variables using the same value codes/labels). We adapt a **variable-by-variable approach**, harmonising the key socio-demographic variables that contribute to education inequality in Europe, as well as the educational outcome variables.

Only a fraction of datasets are included in the harmonisation inventory: PIRLS[4], TIMSS[5], PISA[6], EUROSTUDENT VII[7], PIAAC[8], LIS[9], Growing up in Ireland[10], NEPS SC4[11] and TREE2[12]. These datasets were chosen based on the research needs of the project partners. Two additional datasets that are also selected for the empirical work program, namely the Lithuanian assessment data and the Hungarian administrative data, are not accompanied with the harmonisation guidelines. This is due to restricted access to data and/or lacking technical documentation online or in English. Additionally, the Luxembourgish assessment data ÉpStan is included in the variable definition sections but is excluded from the syntax due to restricted access to the technical documentation, such as codebooks, as well as the data files. For these datasets and any future datasets that might be included during the course of the project, the harmonisation guide will serve as a blueprint for the national teams. Using this blueprint, the teams have the tools to harmonise their country's data to be comparable with the datasets in the inventory.

Additionally, not every dataset is covered for every variable. Datasets were included in the variable subsections that are relevant to the research proposed. This approach was dictated by the short duration dedicated to the harmonisation exercise across numerous datasets and allowed us to have more flexibility in the harmonisation of variables that are specific to certain research questions.

Variable-by-variable harmonisation below follows the same structure: we briefly discuss the measurement of the target variable across different surveys, referring to existing harmonisation literature, whenever applicable. It is followed by a table with the original coding in each dataset – both cross-sectional and longitudinal. We proceed afterwards with the suggested harmonisation approach to ensure maximum comparability across these datasets. The Stata syntax file can be provided upon request. With respect to the missing values, the harmonisation guidelines ensure that the original missing value schemes remain as is.

[4] PIRLS Codebooks: (International Association for the Evaluation of Educational Achievement (IEA) 2003a, 2007a, 2013a, 2019)
[5] TIMSS Codebooks: (International Association for the Evaluation of Educational Achievement (IEA) 2005a, 2009a, 2013b, 2013c, 2017a, 2017b, 2021a, 2021b)
[6] PISA Codebooks: (OECD 2001a, 2001b, 2004a, 2004b, 2007a, 2007b, 2010b, 2010c, 2013b, 2013c, 2016, 2019e)
[7] EUROSTUDENT VII Data: (Cuppen, Muja, Hauschildt, Daniel, *et al.* 2021)
[8] PIAAC Codebook: (OECD n.d.)
[9] LIS Codebook: (Luxembourg Income Study (LIS) 2019)
[10] Growing Up in Ireland Codebooks: (Growing Up in Ireland n.d., n.d., n.d., n.d.)
[11] NEPS Starting Cohort 4 Codebook: (National Education Panel Study (NEPS) 2020)
[12] TREE 2 Codebooks: (TREE 2016b, 2017, 2018)

3.1 Socio-Demographic Variables (Axes of Inequality in Education)

3.1.1 Sex / Gender

Sex/Gender is a valuable variable in educational inequality research and serves as an important role in identifying disadvantaged groups. All the surveys in the inventory report the sex of each student. The harmonisation is relatively straightforward given the majority of the surveys report sex based on biological and physiological characteristics, meaning two binary categories of male and female. The only exception lies in the EUROSTUDENT VII survey, which asks multiple questions about the student's gender identity (see *Table 2* and *Table 3*).

Some variation in the reporting of sex could exist between surveys that allow self-completion vs. surveys that instruct the interviewer to provide the answer without asking the interviewee. An interviewer will most likely report the respondent's biological sex, where the respondent may have chosen to respond with their gender identity. Another source of variation is amongst surveys that report non-binary as a third option for gender. In our inventory, this is true for the EUROSTUDENT VII survey (only in the Finnish sample). The inclusion of a non-binary gender identification is optimal for analysing transsexuality and intersexuality, although it poses some complications for harmonising with surveys that omit it. Overall, the surveys included allow for a straightforward harmonisation of sex.

Table 2: Cross sectional datasets

Dataset	Year	Variable label/definition *(Value scheme)*
PIRLS	2016, 2011, 2006, 2003	Sex/Gender *(1: Girl; 2: Boy)*
TIMSS G4	2019, 2015, 2011, 2007, 2003	Sex *(1: Girl; 2: Boy)*
TIMSS G8	2019, 2015, 2011, 2007, 2003	Sex *(1: Girl; 2: Boy)*
PISA	2018, 2015, 2012, 2009, 2006, 2003, 2000	Gender *(1: Female; 2: Male)*
EUROSTUDENT	2019	-What is your sex? (1: Female; 2: Male; 3: Other; 4: I prefer not to assign) -Sex used to enroll at higher education institution *(1: Male; 2: Female; 3: Other)*
PIAAC	All	Is respondent male or female? *(1: Male; 2: Female)*
LIS	All	Sex *(1: Male; 2: Female)*

Table 3: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
GUI '98	All	Gender study child (*1: Male; 2: Female)*
NEPS SC4	All	Gender *(1: Male; 2: Female)*
TREE2	All	Gender (*1 female; 2 male)*
ÉpStan Grade 3	2013	Gender *(1: Female; 2: Male)*
ÉpStan Grade 9	2019	Gender *(1: Male; 2: Female)*

We propose the following harmonisation for sex: *Sex (1: Female; 2: Male)*.

Overall the harmonisation of the gender/sex variable is straightforward and we follow the standardization as proposed by the European Commission (2020) where sex "refers to the biological and physiological characteristics that define a person to be either male or female" (p. 5). Similarly, Hoffmeyer-Zlotnik (2016) standardises sex into two categories where trans-sexuality and inter-sexuality are not included. Using this standardized variable allows for a comprehensive harmonisation across all surveys and facilitates comparisons across studies.

EUROSTUDENT VII is the only survey that deviates from the desired harmonisation. It begins with a more open question on gender, allowing a third and fourth response option where students can choose not to assign themselves a gender category. It continues by asking the students who originally did not identify as either male or female to specify the gender with which they are registered at their higher education institution. All EUROSTUDENT VII participating countries follow this above mentioned two-question sequence; with the exception of Estonia, Georgia, Hungary and Poland, which omit the non-binary response options altogether and thus, ask only the first question. In the countries that use the two-question sequence, combining these two survey questions generates a sex variable comparable with the other surveys with only two binary response options. This is true for all countries except Finland, which allows students to identify in an "other" category at their higher institutions, resulting in a third non-binary category for gender remaining in the sample of Finnish students (Cuppen, Muja, Hauschildt, Buck, *et al.* 2021).

For the Finnish sample in the EUROSTUDENT data there exists the possibility to observe the effects of non-binary identified individuals. This is evidently an advantage for this sample, as often data availability hinders researcher's ability to observe this particular group or participants. However, for the purpose of the harmonisation of sex/gender and comparability across all the surveys we propose a target variable that recodes as missing the non-binary identifying individuals in the Finnish sample which make up only 40 of the entire sample (N=134, 255). All other surveys will be coded to match the value scheme of the desired target variable.

Proposed harmonisation:

Operationalization	Available datasets
Sex (1: Female; 2: Male)	All

3.1.2 Country of Birth – Mother and Father

The parental country of birth variable is an important migration indicator as it allows researchers to identify individuals with migratory background. All the surveys include information about the parent's country of birth, although they differ significantly in detail. While some surveys report exact countries of birth, others provide multiple categories with the options for the most common countries of birth amongst their residents and others report a simple yes/no variable for those born in the country of the survey. Generally, the students report this information, with the exception of GUI, where the parents respond to this question. Below in *Table 4**: Cross sectional datasets*** and *Table 5**: Longitudinal datasets*** is a detailed description of the variables in the inventory.

Table 4: Cross sectional datasets

Dataset	Year	Variable label/definition *(Value scheme)*
PIRLS	2016, 2011	N/A
PIRLS	2006, 2001	-Was your mother born in <country>? -Was your father born in <country>? *(1: Yes; 2: No; 3: I don't know)*
TIMSS G4 and G8	2019	-Was your <parent/guardian A> born in <country>? -Was your <parent/guardian B> born in <country>? *(1: Yes; 2: No; 3: I don't know; 4: Not Applicable)*
TIMSS G4 and G8	2015	-Was your mother born in <country>? -Was your father born in <country>? *(1: Yes; 2: No; 3: I don't know)*
TIMSS G4	2011	N/A
TIMSS G8	2011	-Was your mother born in <country>? -Was your father born in <country>? *(1: Yes; 2: No)*
TIMSS G4 and G8	2007, 2003	-Was your mother born in <country>? -Was your father born in <country>? *(1: Yes; 2: No)*
PISA	2018-2000	-In what country were your parents born? Mother -In what country were your parents born? Father *(1: Country of test; 2: Other country)*
EUROSTUDENT	2019	-In which country were your parents born? Mother or female guardian -In which country were your parents born? Father or male guardian *(0: Country of questionnaire; 1: Europe- Nordic and Baltic countries (DK, FI, IS, NO, SE, EE, LV, LT); 2: Europe- British Isles and Western countries ((AT, BE, FR, DE, LI, LU, NL, CH, UK; 3: Europe- Southern countries (CY, GR, IT, ES, PT, MT, AD, MC, VT, SM); 4: Europe- Central and Eastern countries (SI, HR, BG, CZ, HU, PL, RO, SK); 5: Non-EU/EFTA, EHEA (KZ, AM, AZ, GE, UA, BY, MD, RU, TR, MK, AL, BA, ME, RS, XK); 6: North America (incl. GL); 7: Latin America and Caribbean; 8: Asia; 9: Australis and Oceania + Antarctica; 10: Africa; 11: Other territories (predominately small islands))*
PIAAC		-Was your mother or female guardian born in <country name>? -Was your father or male guardian born in <country name>? *(1: Yes; 2: No)*

Table 5: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
GUI '98	(waves 1, 2, 3)	-Were you born in Ireland? Mother/Primary caregiver -Were you born in Ireland? Father/Secondary caregiver *(1: Yes; 2: No)*
NEPS SC4	(waves 1, 2, 3)	-Mother: Country of birth -Father: Country of birth *(0: Germany; 1: Not Germany)*
NEPS SC4	(wave 9)	-Country of birth, mother of respondent -Country of birth, father of respondent *(0: Germany; 1: Not Germany)*
TREE2	(wave 0)	-Mother born abroad -Father born abroad *(0: Born in Switzerland; 1: Born abroad)*
ÉpStan Grade 3	2013	Country of birth *(1: Luxembourg; 2: Portugal; 3: Italy; 4: Spain; 5: Belgium; 6: Germany; 7: France; 10: Bosnia, Serbia, Montenegro, Macedonia, Bosnia-Herzegovina, Albania, Croatia; 9: Cap Verde; 10: other country)*
ÉpStan Grade 9	2019	Country of birth *(1: Luxembourg; 2: Portugal; 3: Italy; 4: Spain; 5: Belgium; 6: Germany; 7: France; 8: Croatia; 9: Other EU country; 10: Bosnia; 11: Serbia; 12: Montenegro; 13: Macedonia; 14: Bosnia and Herzegovina; 15: Albania; 16: Cap Verde; 17: other country; 18: unknown)*

We propose the following harmonisation: *Country of birth – Mother/Father (1: Country of test; 2: Other country).*

Although some surveys offer more detailed responses, in order to remain comparable with all the surveys we propose a two-choice categorical variable that identifies those who are native born to the country of the test and those who are not.

In PIRLS 2001-2006 and TIMSS 2015-2019, the students report their parents' country of birth. These surveys allow the students to select the response option "I don't know" when responding to the question about their parents' birth country. There are a significant number of students that report "I don't know", for mother's country of birth between 7,616 and 28,365 students report they do not know the country, and for fathers between 9,724 and 31,168 students report they do not know (International Association for the Evaluation of Educational Achievement (IEA) 2003b, 2007b, 2017c, 2021c). To avoid a large sample reduction by coding these values as missing we will keep them as is, coding them as 3: I don't know. For any other response options, which indicate that the birth country is not applicable, the response is coded as missing.

Proposed harmonisation:

Operationalization	Available datasets
Country of birth – Mother Country of birth – Father (1: Country of test; 2: Other country)	PIRLS, TIMSS, PISA, EUROSTUDENT, PIAAC, GUI, NEPS, TREE2

3.1.3 Country of Birth –Self

Country of birth of the student or participant allows first generation migrants to be identified and is a key indicator in education inequalities. Much like the parental country of birth, the student's country of birth varies in details amongst the surveys. A detailed description of the variables in the inventory follows in *Table 6* and *Table 7*.

Table 6: Cross sectional datasets

Dataset	Year	Variable label/definition *(Value scheme)*
PIRLS	2016	Was your child born in <country of test>? *(1: Yes; 2: No)*
PIRLS	2011	N/A
PIRLS	2006, 2001	Were you born in <country>? *(1: Yes; 2: No)*
TIMSS G4	2019, 2015, 2007, 2003	Were you born in <country>? *(1: Yes; 2: No)*
TIMSS G4	2011	N/A
TIMSS G8	2019-2003	Were you born in <country>? *(1: Yes; 2: No)*
PISA	2018-2000	In what country were you born? *(1: Country of test; 2: Other country)*
EUROSTUDENT	2019	In which country were you born? *(0: Country of questionnaire; 1: Europe- Nordic and Baltic countries (DK, FI, IS, NO, SE, EE, LV, LT); 2: Europe- British Isles and Western countries ((AT, BE, FR, DE, LI, LU, NL, CH, UK; 3: Europe- Southern countries (CY, GR, IT, ES, PT, MT, AD, MC, VT, SM); 4: Europe- Central and Eastern countries (SI, HR, BG, CZ, HU, PL, RO, SK); 5: Non-EU/EFTA, EHEA (KZ, AM, AZ, GE, UA, BY, MD, RU, TR, MK, AL, BA, ME, RS, XK); 6: North America (incl. GL); 7: Latin America and Caribbean; 8: Asia; 9: Australis and Oceania + Antarctica; 10: Africa; 11: Other territories (predominately small islands))*
PIAAC		Were you born in <country name>? *(1: Yes; 2: No)*

Table 7: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
GUI '98	(wave 1)	Was study child born in Ireland? *(1: Yes; 2: No)*
NEPS SC4	(waves 1, 2, 3)	Country of birth, Germany or abroad? *(0: Germany; 1: Abroad)*
NEPS SC4	(wave 9)	Respondent born in Germany? *(1: in Germany; 2: Abroad)*
TREE2	(wave 0)	Respondent born abroad *(0: Born in Switzerland; 1: Born abroad)*
ÉpStan Grade 3	2013	Country of birth *(1: Luxembourg; 2: Portugal; 3: Italy; 4: Spain; 5: Belgium; 6: Germany; 7: France; 10: Bosnia, Serbia, Montenegro, Macedonia, Bosnia-Herzegovina, Albania, Croatia; 9: Cap Verde; 10: other country)*
ÉpStan Grade 9	2019	Country of birth *(1: Luxembourg; 2: Portugal; 3: Italy; 4: Spain; 5: Belgium; 6: Germany; 7: France; 8: Croatia; 9: Other EU country; 10: Bosnia; 11: Serbia; 12: Montenegro; 13: Macedonia; 14: Bosnia and Herzegovina; 15: Albania; 16: Cap Verde; 17: other country; 18: unknown)*

We propose the following harmonisation: *Country of birth –Self (1: Country of test; 2: Other county).*

This harmonisation is straightforward as none of the surveys deviate from this target variable.

Proposed harmonisation:

Operationalization	Available datasets
Country of birth –Self (1: Country of test; 2: Other county)	PIRLS, TIMSS, PISA, EUROSTUDENT, PIAAC, GUI, NEPS, TREE2

3.1.4 Age at Migration – Self

Age at migration focuses on the students or participants (of adult surveys) and is defined by the age at which the individual immigrated to the country of the survey. Information concerning the age at migration is available in most of the surveys but varies in detail. PISA, PIAAC, NEPS SC4 and TREE2 record the actual age in years old (or date of migration, which allows exact age to be calculated), where the remaining surveys report the age in pre-assigned categories. These differences in detail are outlined in *Table 8* and *Table 9*.

Table 8: Cross sectional datasets

Dataset	Year	Variable label/definition *(Value scheme)*
PIRLS	2016	How old was child when he/she came to <country of test>? *(1: Younger than 3 years old; 2: 3 to 5 years old; 3: 6 to 7 years old; 4: 8 years old or older)*
PIRLS	2011, 2006	N/A
PIRLS	2001	How old were you when you came to <country>? *(1: 10 years old or older; 2: 9yo; 3: 8yo; 4: 7yo; 5: 6yo; 6: 5yo; 7: 4yo; 8: 3yo; 9: 2yo; 10: 2 year old or younger)*
TIMSS G4	2019, 2015	How old was your child when he/she came to <country>? *(1: Younger than 3 years old; 2: 3 to 5 years old; 3: 6 to 7 years old; 4: 8 years old or older)*
TIMSS G4	2011	N/A
TIMSS G4	2007, 2003	How old were you when you came to <country>? *(1: Older than 5 years old; 2: 1 to 5 years old; 3: Younger than 1 year old)*
TIMSS G8	2019-2003	How old were you when you came to <country>? *(1: Older than 10 years old; 2: 5 to 10 years old; 3: Younger than 5 years old; 9: Omitted or invalid; 6: Logically not applicable; Sysmis: Not administered)*
PISA	2018, 2015	How old were you when you arrived in <country of test>? *(1: age 0 – 1; 2: age 1; 3: age 2; 4: age 3; 5: age 4; 6: age 5; 7: age 6; 8: age 7; 9: age 8; 10: age 9; 11: age 10; 12: age 11; 13: age 12; 14: age 13; 15: age 14; 16: age 15; 17: age 16)*
PISA	2012, 2009, 2006, 2003	How old were you when you arrived in <country of test>? *(VALUE: 0-16 years old)*
PISA	2000	N/A
EUROSTUDENT	2019	N/A
PIAAC	All	-At what age did you first immigrate to <country of test>? *(VALUE: age)* -Age of immigration, categories *(1: Aged 0-5; 2: Aged 6-10; 3: Aged 11-15; 4: Aged 16-20; 5: Aged 21-25; 6: Aged 26-30; 7: Aged 31-35; 8: Aged 36-40; 9: Aged 41 or older)*

Table 9: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (*Value scheme)*
GUI '98	(wave 1)	-How long ago did the Study Child first come to live in Ireland? *(1: 5 years or less; 2: 6-10 years ago; 3: 11-20 years ago; 4: More than 20 years ago)* -Age of person, Study child *(VALUE: age)*
NEPS SC4	(wave 1)	Age upon arrival in Germany *(VALUE: age)*
NEPS SC4	(wave 9)	-Date of respondent's move to Germany month *(VALUE: 1-12)* -Date of respondent's move to Germany year *(VALUE: 2011-2019)*

		-List of children/students: Date of birth month Target child *(VALUE: 1-12)* -List of children/students: Date of birth year Target child *(VALUE: 1991-1999)*
TREE2	(wave 0)	-How old were you when you arrived in Switzerland? *(VALUE: 0-16 years old)*
ÉpStan Grade 3	2013	Which grades did you attend in Luxembourg? -Précoce -Kindergarten 1st year -Kindergarten 2nd year -Grundschule 1st class -Grundschule 2nd class *(1: Attended Luxembourgish school; 2: Did not attend Luxembourgish school)*
ÉpStan Grade 9	2019	Which grades did you attend school in Luxembourg? -Précoce -Kindergarten, Cycle 1 -Grundschule, Cycle 2-4 -Sekundarshule *(1 : Entirely attended in Luxembourg; 2 : Entirely or partially attended abroad; 3 : Not at all attended)*

We propose the following harmonisation: *Age at migration (1: Younger than 5 years old; 2: Between 5 and 10 years old; 3: Older than 10 years old).*

We base our harmonisation on age categories in order to include the most surveys as possible. Some surveys deviate slightly from the proposed categories. PIRLS 2016, TIMSS Grade 4 2015-2019 and PIAAC include students at and below the age of 5 in category *1: Younger than 5 years old*, as their original categories do not allow for differentiation between 4 year-olds and 5 year-olds (International Association for the Evaluation of Educational Achievement (IEA) 2017c, 2018, 2021c, OECD n.d.).

Additionally, for PIRLS and TIMSS Grade 4 we assume the maximum age of migration of a child is 10 years old since the students participating in these surveys are in grade 4 and most likely are between the ages of 9 and 10. Thus, in PIRLS and TIMSS Grade 4 we omit the last category 3: Older than 10 years old.

In PISA and PIAAC the age at migration is a continuous variable measured in years, which was straightforward to harmonise. The NEPS harmonisation for age at migration was derived using the date of arrival in Germany and year of birth of the student. The difference between these two dates was used to deduce the age of the student when they moved to Germany.

The Growing up in Ireland ’98 variable is the only one not included in the harmonisation. The categories for how long ago a child moved to Ireland are informative but unfortunately, cannot precisely be converted into the age at migration. Changing the measurement into years old at migration creates a vague variable that is not comparable with the target variable. For this reason, there is no proposed target variable for this study.

The remainder of the surveys are harmonised into the target categories without any deviations.

Proposed harmonisation:

Operationalization	Available datasets
Age at migration –Self (1: Younger than 5 years old; 2: Between 5 and 10 years old; 3: Older than 10 years old)	PIRLS, TIMSS, PISA, PIAAC, NEPS, TREE2

3.1.5 Language Spoken at Home

The language spoken at home variable is often used as a proxy for migration status and is a valuable variable in education inequality research as it allows for students who do not speak the national language or main language(s) of instruction at home to be identified. Additionally, students of any age group can easily report the language spoken at home, compared to questions about migration status or parental country of origin, which younger students may report with less accuracy.

Language spoken at home varies in detail across the surveys in the inventory. While PISA, PIAAC, NEPS SC4, TREE2 and ÉpStan surveys record the exact language spoken at home, others record only whether or not the language of the test is most often spoken at home.

The optimal language at home variable records the exact language spoken using ISO 639.2 codes. This offers and extremely detailed identification of the precise language spoken in the home. Unfortunately, this level of detail is only available in a handful of the surveys in our inventory, including PIAAC, PISA and NEPS. Otherwise, the language spoken at home is aggregated and offers much less detail generally only with two response options that indicate the test language or another language is most often spoken at home. In TIMSS and PIRLS a slightly different definition of the variable is used. It records how often the language of the test is spoken at home with four response options. *Table 10* and *Table 11* outline these differences.

Table 10: Cross sectional datasets

Dataset	Year	Variable label/definition *(Value scheme)*
PIRLS	2016	How often do you speak language of test at home *(1: Always; 2: Almost always; 3: Sometimes; 4: Never)*
PIRLS	2011, 2006, 2001	How often do you speak language of test at home *(1: Always or almost always; 2: Sometimes; 3: Never)*
TIMSS G4	2019, 2015, 2007, 2003	How often do you speak language of test at home *(1: Always; 2: Almost always; 3: Sometimes; 4: Never)*
TIMSS G4	2011	How often do you speak language of test at home *(1: Always or almost always; 2: Sometimes; 3: Never)*
TIMSS G8	2019, 2015, 2011, 2007, 2003	How often do you speak language of test at home *(1: Always; 2: Almost always; 3: Sometimes; 4: Never)*
PISA	2018, 2015, 2012, 2009	-Language at home *(1: Language of test; 2: other)* -Language at home *(ISO 639.2 codes)*
PISA	2006	-Language at home *(1: language of test; 2: other national language; 3: other language)* -Language at home *(ISO 639.2 codes)*
PISA	2001, 2003	Language at home *(1: language of test; 2: other official language; 3: national dialects; 4: other languages)*
EUROSTUDENT	2019	Native language *(1: Yes, official language or main teaching language of country; 2: No, other language)*
PIAAC	All	Language most often spoken at home *(ISO 639.2 codes)*

Table 11: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (*Value scheme)*
GUI '98	(wave 1)	Language most often spoken at home *(1: English; 2: Irish; 3: other)*
GUI '98	(waves 3, 4)	Language spoken most often at home *(1: English; 2: Other)*

NEPS SC4	(waves 1, 2, 3, 5, 8, 11)	-Language most often spoken at home *(0: German; 1: not German)* -Language most often spoken at home *(ISO 639.2 codes)*
TREE2	(wave 0)	-Language at home *(1 Swiss German; 2 Liechtenstein dialect; 3 German; 4 French ; 5 Ticino dialect or dialect of an Italian region ; 6 Italian; 7 Romansch (Rhaeto-Romance) ; 8 Spanish; 9 Portuguese; 10 Turkish; 11 Southern Slavic language (Bosnian, Croatian, Macedonian, Serbian, Slovak); 12 Albanian; 13 English; 14 Other language)* -Test language is spoken at home *(0: Test language is not spoken at home; 1: Test language is spoken at home)*
ÉpStan Grade 3	2013	-Language spoken at home with Mother -Language spoken at home with Father *(1: Luxembourgish; 2: Portuguese; 3: French; 4: Italian; 5: Spanish; 6: German; 7: Bosnian, Croatian, Serbian, Serbo-Croatian, Montenegrin, Macedonian, Albanian; 8: English; 9: other; 10: answer not possible)*
ÉpStan Grade 9	2019	-Language spoken at home with Mother -Language spoken at home with Father *(1: Luxembourgish; 2: Portuguese; 3: French; 4: Italian; 5:Spanish; 6: German; 7: Bosnian; 8: Croatian; 9: Serbian; 10: Serbo-Croatian; 11: Montenegrin; 12: Macedonian; 13: Albanian; 14: English; 15: other; 16: N/A)*

For the harmonisation, we propose the use of ISO 639.2 codes when possible, but recognize that this will only be possible in a handful of surveys. Thus, for a comprehensive harmonisation that is comparable across all surveys, we will use the following definition for language at home: *Language at home (1: Language of test or national language; 2: Other).*

In order to compare, we recode the TIMSS and PIRLS variables into a new target variable to best match this definition. We will obtain the target variable using the reasonable assumption that the students responding always or almost always to how often they speak the language of the test at home, speak the language of the test most often at home, while those responding with sometimes or never would most often speak another language at home. For the remaining surveys anyone responding that they speak their country's national language at home, or they speak the language of the test at home, will be grouped together, while any other language most often spoken at home will be grouped into a second response option labelled "other".

In PIAAC, we recode the ISO-639.2 codes for language spoken into the target variable by matching the language spoken at home with the language of the survey. The language of the survey consists of two variables, one that describes the language of the background questionnaire, and the second that describes the language of the exercises. Individuals that speak at home the same language as either the background questionnaire and/or the exercises will be grouped together as speaking the language of the test at home.

Proposed harmonisation:

Operationalization	Available datasets
Language spoken at home (ISO codes)	PISA, PIAAC, NEPS
Language spoken at home (1: National language or language of test; 2: Other)	Above datasets and PIRLS, TIMSS, GUI, TREE2

3.1.6 Education –Mother and Father

As an important indicator of socio-economic status, parental education is a key variable in educational inequality research. This variable is a core social background variable in standardized surveys and is most commonly standardised internationally using the International Standard Classification of Education (ISCED). ISCED standardizations were introduced in 1997 and updated in 2011. Since then, most surveys pre-harmonise their survey questions either directly using the ISCED levels or such that they can be easily linked to the standardized ISCED levels. This eases the post-harmonisation efforts as most of the variables in our surveys have already been recorded with the ISCED levels in mind.

Some common errors in reporting the highest level of education to be aware of include the inadequate measure of the education of migrants or those who have completed their education abroad, especially amongst national surveys. Often national surveys use a fixed number of response options containing the qualifications most common for the country of the survey, but this limited detail in response options and yet increasing differences in educational systems internationally, complicate the measure of highest educational achievement for those holding degrees uncommon in their survey country (Schneider and Ortmanns 2019).

The vast majority of the surveys in our inventory record parental education levels using some categorization of the ISCED levels. In PIRLS, TIMSS grade 4, GUI, NEPS and ÉpStan, the parents themselves report their highest level of education, while in the other surveys it's the students who report on behalf of their parents. For full descriptions of the variables available, see *Table 12* and *Table 13*.

Table 12: Cross sectional datasets

Dataset	Year	Variable label/definition *(Value scheme)*
PIRLS	2016	- Highest education level, mother - Highest education level, father *(1: Did not go to school; 10: Not applicable; 2: Some <Primary education—ISCED Level 1 or Lower secondary education—ISCED Level 2>; 3: <Lower secondary education—ISCED Level 2>; 4: <Upper secondary education—ISCED Level 3>; 5: <Post-secondary, non-tertiary education— ISCED Level 4>; 6: <Short-cycle tertiary education—ISCED Level 5>; 7: <Bachelor´s or equivalent level—ISCED Level 6>; 8: <Master´s or equivalent level—ISCED Level 7>; 9: <Doctor or equivalent level—ISCED Level 8>)*
PIRLS	2011	- Highest education level, mother - Highest education level, father *(1: No school; 2: ISCED 1 or 2; 3: ISCED 2; 4: ISCED 3; 5: ISCED 4; 6: ISCED 5B; 7: ISCED 5A,1ST DEG; 8: Beyond ISCED 5A,1st Deg; 9: Not applicable)*
PIRLS	2006	-Mother level of education -Father level of education *(1: Some ISCED 1 or 2, did not go; 2: ISCED 2; 3: ISCED 3; 4: ISCED 4; 5: ISCED 5B; 6: ISCED 5A,1st degr; 7: Beyond ISCED 5A,1st degr>; 10: Not admin.)*
PIRLS	2000	- Highest education level, mother - Highest education level, father

		(1: Some ISCED 1 or 2 or did not go; 2: ISCED 2; 3: ISCED 3A or 3B; 4: ISCED 3C; 5: ISCED 4A; 6: ISCED 4B; 7: ISCED 5A or higher; 8: ISCED 5B or higher)
TIMSS G8 TIMSS G4	2019, 2015	- Level of education, <parent/guardian A> - Level of education, <parent/guardian B> *(1: Did not go to school; 2: Some <Primary education—ISCED Level 1 or Lower secondary education—ISCED Level 2>; 3: <Lower secondary education—ISCED Level 2>; 4: <Upper secondary education—ISCED Level 3>; 5: <Post-secondary, non-tertiary education—ISCED Level 4>; 6: <Short-cycle tertiary education—ISCED Level 5>; 7: <Bachelor's or equivalent level—ISCED Level 6>; 8: <Postgraduate degree: Master's—ISCED Level 7 or Doctor—ISCED Level 8>; 9: Not applicable)*
TIMSS G8 TIMSS G4	2011	- Level of education, <Step mother or female guardian> - Level of education, <Step father or male guardian> *(1: No school; 2: ISCED 1 OR 2; 3: ISCED 2; 4: ISCED 3; 5: ISCED 4; 6: ISCED 5B; 7: ISCED 5A, first degree; 8: Beyond ISCED 5A, first degree; 9: Not applicable)*
TIMSS G8	2007	- Highest education level, mother - Highest education level, father *(1; ISCED Level 1 or 2, or no school; 2: ISCED 2; 3: ISCED 3; 4: ISCED 4; 5: ISCED 5B; 6: ISCED 5A, first degree; 7: Beyond ISCED 5A, first degree; 8: I don't know)*
TIMSS G4	2007	N/A
TIMSS G8 TIMSS G4	2003	N/A
PISA	All	-Education level of Mother -Education level of Father *(0: None; 1: ISCED 1; 2: ISCED 2, 3:ISCED3B,C; 4: ISCED 3A-4; 5: ISCED 5B; 6: ISCED 5A, 6)*
EUROSTUDENT	2019	N/A
PIAAC	All	-Mother highest education level -Father highest education level *(1: ISCED 1-3C; 2: ISCED 3B-4; 3: ISCED 5-6)*

Table 13: Longitudinal datasets

Dataset	**Year (wave)**	**Variable label/definition *(Value scheme)***
GUI '98	(waves 1, 2, 3)	-Primary caregivers highest level of education -Secondary caregivers highest level of education *(1: None or primary; 2: Lower Sec; 3: Hi Sec/ TechVoc/ UppSec+Tech/ Voc; 4: Non Degree; 5: Primary; 6: Postgrad)*
GUI '98	(wave 4)	Primary caregivers highest level of education *(1: None or primary; 2: Lower Sec; 3: Hi Sec/ TechVoc/ UppSec+Tech/ Voc; 4: Non Degree; 5: Primary; 6: Postgrad)*
NEPS SC4	(waves 1, 5, 7)	-Highest educational qualification Respondent ISCED -Highest education qualification Partner ISCED *(0: 0A/1A no qualification; 1: [2B] school-leaving qualification of a Hauptschule, Volksschule, vocational preparation measure; 2: [2A] Mittlere Reife (intermediate school-leaving qualification), school-leaving qualification of a Realschule; 3:[3A] Fachhochschulreife, higher education entrance qualification; 4: [3B] apprenticeship,*

		vocational school, school of public health (less than two years), civil servant subclerical class, basic professional skills; 5: [3C] civil servant, clerical class; 6: [4A] Fachhochschulreife, higher education entrance qualification (second cycle); 7: [4B] apprenticeship, vocational school, school of public health (less than two years), civil servant subclerical class, basic professional skills (second cycle); 8: [5B] diploma from a Fachakademie and Berufsakademie, university of applied sciences for public administration, school of public health (at least two years), master craftsman/ craftswoman/technician certificate, other qualification from a Fachschule, civil servant executive class; 9: [5A] Bachelor, Master, Diploma, Magister, Staatsexamen, civil servants in administrative class; 10: [6] doctorate, habilitation)
TREE2	(wave 0)	-Mothers highest educational attainment -Fathers highest educational attainment *(0: Pre-primary education; 1: Primary education or first stage of basic education; 2: Lower secondary or second stage of basic education; 3: (Upper) secondary education; 4: Post-secondary non-tertiary education; 5: First stage of tertiary education; 6: Second stage of tertiary education)*
ÉpStan Grade 3	2013	-Mother highest education level -Father highest education level *(1:* No diploma; 2: Technical secondary education; 3: Secondary education; 4: Craftsman's diploma; 5: Non-university degree; 6: University degree*)*
ÉpStan Grade 9	2019	N/A

We propose the following harmonisation, *Mother/Father/Parent highest education level (1: None/ Primary/ Incomplete secondary; 2: Secondary/ Post-secondary non-tertiary (for some vocational); 3: Tertiary)*.

As can be concluded from the above two tables, most of the surveys have their own categories for parental education levels. Often surveys will combine multiple ISCED levels within one response option. For this reason, we propose a target variable of three general categories for parental education. We propose the following categorization, in ISCED-11 levels the three categories are defined as, Group 1: ISCED level 1 or 2, Group 2: ISCED 3 or 4, and Group 3: ISCED 5, 6, 7 or 8. The only deviation from these categories is in the PIAAC sample, which will include ISCED level 3C in Group 1, as it is not possible to separate this ISCED level in the original data.

In the longitudinal studies, we allow for changes in the parent's highest educational level over time if the survey reports highest education in more than one wave, this is the case for GUI and NEPS. Since these surveys do not report educational level in all waves of data, we propose two different target variables. The first, which records the changes in highest education only in the wave in which it is reported. The second, which fills in the missing waves with the most recent reported education level.

Proposed harmonisation:

Operationalization	Available datasets
Highest education level –Mother Highest education level –Father Highest educational level –Parents (1: None/ Primary/ Incomplete secondary; 2: Secondary/ Post-secondary non-tertiary (for some vocational); 3: Tertiary)	PIRLS, TIMSS, PISA, PIAAC, GUI 98', NEPS SC4, TREE2

3.1.7 Occupation – Mother and Father

Occupation itself is not an easy variable to define. Internationally the job market contains millions of different job titles where the duties and responsibilities are not fixed, and may differ between companies and nations. Harmonising occupation levels internationally is thus a very lengthy and time-consuming task. Thanks to the International Labour Organization (ILO) of the United Nations this task has become considerably easier with the creation of the International Standard Classification of Occupations (ISCO) in 1958. Surveys can thus be harmonised ex-ante, linking response options to the ISCO codes for easier harmonisation. This also simplifies the post-harmonisation required to make comparisons across different surveys or to compare internationally. These codes were updated in 2008, and ISCO-08 codes are still being used today. The codes are 4-digits long, linking jobs with the same set of main tasks and duties together (International Labour Organization 2012). We can further group the codes into smaller 3-digit, 2-digit and 1-digit groups based on these codes.

One source of variation in parental occupation may occur when the student's report parental occupations. This is the case in the PISA survey, which derives parental occupations from student reported answers on their parents' job titles and duties. Students are more prone to misreporting their parents' occupations due to a misunderstanding of the jobs of their parents and this would increase the risk of respondent error.

The level of detail recorded in occupation differs amongst the surveys in the inventory. NEPS, TREE2, PIAAC and PISA use the international standard classification of occupations (ISCO-08) 4-digit values. These values offer a very detailed occupation variable by grouping jobs based on the similarity of the task and duties undertaken (European Commission 2020). In the other surveys, categories of occupations are used instead of the in-depth codes for each occupation title. See *Table 14* and *Table 15* below for a detailed overview of the parental education variables available in the data.

Table 14: Cross sectional datasets

Dataset	Year	Variable label/definition *(Value scheme)*
PIRLS	2016, 2011, 2006, 2001	-Main job – Mother -Main job – Father *(1: Never worked; 2: Small Business Owner; 3: Clerical Worker; 4: Service or Sales Worker; 5: Skilled Agriculture or Fishery Worker; 6: Craft or Trade Worker; 7: Plant or Machine Operator; 8: General Laborers; 9: Manager or Senior Official 10: Professional; 11: Technician or Associate Professional; 12: Not applicable)*
TIMSS G4	2019, 2015, 2011	-Main job – Mother/Guardian B -Main job – Father/Guardian A *(1: Never worked; 2: Small Business Owner; 3: Clerical Worker; 4: Service or Sales Worker; 5: Skilled Agriculture or Fishery Worker; 6: Craft or Trade Worker; 7: Plant or Machine Operator; 8: General Laborers; 9: Manager or Senior Official 10: Professional; 11: Technician or Associate Professional; 12: Not applicable)*
TIMSS G4	2007, 2003	N/A
TIMSS G8	All	N/A
PISA	2018, 2015, 2012	-Mother ISCO code -Father ISCO code *(<ISCO-08> codes & 9701: Doing housework, bringing up children; 9702: Learning, studying; 9703: Retired, pensioner, on unemployment benefits; 9704: I do not know, it's hard to say;*

		9705: Answer too general or vague (e.g. good job, well paid job, employee, worker))
PISA	2009, 2006	-Mother ISCO code -Father ISCO code *(<ISCO-88> codes; 1251: Higher military officers (Captain and above); 1252: Lower grade commissioned officers (incl. Army Lieutenant); 6131: Mixed farmers; 6132: Farm foremen/supervisor; 6133: Farmers nfs ; 6134: Skilled farm workers nfs; 7500: Skilled workers NFS; 8400: Semi-skilled workers NFS (Incl. Production Process Worker nfs, Factory Worker nfs); 9501: Doing housework, bringing up children; 9502: Learning, studying; 9503: Retired, pensioner, on unemployment benefits; 9504: I do not know, it's hard to say; 9505: Answer too general or vague (e.g. good job, well paid job, employee, worker))*
PISA	2003	-Mother ISCO code -Father ISCO code (<ISEI-88> codes)
PISA	2000	N/A
EUROSTUDENT	2019	N/A
PIAAC	All	N/A
EU-SILC		-Mother occupation when respondent was 14 years old -Father occupation when respondent was 14 years old *(<ISCO 08>)*

Table 15: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	All	N/A in AMF
NEPS SC4	(waves 1, 5, 7)	-Mother: Profession -Father: Profession *(<ISCO-08>)*
TREE2	(wave 0)	-Mother's Occupation -Father's Occupation *(<ISCO-08>; -960: Response not in code list; -957:Houseman/wife; -950: No codable response)*
ÉpStan Grade 3	2013	-Mother's occupation -Father's Occupation *(1: Has never worked for pay; 2: Service of sales worker; 3: Academic of academic related occupation (professional); 4: Elementary occupation; 5: Craft or related trades worker; 6: Clerical support worker (clerk); 7: Technical or associate professionals; 8: Motorist, machine operator of assemblers; 9: Executive or senior official; 10: Skilled agriculture worker; 11: Answer not possible)*
ÉpStan Grade 9	2019	-Mother's occupation -Father's Occupation *(ISEI-08)*

We recommend the following harmonised target variables: *Mother/Father Occupation (1: Managers; 2: Professional; 3: Technicians and Associate Professionals; 4: Clerical support workers; 5: Services and Sales workers; 6: Skilled Agriculture, Forestry and Fishery workers; 7: Craft and related Trades workers;*

8: Plant and Machine Operators and Assemblers; 9: Elementary occupations; 10: Small Business Owners; 11: Never worked; 12: Other).

The ideal occupation variable is the detailed ISCO-08 codes. Given not all of the surveys offer this level of detail our proposed harmonisation will be occupation categories based on the ISCO-08 major groups and the categories presented in TIMSS and PIRLS. From the first digit of the ISCO-08 code the major groups are: (1) Managers, (2) Professionals, (3) Technicians and Associate Professionals, (4) Clerical support workers, (5) Services and Sales workers, (6) Skilled Agriculture, Forestry and Fishery workers, (7) Craft and related Trades workers, (8) Plant and Machine Operators and Assemblers, (9) Elementary occupations, and (0) Armed Forces (International Labour Organization 2012). The occupation categories used in both TIMSS and PIRLS can be found in the table above, and for the exception of two problematic groups, which we will discuss later, matching is a straightforward process.

In the original PIRLS and TIMSS definition of occupation, the "Small Business Owner" category poses some complications when linking to the ISCO-08 major groups. This is because ISCO-08 codes do not differentiate between working proprietors and workers; it focusses solely on the tasks and duties of the worker. This issue is further outlined in the International Standard of Occupations where it discusses how small business owners are to be assigned an ISCO-08 code, explaining that "in ISCO-08 such jobs are only classified in Major Group 1: Managers if there is evidence that management, including supervision of staff, is the major component of the work performed. For example, a motor vehicle mechanic who owns and operates his or her own workshop and employs a small number of other mechanics, a receptionist and a cleaner, but spends most of the time repairing vehicles and/or supervising the work of the other mechanics, is classified in Unit Group 7231: Motor Vehicle Mechanics and Repairers" (International Labour Organization 2012). Thus, determining where to group the Small Business Owners from PIRLS and TIMSS is problematic since we do not know the main tasks of said individuals. We will leave this group as is in the harmonised target variable, but will emphasise that it only exists in the PIRLS and TIMSS data and may require to be dropped or merged to another group for comparison with the surveys using ISCO-08 codes.

Additionally, PIRLS and TIMSS do not distinguish an Armed Forces group; instead, they group those who work for the armed forces in either the Manager or Services and Sales category. They do so by distinguishing military officers (also known as commissioned officers) as Managers and junior military (also known as non-commissioned or other rank officers) as Service and Sales Workers. We will thus omit the ISCO-08 Major Group 0: Armed Forces from the harmonised target variable and instead code the commissioned officers (with ISCO-08 codes 01) as 1: Managers and non-commissioned and other rank officers (ISCO-08 codes 02 and 03) as 5: Service and Sales Workers.

The PISA definition for parental occupations follows the ISCO-08 and ISCO-88 codes with the addition of some codes. In the years where ISCO-88 codes are used (PISA 2006 and 2009), we will first reassign the codes to their ISCO-08 counterparts and then proceed with the harmonisation. There are not any differences between the top level Major groups of ISCO-88 and ISCO-08 but some major group names have been slightly changed and some occupation categories have been moved from one major group to another (International Labour Organization 2012). For this reason, it is important to first convert the ISCO-88 codes to the ISCO-08 scheme and then proceed. In all the PISA studies there is an additional three pseudo-ISCO codes and two indistinguishable answer categories added to the occupation variable. These pseudo codes distinguish parents, whose work consists of, doing housework and bringing up children, learning/studying, and retired/pensioner/unemployment benefits (see *Table 14*,

codes 9701-9703 for PISA 2018-2012, or code 9501-9503 for PISA 2009-2006). In PISA 2018, in order to calculate parents' highest occupation status, these three pseudo codes were given an ISEI[13] value of 17, this is equivalent to the ISCO-08 Major Group 9: Elementary occupations (OECD 2019f). We however suggest a different harmonisation that compares better to the PIRLS and TIMSS categories for parental occupation and will include these three response options together in Group 12: Other. The two PISA indistinguishable categories "I do not know, it's hard to say" and "Answer too general or vague (e.g. good job, well paid job, employee, worker)", will be coded as missing.

In PISA 2006 and 2009, the parental occupation variable includes an additional eight occupation codes not included in ISCO-88 (see 1251, 1252, 6131-6134, 7500 and 8500 in *Table 14*). For these additional codes, we will exceptionally code them into the appropriate group when needed. Furthermore, in PISA 2003 occupation records ISEI codes and not ISCO codes, thus we will omit this dataset from the harmonisation since it is not possible to regroup the ISEI codes into the desired harmonised target variable.

In the TREE2 data, the ISCO-08 codes are used to classify parental occupation with the addition of some codes (see -960, -957 and -950 in *Table 15*). For the response options "Response not in code list" and "No codable response", we will recode them as missing. For the response "Houseman/wife", we will recode into Group 12: Other, in order to remain persistent with the other surveys.

Proposed harmonisation:

Operationalization	**Available datasets**
Mother's Occupation Father's Occupation (ISCO-08)	PISA, NEPS, TREE2
Mother's Occupation Father's Occupation (1: Managers; 2: Professional; 3: Technicians and Associate Professionals; 4: Clerical support workers; 5: Services and Sales workers; 6: Skilled Agriculture, Forestry and Fishery workers; 7: Craft and related Trades workers; 8: Plant and Machine Operators and Assemblers; 9: Elementary occupations 10: Small Business Owners; 11: Never Worked; 12: Other)	Above datasets and PIRLS, TIMSS

[13] International Socio-Economic Index of Occupational Status

3.1.8 Occupation – Self

We will define occupation of the participant/student using the International Standard Classification of Occupations 2008. Occupation of the students is not in every survey but is included in PIAAC, NEPS and TREE2, where the participants are old enough to be employed. Additionally, in EUROSTUDENT and Growing up in Ireland '98, information about the student's participation in paid work is included, but not the type of occupation they hold. Full details of the variables available is in the tables below.

Table 16: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
PIRLS	All	N/A
TIMSS	All	N/A
PISA	All	N/A
EUROSTUDENT	2019	Do you have paid jobs during your studies? *(1: Yes, I work during the whole lecture period; 2: Yes, I work from time to time during the lecture period; 3: No, I don't work during the lecture period)* Did you have any paid jobs during holiday or lecture free periods in the past 12 months? *(1: Yes; 2: No)*
PIAAC	All	-Current Job, respondent -Last Job, respondent *(ISCO-08, 1-4 digits)* - Occupational classification of respondent's job at 1-digit level, current job - Occupational classification of respondent's job at 1-digit level, last job *(Derived - ISCO 2008 (1 digit))*

Table 17: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	(wave 4)	- Have you ever been in paid employment - do not include the term-time employment or apprenticeship? *(1: Yes; 2: No)* - In relation to the current job you hold / last job you held, how would you describe it? *(1: Regular, full-time; 2: Temporary, full-time; 3: Regular, part-time; 4: Temporary, part-time; 5: Zero hour contract; 6: Internship)*
NEPS SC4	(wave 3)	- Student: Part-time job content *(ISCO-08)*
TREE2	(waves 1, 2)	- Occupation (ISCO-08) [job] *(ISCO-08)*
ÉpStan	All	N/A

We propose the following harmonisation: *Occupation-self (ISCO-08, 1st digit).*

Using the first digit of the ISCO-08 codes will allow us to incorporate all the information in the PIAAC survey. Although PIAAC records detailed 4-digit ISCO-08 codes for some respondents, for others it records the less detailed 3, 2 and 1-digit codes. Additionally, due to the lack of detail about the type of occupation in EUROSTUDENT and GUI '98, we will not include these datasets in the target harmonisation of occupation.

Proposed harmonisation:

Operationalization	Available datasets
Occupation –Self (ISCO-08, 1st digit)	PIAAC, NEPS, TREE

3.1.9 Household Income

Many surveys in our inventory report household income, but they all differ significantly in the unit of measurement and categorization of income groups. Some report net household income, some report gross household income, and others report equalized household income (which uses disposable income by household size). There is an enormous gap in the measurement, and thus harmonising this variable across all the surveys is not feasible.

The national longitudinal datasets and PISA do not provide sufficient information to form a harmonised unit of measure. LIS however uses many different measurements of household income and is the one exception for harmonisation possibilities.

Table 18: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
TIMSS G4	2019	N/A
PISA	2012	Annual household income, gross *(1: Less than < $A>; 2: < $A> or more but less than< $B>; 3: < $B> or more but less than< $C>; 4: < $C> or more but less than< $D>; 5: < $D> or more but less than< $E>; 6: < $E> or more)*
LIS	All	-Disposable household income *(VALUE: income)* -Total household income (*VALUE: income)* *Income is further broken down into many more categories in LIS data.

Table 19: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	All	Equivalised annual household income[14] *(VALUE: disposable household income divided by equivalised household size)*
NEPS SC4	(wave 1)	-Monthly household income, open *(VALUE: Income net amount in euros)* -Monthly household income, split *(1: less than 2,500 euros; 2: 2,500 euros and more)* -Monthly household income, categories under 2,500 euros *(1: less than 1,000 euros; 2: 1,000 to less than 1,500 euros; 3: 1,500 to less than 2,000 euros; 4: 2,000 to less than 2,500 euros)*

[14] In GUI equivalised income is calculated as follows, "In order to make meaningful comparisons between households on their income, household size and structure must be taken into account. This is done by creating an 'equivalised' income. In Growing Up in Ireland, an equivalence scale was used to assign a "weight" to each household member. The equivalence scales assigned a weight of 1 to the first adult in the household, 0.66 to each subsequent adult (aged 14+ years living in the household) and 0.33 to each child (aged less than 14 years). The sum of these weights in each household gives the household's equivalised size – the size of the household in adult equivalents. Disposable household income is recorded as total gross household income less statutory deductions of income tax and social insurance contributions. Household equivalised income is calculated as disposable household income divided by equivalised household size. This gives a measure of household disposable income which has been "equivalised" to account for the differences in size and composition of households in terms of the number of adults and/or children they contain" (Thornton *et al.* 2010).

		-Monthly household income, categories over 2,500 euros *(5: 2.500 to less than 3,000 euros; 6: 3,000 to less than 4,000 euros; 7: 4,000 to less than 5,000 euros; 8 5,000 euros and more)*
NEPS SC4	(waves 5, 7)	-Monthly household income[15], open *(VALUE: Income net amount in euros)* -Monthly household income, split *(1: less than 2,000 euros; 2: 2,000 to less than 4,000 euros; 3: 4,000 euros or more)* -Monthly household income, categories under 2,000 euros *(1: less than 1,000 euros ; 2: 1,000 to less than 1,500 euros; 3: 1,500 to less than 2,000 euros)* -Monthly household income, categories 2,000 - 4,000 euros *(4: 2,000 to less than 2,500 euros; 5: 2,500 to less than 3,000 euros; 6: 3,000 to less than 4,000 euros)* -Monthly household income, categories over 4,000 euros *(7: 4,000 to less than 5,000 euros; 8: 5,000 to less than 6,000 euros; 9: 6,000 euros or more)*
NEPS SC4	(waves 3, 5, 7, 8, 9,10, 11)	-Monthly household income, split *(1: less than 1,500 euros; 2: 1,500 to less than 3,000 euros; 3 3,000 euros and more)* -Monthly household income, categories under 1,500 euros *(1: less than 500 euros; 2: 500 to less than 1,000 euros; 3: 1,000 to less than 1,500 euros)* -Monthly household income, categories 1,500 - 3,000 euros *(1: 1,500 to less than 2,000 euros; 2: 2,000 to less than 2,500 euros; 3: 2,500 to less than 3,000 euros)* -Monthly household income, categories over 3,000 euros *(1: 3,000 to less than 4,000 euros; 2: 4,000 to less than 5,000 euros; 3: 5,000 euros and more)*
TREE2	(waves 1, 2)	-Salary: amount *(VALUE: open)* -Salary: unit *(1: Hourly pay ; 2: Monthly pay; 3: Other form of payment ; 7: Annual pay ; 9: Work without payment; 11: Payment in a foreign currency)* -Salary: type of foreign currency, if foreign currency *(Code list: currencies)* -Salary: unit/reference period of salary, if foreign currency *(1: Hourly pay ; 2: Monthly pay; 3: Other form of payment ; 7: Annual pay ; 13: Daily pay)* -Salary: gross or net salary *(1: Gross (prior to deductions for old age and survivors insurance, disability insurance and pension fund etc.); 2: Net (after deductions for old age and survivors insurance, disability insurance, and pension fund etc.))* -Salary: receives a 13th month of pay *(0: No; 1: Yes)* -Salary: 13th month of pay included in stated annual pay *(0: No; 1: Yes)*
ÉpStan	All	N/A

[15] NEPS household income is determined as follows, "What is the monthly household income of all household members? Please state the net amount, in other words the amount after taxes and social security contributions. Please include regular payments such as pensions, living compensation, child benefits, student financial aid (BAföG), subsistence allowance, unemployment benefits etc.!" (National Education Panel Study (NEPS) 2020).

For the household income variable, there is no reasonable harmonisation possible amongst the national longitudinal surveys and PISA. The reported incomes are either incompatible due to differences in measurement of income (equalized income, disposable income, net income, gross income), or the categorizations of the household incomes are incompatible. Harmonising this variable across all these surveys would result in an enormous loss of information and it is thus not recommended.

LIS provides sufficient information on household income to harmonise, but since it is the only survey remaining, there is no need for harmonisation.

Proposed harmonisation:

Operationalization	Available datasets
Due to heterogeneity across the surveys, the operationalization is to be completed based on specific research needs.	N/A

3.1.10 Number of Books at Home

The number of books in the home is another common socio-economic status indicator in educational surveys and data collection. Students of all age groups can record the approximate number of books in their household, and for this reason, the number of books in the household is available in all the surveys except EUROSTUDENT VII.

The main source of variation in the recorded variable is in the categorisation of the number of books. PIRLS and TIMSS use the same categories, as does PISA, PIAAC and ÉpStan, but the remaining national longitudinal surveys differ slightly in their choice of categories. Given the number of books variable reports the approximate amount of books in the house and not the exact count, we will harmonise the categories such that they are approximately similar.

Table 20: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
PIRLS	2016, 2011, 2006, 2001	Amount of books in your home *(1: None or very few (0–10 books); 2: Enough to fill one shelf (11–25 books); 3: Enough to fill one bookcase (26–100 books); 4: Enough to fill two bookcases (101–200 books); 5: Enough to fill three or more bookcases (more than 200))*
TIMSS G4 and G8	2019, 2015, 2011, 2003	Amount of books in your home *(1: None or very few (0–10 books); 2: Enough to fill one shelf (11–25 books); 3: Enough to fill one bookcase (26–100 books); 4: Enough to fill two bookcases (101–200 books); 5: Enough to fill three or more bookcases (more than 200))*
PISA	2018, 2015, 2012, 2009, 2006, 2003	How many books at home *(1: 0-10 books; 2: 11-25 books; 3: 26-100 books; 4: 101-200 books; 5: 201-500 books; 6: More than 500 books)*
PISA	2000	How many books at home *(1: None; 2: 0-10 books; 3: 11-25 books; 4: 26-100 books; 5: 101-200 books; 6: 201-500 books; 7: More than 500 books)*
EUROSTUDENT	2019	N/A
PIAAC	All	Number of books at home *(1: 0-10 books; 2: 11-25 books; 3: 26-100 books; 4: 101-200 books; 5: 201-500 books; 6: More than 500 books)*

Table 21: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
GUI '98	(wave 1)	How many children's books does child have access to in your home *(1: None; 2 Less than 10; 3: 10 to 20; 4: 21 to 30; 5: More than 30)*
	(wave 2)	How many books does child have access to in your home *(1: None; 2: One to 10; 3: 11 to 30; 4: 31 to 50; 5: 51 to 100; 6: More than 100)*
NEPS SC4	(waves 1, 9, 11)	Number of books *(1: 0-10 books; 2: 11-25 books; 3: 26-100 books; 4: 101-200 books; 5: 201-500 books; 6: More than 500 books)*
TREE2	(wave 0)	Number of books in home *(1: None; 2: 0-10 books; 3: 11-50 books; 4: 51-100 books; 5: 101-250 books; 6: 251-500 books; 7: More than 500 books)*
ÉpStan Grade 3	2013	N/A

ÉpStan Grade 9	2019	Number of books *(1: 0-10 books, 2: 11-25 books, 3: 26-100 books, 4: 101-200 books, 5: 201-500 books, 6: >500 books)*

We propose the following harmonisation, *Number of books (1: None or very few; 2: Enough to fill one shelf; 3: Enough to fill one bookcase; 4: Enough to fill two or more bookcases)*

Since TIMSS and PIRLS use both numbers and visual aids to approximate the number of books in the house, we will use their categories as the base, with the exception of combining the last two categories (enough books to fill two or more shelves). We will best match the other surveys to this proposed harmonisation. We will be unable to match the exact ranges of the number of books, but we will approximately match the categories that most resemble one another.

In Growing Up in Ireland '98, the first wave of data records the number of children's books in the house, and not the total number of books. Since this does not easily compare with the total number of books in the house, we will omit this variable and instead only use the number of books recorded in the second wave, which records total number of books.

Proposed harmonisation:

Operationalization	**Available datasets**
Number of books (1: None or very few; 2: Enough to fill one shelf; 3: Enough to fill one bookcase; 4: Enough to fill two or more bookcases)	PIRLS, TIMSS, PISA, PIAAC, GUI, NEPS, TREE2,

3.1.11 Degree of Urbanisation

The degree of urbanisation defines the area where the student's school is located. This variable is available in PIRLS, TIMSS, PISA and EUROSTUDENT and is collected from the principals. NEPS also includes information on the school area, but unfortunately, this data is not available in their downloadable files. Growing Up in Ireland also records the degree of urbanisation, but provides information about the participant's main residence instead of their school. This is the main source of variation in the data, as we know students may commute from rural areas into more densely populated areas for their education, thus the degree of urbanisation of the school may not be the same as the student's main residence.

The most recent standardization for degree of urbanisation suggested by the European Commission includes three categories - the first being densely populated area or alternatively named cities, the second being intermediate density area or alternatively named towns and suburbs and the third being thinly populated area or alternatively named rural area (Dijkstra and Poelman 2014). This proposed standardization uses a population grid and the percentage of the population living in high-density clusters and rural grid cells. We will not be able to apply this exact definition to our datasets as each survey uses its own definition of urbanisation and we do not have access to the exact addresses to precisely locate the schools. We will however implement the same three categories of urbanisation in our harmonised variable.

Table 22: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
PIRLS	2016	-School location immediate area *(1: Urban–Densely populated; 2: Suburban–On fringe or outskirts of urban area; 3: Medium size city or large town; 4: Small town or village; 5: Remote rural)* -How many people live in area *(1: More than 500,000 people; 2: 100,001 to 500,000 people; 3: 50,001 to 100,000 people; 4: 30,001 to 50,000 people; 5: 15,001 to 30,000 people; 6: 3,001 to 15,000 people; 7: 3,000 people or fewer)*
PIRLS	2011	-School location immediate area *(1: Urban; 2: Suburban; 3: Medium city; 4: Small town; 5: Remote rural)* -How many people live in area *(1: More than 500,000 people; 2: 100,001 to 500,000 people; 3: 50001 to 100,000 people; 4: 15,001 to 50,000 people; 5: 3,001 to 15,000 people; 6: 3,000 people or fewer)*
PIRLS	2006	-School location immediate area *(1: Urban; 2: Suburban; 3: Rural)* - How many people live in area *(1: Fewer than 3,000 people; 2: 3,001 to 15,000 people; 3: 15,001 to 50,000 people; 4: 50,001 to 100,000 people; 5: 100,001 to 500,000 people; 6: More than 500,000 people)*
PIRLS	2001	-School location immediate area *(1: Urban; 2: Suburban; 3: Rural)* -How many people live in area *(1: Less than 3,000 people; 2: 3,001 to 100,000 people; 3: 100,001 to 500,000 people; 4: More than 500,000 people)*
TIMSS G4 and G8	2019, 2015	-School location immediate area *(1: Urban–Densely populated; 2: Suburban–On fringe or outskirts of urban area; 3: Medium size city or large town; 4: Small town or village; 5: Remote rural)*

		-How many people live in area *(1: More than 500,000 people; 2: 100,001 to 500,000 people; 3: 50,001 to 100,000 people; 4: 30,001 to 50,000 people; 5: 15,001 to 30,000 people; 6: 3,001 to 15,000 people; 7: 3,000 people or fewer)*
TIMSS G4 and G8	2011	-School location immediate area *(1: Urban; 2: Suburban; 3: Medium city; 4: Small town; 5: Remote rural)* -How many people live in area *(1: More than 500,000 people; 2: 100,001 to 500,000 people; 3: 50,001 to 100,000 people; 4: 15,001 to 50,000 people; 5: 3,001 to 15,000 people; 6: 3,000 people or fewer)*
TIMSS G4 and G8	2007, 2003	-How many people live in area *(1: More than 500,000; 2: 100,001 to 500,000; 3: 50,001 to 100,000; 4: 15,001 to 50,000; 5: 3001 to 15,000; 6: 3,000 or fewer)*
PISA	2018, 2015, 2012, 2009, 2006, 2003, 2000	School location *(1: A village, hamlet or rural area (fewer than 3 000 people); 2: A small town (3 000 to about 15 000 people); 3: A town (15 000 to about 100 000 people); 4: A city (100 000 to about 1 000 000 people); 5: A large city (with over 1 000 000 people))*
EUROSTUDENT	2019	Where are you studying *(1: Less than 300,000 ; 2: More than 300,000)*
PIAAC	All	N/A

Table 23: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	(waves 1, 4)	Region of household *(1: Urban; 2: Rural)*
NEPS SC4	All	N/A (not available in download data) Can be found in RemoteNEPS or OnSite data
TREE2	All	N/A
ÉpStan	All	N/A

We propose the following harmonisation: *Degree or Urbanisation (1: City; 2: Town or Suburb; 3: Rural Area).*

All of the datasets define the degree of urbanisation either by using the approximate size of the population in the area, or by providing a choice of categorical responses. These categories include but are not limited to the following options: urban centre, city, suburb, town, medium town, small town, village and rural. For the harmonisation, we will group all categories mentioning urban centres or cities as Group 1, all categories including medium towns, towns and suburbs as Group 2 and all categories including villages, small towns and rural areas as Group 3.

For the two years of TIMSS (2007 and 2003), which only report degree of urbanisation in terms of population size, we will match the population categories to those in PISA, which defines urbanisation in terms of both population size and categories. Then we will recode the responses into the three harmonised categories.

The EUROSTUDENT survey proposes only two ranges of population size for the degree of urbanisation. For this survey, we will group populations greater than 300,000 as Group 1, and those less than 300,000 as Group 2. We will omit the Group 3 in the EUROSTUDENT sample, since the survey sample is drawn

uniquely from current higher education students and typically, higher education institutions are not located in rural areas.

Growing up in Ireland's variable varies from the others as it records the degree of urbanisation of the household and not the school. This variable will thus be left as is and not harmonised, since comparability with the other surveys is not possible.

Proposed harmonisation:

Operationalization	Available datasets
Degree or Urbanisation – School Area (1: City; 2: Town or Suburb; 3: Rural Area)	PIRLS, TIMSS, PISA, EUROSTUDENT

3.1.12 School Size

School size defines the enrolment of students attending a specific institution. The principals of the school generally report this value. PIRLS, TIMSS, PISA and GUI '98 all report the school size in terms of enrolment. NEPS SC4 also reports the school size, but this information is not available in the downloadable data due to anonymization and data restrictions associated with off-site data use. To facilitate the comparison of school size across all the surveys we will focus on the total enrolment of pupils, but we do acknowledge that some surveys also offer the enrolment of boys and girls separately, which could also be of interest for certain research questions.

Table 24: Cross sectional datasets

Dataset	Year	Variable label/definition *(Value scheme)*
PIRLS	2016	N/A
PIRLS	2011, 2006	Total enrolment of students *(VALUE)*
PIRLS	2001	-Total enrolment of girls -Total enrolment of boys *(VALUE)*
TIMSS G4	2019, 2015	N/A
TIMSS G4	2015, 2011, 2007, 2003	Total enrolment of students *(VALUE)*
TIMSS G8	2019, 2015	N/A
TIMSS G8	2015, 2011, 2007, 2003	Total enrolment of students *(VALUE)*
PISA	All	School size *(VALUE)*
EUROSTUDENT	All	N/A
PIAAC	All	N/A

Table 25: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	(wave 1)	-How many pupils are enrolled, Categories *(1: 1 - 80; 2: 81 - 120 ;3: 121 - 160; 4: 161 - 200; 5: 201 - 240; 6: 241 - 280; 7: 281 - 320; 8: 321 - 360; 9: 361 - 400; 10: 401+)* -How many boys are enrolled, Categories -How many girls are enrolled, Categories *(1: 0; 2: 1 - 40; 3: 41 - 80; 4: 81 - 120; 5: 121 - 160; 6: 161 - 200; 7: 201 - 240; 8: 241 - 280; 9: 281 - 320; 10: 321+)*
	(wave 2)	-How many pupils are enrolled, Categories *(3: 1- 160; 4: 161 - 200; 5: 201 - 240; 6: 241 - 280; 7: 281 - 320; 8: 321 - 360; 9: 361 - 400; 10: 401+)* -How many boys are enrolled, Categories -How many girls are enrolled, Categories *(1: 0; 2: 1 - 40; 3: 41 - 80; 4: 81 - 120; 5: 121 - 160; 6: 161 - 200; 7: 201 - 240; 8: 241 - 280; 9: 281 - 320; 10: 321+)*
NEPS SC4	All	Not available in Download data
TREE2	All	N/A
ÉpStan	All	N/A

We propose the following harmonisation: *School Size (1: 0 - 160 pupils; 2: 161 - 200 pupils; 3: 201 - 240 pupils; 4: 241 - 280 pupils; 5: 281 - 32 pupils; 6: 321 - 360 pupils; 7: 361 - 400 pupils; 8: 401-700 pupils; 9: 701-1000 pupils; 10: 1001-1500 pupils; 11: 1501+ pupils)*

The Growing up in Ireland '98, categories for school size will serve as our base for the harmonisation as it is the only survey to categorize the school size values. We will exceptionally add groups 8: 401-700 pupils, 9: 701-1000 pupils, 10: 1001-1500 pupils and 11: 1501+ pupils, in order to capture more information about the larger schools in PIRLS, TIMSS and PISA. However, for the Growing up in Ireland variable, the maximum categories will range from only 1 to 8. The last category will include schools with a total enrolment of 401 or more pupils, since we are unable to deduce larger schools from the original variable.

Proposed harmonisation:

Operationalization	Available datasets
School Size (1: 0 - 160 pupils; 2: 161 - 200 pupils; 3: 201 - 240 pupils; 4: 241 - 280 pupils; 5: 281 - 32 pupils; 6: 321 - 360 pupils; 7: 361 - 400 pupils; 8: 401 - 700 pupils; 9: 701-1000 pupils; 10: 1001 - 1500 pupils; 11: 1501+ pupils)	PIRLS, TIMSS, PISA,
School Size (1: 0 - 160 pupils; 2: 161 - 200 pupils; 3: 201 - 240 pupils; 4: 241 - 280 pupils; 5: 281 - 32 pupils; 6: 321 - 360 pupils; 7: 361 - 400 pupils; 8: 401+ pupils)	GUI '98

1.1.1 School Type: Public vs Private

School type defines the management of the school. A public education authority manages public schools. Private institutions (including churches, businesses and other non-government organizations) manage private schools. Information on school type is only available in two of the surveys in the inventory, PISA and NEPS SC4. They both report school type in a very similar manner, thus the harmonising procedure is straightforward.

Table 26: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
PIRLS	All	N/A
TIMSS G4	All	N/A
TIMSS G8	All	N/A
PISA	2018	School type (1: A public school (Managed by a public education authority, government agency, or governing board); 2: A private school (Managed by a non-government org; e.g. a church, trade union, business, or other private institution))
PISA	2015, 2012, 2003	School type (1: Private Independent; 2: Private Government-dependent; 3: Public)
PISA	2009, 2006, 2000	School type (1: Public; 2: Private)
EUROSTUDENT	All	N/A
PIAAC	All	N/A

Table 27: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	All	N/A
NEPS SC4	(waves 5, 7, 8, 9, 10, 11)	School authority (1: Public school; 2: Church school; 3: Another kind of private or free school)
TREE2	All	N/A
ÉpStan	All	N/A

We propose the following harmonisation*: School Type (1: Public; 2: Private).*

Proposed harmonisation:

Operationalization	Available datasets
School Type (1: Public; 2: Private)	PISA, NEPS

3.2 Educational Outcomes

3.2.1 Educational Achievement – Standardized Test Scores

Standardized test scores are the ideal measure for educational achievement. Most of the surveys in our inventory use standardized testing in order to measure student achievement in different subjects. This section will focus on assessments in the three following domains: reading, mathematics and science, as they are the most common subjects across the international surveys as well as the national surveys. In the international cross-sectional studies, PIRLS and PISA assess reading abilities, while TIMSS and PISA include assessments in mathematics and science. PIAAC assesses numeracy and literacy. In the national studies, GUI, NEPS, TREE2 and ÉpStan include assessments in mathematics in at least one wave of data, while GUI, NEPS and ÉpStan additionally include assessments in reading.

The majority of the surveys report achievement scores using plausible values, with the exception of GUI, which uses percentage correct and logit scores and TREE2, which uses WLE scores. In addition to plausible values, the international assessments (PIRLS, TIMSS, PISA and PIAAC) include benchmarks or proficiency levels, which offer some more understanding of the abilities associated with the different plausible value scores.

After an extensive literature review it is evident that the methods used to harmonise assessment scores across the international large-scale assessments are broad with no agreed upon "golden standard" for harmonisation. The surveys test different populations of students of differing ages and abilities, additionally the assessment have different content and formats. When combining assessments across surveys with similar target populations, for example PISA and TIMSS grade eight, Wu (Wu 2009) argues that the differences can be mostly attributed to the time lag between the tests and the differences in content. Although the time lag is easy to control for and practiced by researchers, there are no relevant examples of controls for balancing the content and the best method for doing so in unclear (Chmielewski and Dhuey 2017). Additionally, the approach of rescaling the scores by assuming the pooled standard deviation of a core group of countries is constant over time is used by researchers matching synthetic cohorts and comparing the dispersion of scores over time (Ammermueller 2013). Furthermore, researchers interested in observing the differences in performance between different groups have used achievement score gaps. For example, Chmielewski and Reardon (Chmielewski and Reardon 2016) measure income achievement gaps in PIRLS and PISA by standardizing the scores within each country than measuring the effect sizes of test scores for students in the 10^{th} and 90^{th} percentiles of the income distribution.

There is also a more recent strand of literature focusing on linking methodologies for harmonising assessment scores. Altinok, Anfrist and Patrinos (Altinok *et al.* 2018) describe the different linking methods and use both equipercentile and pseudo-linear linking methods in order to construct a global dataset with comparable achievement outcomes. The equipercentile linking method has further been used to link NEPS SC4 grade nine mathematics scores with PISA 2012 mathematics score in order to assign PISA proficiency levels to the NEPS test scores (Ehmke *et al.* 2020). These linking methods require a fully merged dataset as they use an anchoring process, which either adjusts results from the same population between two studies or with the same test items used in two different studies.

With regard to the national assessments, which have significant differences in their measurement of achievement the harmonisation of the these scores will not be covered in this guide, as any potential need for harmonisation of these scores will depend heavily on the research questions of the partners.

Table 28: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
PIRLS	2016, 2011, 2006, 2001 All	-Plausible value: overall reading PV1 – PV5 *(scores are scaled to have an international average value of 500 and a standard deviation of 100)* - International reading scale benchmark reached with PV1 – PV5 *(1: Not reaching low benchmark, below 400; 2: Low, at or above 400 but below 475; 3: Intermediate, at or above 475 but below 550; 4: High, at or above 550 but below 625; 5: Advanced, at or above 625)*
TIMSS G4 and G8	2019, 2015, 2011, 2007, 2003 All	-Plausible value mathematics PV1 – PV5 - Plausible value science PV1 – PV5 *(scores are scaled to have an international average value of 500 and a standard deviation of 100)* - International math bench reached with PV1 – PV5 - International science bench reached with PV1 – PV5 *(1: Not reaching low benchmark, below 400; 2: Low, at or above 400 but below 475; 3: Intermediate, at or above 475 but below 550; 4: High, at or above 550 but below 625; 5: Advanced, at or above 625)*
PISA	2018, 2015	-Plausible value in mathematics, PV1-PV10 -Plausible value in science, PV1-PV10 -Plausible value in reading, PV1-PV10 *(scale with a mean of 500 and a standard deviation of 100)*
PISA	2012, 2009, 2006, 2003, 2000	-Plausible value in mathematics, PV1-PV5 -Plausible value in science, PV1-PV5 -Plausible value in reading, PV1-PV5 *(scale with a mean of 500 and a standard deviation of 100)*
PISA	All	-Proficiency levels reading (*1c: 189.33 to less than 262.04; 1b: 262.04 to less than 334.75; 1a: 334.75 to less than 407.47; 2: 407.47 to less than 480.18; 3: 480.18 to less than 552.89; 4: 552.89 to less than 625.61; 5: 625.61 to less than 698.32; 6: 698.32-1000)* -Proficiency levels mathematics *(1a: 357.77 to less than 420.07; 2: 420.07 to less than 482.38; 3: 482.38 to less than 544.68; 4: 544.68 to less than 606.99; 5: 606.99 to less than 669.30; 6: 669.30 to 1000)* -Proficiency levels science *(1b: 260.54 to less than 334.94; 1a: 334.94 to less than 409.54; 2: 409.54 to less than 484.14; 3: 484.14 to less than 558.73; 4: 558.73 to less than 633.33; 5: 633.33 to less than 707.93; 6: 707.93 to 1000)*
EUROSTUDENT	2019	N/A
PIAAC	All	-Literacy scale score, PV1 – PV10 -Numeracy scale score, PV1 – PV10 -Problem-solving scale score, PV1 – PV10 *(scales ranges from 0-500)*

		-Literacy proficiency levels -Numeracy proficiency levels *(Below Level 1: 0-175; Level 1: 176-225; Level 2: 226-275; Level 3: 276-325; Level 4: 326-375; Level 5: 376-500)*

Table 29: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
GUI '98	(wave 1)	-Drumcondra reading test -Drumcondra math test *(VALUE: percentage correct)* *(VALUE: number of questions attempted)* *(VALUE: logit score and logit score standard error)*
	(wave 2)	-Drumcondra verbal reasoning -Drumcondra numerical ability *(VALUE: percentage correct)* *(VALUE: number of questions attempted)* *(VALUE: logit score and logit score standard error)*
NEPS SC4	All	-Mathematic competence, grade 9 -Reading competence, grade 9 -Scientific literacy, grade 9 -Scientific literacy, grade 11 -Mathematic competence, grade 12 -Reading competence, grade 12 *(VALUE: WLE scores and standard errors)* -Plausible values for all competence scores[16]
TREE2	(wave 0)	Math test, AES *(VALUE: WLE)*
ÉpStan Grade 3	2013	-Mathematics, grade 3 -German reading, grade 3 *(VALUE: mean of 500 and SD of 100)* - Proficiency levels[17] Mathematics grade 3 *(2: Niveau Socle, cutoff 437,565239; 3: Niveau Avancé, cutoff 520,5299777)* - Proficiency levels German reading, grade 3 *(2: Niveau Socle, cutoff 484,9013093; 3: Niveau Avancé, cutoff 543,9468151)*
ÉpStan Grade 9	2019	-Mathematics, grade 9 -German reading, grade 9 -French reading, grade 9 *(VALUE: Plausible value with mean of 500 and SD of 100)* - Proficiency levels Mathematics, grade 9 *(1: Niveau 1, cutoff* 440,0255126*; 2: Niveau 2, cutoff* 519,0609976*; 3: Niveau 3, cutoff* 685,2269239*)*

[16] NEPS competence scores can be used to generate plausible values specific to the individual research question using an adapted background model. An R package called NEPSscaling is available for this purpose NEPS Plausible Values.

[17] For all ÉpStan proficiency levels, the technical report outlines the procedures used in defining levels ÉpStan Technical Report.pdf (uni.lu).

		- Proficiency levels German reading, grade 9 *(1: Niveau 1, cutoff* 430,5543256*; 2: Niveau 2, cutoff* 486,9730662*; 3: Niveau 3, cutoff* 530,2337449*; 4 : Niveau 4, cutoff* 608,8507988*)* - Proficiency levels French reading, grade 9 *(1: Niveau 1, cutoff* 447,7770712*; 2: Niveau 2, cutoff* 548,7070539*; 3: Niveau 3, cutoff* 611,4473134 *; 4 : Niveau 4, cutoff* 662,7516393*)*

For the harmonisation, we propose two options: first, we propose the use of ***Plausible values (PVs).***

The international assessments along with some of the national assessments use PVs to report overall scores in a given subject. Generally reporting five PVs per student per subject, with the exception of the most recent PISA scores (2018 and 2015) and PIAAC, which report ten PVs each. The plausible values in PIRLS, TIMSS, PISA and PIAAC are scaled to have a mean of 500 and standard deviation of 100. Margaret Wu gives an excellent description of plausible values:

"One way to describe plausible values is to say that plausible values represent of the range of abilities that a student might reasonably have, given the student's item responses. […] Instead of directly estimating a student's Θ [students ability parameter], a probability distribution for a student's Θ is estimated. That is, instead of obtaining a point estimate for Θ, a range of possible values for a student's Θ, with an associated probability for each of these values is estimated. Plausible values are random draws from this (estimated) distribution for a student's Θ" (Wu 2005)

Using PVs correctly in empirical work is important. Some common shortcuts such as taking an average of the available PVs and applying it as if it were the only estimate of performance should be avoided (Aparicio *et al.* 2021). For proper use of PVs in secondary analysis, the technical reports and user manuals of the respective large-scale study should be addressed. Additionally, most data processing software is equipped with routines or commands that correctly perform estimations using PVs, for example PV (Macdonald 2008) and REPEST (Avvisati and Keslair 2014) in Stata.

Combining the international large-scale assessments is evidently not a straightforward process and depends heavily on the research question and structure. Although surveys such as PIRLS, TIMSS and PISA report plausible values with the same mean score of 500 and standard deviation of 100 does not mean they are directly comparable. As can be deduced from the quick summary of some of the literature above, there are many different methods for dealing with the comparability of the PV scores, and each method is designed for the different research questions posed. We thus propose that the researchers use the PVs along with the method of analysis that best fits their research question.

Proposed harmonisation:

Operationalization	**Available datasets**
Plausible values	PIRLS, TIMSS, PISA, PIAAC, NEPS, ÉpStan

As a second recommendation, we propose the use of ***Proficiency Levels (PLs)***.

The PLs in PIRLS, TIMSS, PISA and PIAAC group students into five or six levels of proficiency. Each PL is accompanied by a detailed definition of the skills and competencies of the tasks at that level, to help users understand the meaning of each plausible value score. The benchmarks used in TIMSS and PIRLS include five levels, 1: Not reaching low benchmark, 2: Low, 3: Intermediate, 4: High and 5: Advanced (Mullis *et al.* 2017, 2020). PISA uses a six level proficiency scale, with the inclusion of some sub-levels

(for example levels 1a and 1b) depending on the subject of the assessment (OECD 2014c, 2017b, 2018b). As the levels increase so does the difficulty of the task items, thus students in higher levels are assigned a relatively high proficiency. PIAAC PLs range from 1 to 5, where the higher levels refer to higher proficiency in the given subject (OECD 2013d). Full detailed descriptions of the tasks associated with each level for the above-mentioned surveys can be found in the Annex 6.1.

Since the PLs in PISA and PIAAC are not included in the original datasets, we provide the syntax to recode the plausible values into PLs for these two surveys. The level cut-offs used in the syntax for the PISA scores follow the most recent technical report for PISA 2018. The PIAAC syntax is based on the PLs defined in the most recent technical report (Kentaro Yamamoto *et al.* 2019).

Alternatively, we can also identify the specific group of **low achievers** separately for both math and reading achievement based on the minimum PLs defined by UNESCO Institute of Statistics. Based on their recommendations the minimum reading proficiency in PISA 2012-2015 is level 2, and in PIRLS 2011-2016 is low, while the minimum math proficiency in PISA 2012-2015 is level 2 and in TIMSS grades 4 and 8 2015 is intermediate (UNESCO Institute for Statistics 2021).

For the comparability of the PLs across the large-scale assessments, please refer to the survey given descriptions of each level in Annex 6.1. We outline the most recent benchmark summaries and PL descriptions for PIRLS, TIMSS, PISA and PIAAC. For previous years descriptions you can refer to the respective PISA technical reports or TIMSS and PIRLS international results reports, the previous definitions may vary slightly, but with no significant changes.

Proposed harmonisation:

Operationalization	Available datasets
Proficiency levels	PIRLS, TIMSS, PISA, PIAAC

3.2.2 Educational Attainment

Educational attainment is an important outcome variable in education research. In the surveys, we will outline two different definitions for educational attainment. First, completed educational attainment is the highest level of education achieved to date by the participant. Second, current educational attainment is the highest in-progress educational attainment of the participant. This is an important distinction since our inventory includes many datasets with sample populations that cover different stages of one's educational career, thus students still completing primary or secondary would not have yet had the chance to attain tertiary levels of education. Educational attainment is available in PISA, EUROSTUDENT, PIAAC, LIS, GUI, NEPS and TREE2.

Table 30: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
PRILS	All	N/A
TIMSS	All	N/A
PISA	2018	ISCED Level *(1: ISCED level 1; 2: ISCED level 2; 3: ISCED level 3; 4: ISCED level 4; 5: ISCED level 5)*
PISA	2015	ISCED Level *(2: ISCED level 2; 3: ISCED level 3; 4: ISCED level 4)*
PISA	2012, 2009, 2006, 2003	ISCED Level *(1: ISCED level 1; 2: ISCED level 2; 3: ISCED level 3)*
PISA	2000	N/A
EUROSTUDENT	2019	With which degree does your current study programme conclude? *(1: ISCED 5 short cycle degree; 2: ISCED 6 bachelor degree; 3: ISCED 6-UP TO 3 YEARS short national degree; 4: ISCED 7 master degree; 5: ISCED 7-more than 3 years long national degree; 6: ISCED 7 other postgraduate degree)*
PIAAC	All	-Education - Highest qualification – Level *(1: No formal qualification or below ISCED 1; 2: ISCED 1; 3: ISCED 2; 4: ISCED 3C shorter than 2 years; 5: ISCED 3C 2 years or more; 6; ISCED 3A-B; 7: ISCED 3 (without distinction ABC 2+ years); 8: ISCED 4C; 9: ISCED 4A-B; 10: ISCED 4 (without distinction ABC); 11: ISCED 5B; 12: ISCED 5A, bachelor; 13: ISCED 5A, master; 14: ISCED 6; 15: foreign qualification; 16: ISCED 5A and 6 (without distinction))*
LIS	All	Highest completed education level *(100 low, less than upper secondary; 110 less than primary ISCED 0; 111 never attended; 120 primary ISCED 1; 130 lower secondary ISCED 2; 200 medium, upper secondary and post-secondary non-tertiary; 210 upper secondary ISCED 3; 220 post-secondary non-tertiary ISCED 4; 300 high, tertiary; 310 BA, MA or equivalent, short-cycle tertiary; 311 short-cycle tertiary ISCED 5; 312 bachelor or equivalent ISCED 6; 313 master or equivalent ISCED 7; 320 doctorate or equivalent ISCED 8)*

Table 31: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	(wave 4)	Did you complete the course: -Honours Bachelor Degree (NFQ Level 8) -Ordinary Bachelor Degree (NFQ Level 7) -Higher Certificate Course (NFQ Level 6)

		-Post-Leaving Cert Course (NFQ Level 5/6) -Apprenticeship -Solas (FÁS), Fáilte Ireland, Teagasc, etc. *(1: Yes; 2: No; 3: Still on course)*
NEPS SC4	All	Recent ISCED-97 *(0 [0A/1A] no qualification; 1 [2B] school-leaving qualification of a Hauptschule, Volksschule, vocational preparation measure; 2[2A] Mittlere Reife (intermediate school-leaving qualification), school-leaving qualification of a Realschule; 3 [3A] Fachhochschulreife, higher education entrance qualification; 4 [3B] apprenticeship, vocational school, school of public health (less than two years), civil servant subclerical class, basic professional skills; 5 [3C] civil servant, clerical class; 6 [4A] Fachhochschulreife, higher education entrance qualification (second cycle); 7 [4B] apprenticeship, vocational school, school of public health (less than two years), civil servant subclerical class, basic professional skills (second cycle); 8 [5B] diploma from a Fachakademie and Berufsakademie, university of applied sciences for public administration, school of public health (at least two years), master craftsman/ craftswoman/technician certificate, other qualification from a Fachschule, civil servant executive class; 9 [5A] Bachelor, Master, Diploma, Magister, Staatsexamen, civil servants in administration; 10 [6] doctorate, habilitation)*
TREE2	All	Type of educational program *(34: General/academic upper secondary; 35: Vocational/professional upper secondary; 341: General/academic upper secondary education (insufficient for level completion))*
ÉpStan	All	N/A

We propose the following harmonisation: *Completed/Current Attainment (1: Primary or less (ISCED 1 or less); 2: Lower secondary (ISCED 2); 3: Upper secondary (ISCED 3); 4: Post-secondary, non tertiary (ISCED-11 4/ISCED-97 4a-b); 5: Short-cycle degree (ISCED-11 5/ ISCED-97 5b); 6: Bachelors or Masters degree (ISCED-11 6-7/ ISCED-97 5a); 7: Doctorate degree (ISCED-11 8/ ISCED-97 6); 88: Other, including foreign)*

In PISA, the harmonised current educational attainment variable includes only ISCED-11 levels 2 and 3 due to the age of the sample population of the survey. The EUROSTUDENT data has similar restriction and only includes ISCED-11 levels 5 and 6, and the TREE2 data includes only ISCED-11 level 3 due to the participating sample population.

In PIAAC, the completed educational attainment variable includes ISCED levels that could not be distinguished between. The non-distinguishable ISCED-97 5A and 6 levels (including bachelor, masters and doctorate studies) were grouped with Group 6: Bachelors or Master's degree. Additionally, PIAAC is the only study, which identifies foreign qualifications as a separate variable. This response option will remain as is, but is only relevant in the PIAAC sample.

In the LIS data, three categories are problematic to assign to an ISCED level (see values 100, 200 and 300 in *Table 30*). For this reason, educational attainment in LIS will be categorised into three main categories based off of the desired harmonised variable, 1: Lower-secondary or less (ISCED 2 or less), 2: Upper-secondary or Post-secondary, non-tertiary (ISCED 3 and 4), 3: Short-cycle tertiary, Bachelors,

Masters or Doctorate degree (ISCED 5-8). By assigning less categories, we allow the use of all the data in the original variable.

In Growing Up in Ireland, the educational attainment information is only available in the fourth wave of data, and focuses only on attainment of higher education. In the previous waves of data, the current ISCED levels cannot be extracted. In order to remain in line with the proposed harmonisation, we propose two categories for educational attainment. We will group the students as either attaining higher education (ISCED-11 levels 5+) or not (ISCED-11 level 4 or less). We will match the wave 4 national attainment degrees with the corresponding ISCED level based on the recommendations in The Irish Educational System: A Note on Classification (Smyth 2008).

Proposed harmonisation:

Operationalization	Available datasets
Completed educational attainment (1: Primary or less (ISCED 1 or less); 2: Lower secondary (ISCED 2); 3: Upper secondary (ISCED 3); 4: Post-secondary, non tertiary (ISCED-11 4/ISCED-97 4a-b); 5: Short-cycle degree (ISCED-11 5/ ISCED-97 5b); 6: Bachelors or Masters degree (ISCED-11 6-7/ ISCED-97 5a); 7: Doctorate degree (ISCED-11 8/ ISCED-97 6); 88: Other, including foreign)	PIAAC, NEPS
Completed educational attainment (1: Lower-secondary or less (ISCED 2 or less), 2: Upper-secondary or Post-secondary, non-tertiary (ISCED 3 and 4), 3: Short-cycle tertiary, Bachelors, Masters or Doctorate degree (ISCED 5-8))	Above datasets and LIS

Operationalization	Available datasets
Current educational attainment (1: Primary or less (ISCED 1 or less); 2: Lower secondary (ISCED 2); 3: Upper secondary (ISCED 3); 4: Post-secondary, non tertiary (ISCED-11 4/ISCED-97 4a-b); 5: Short-cycle degree (ISCED-11 5/ ISCED-97 5b); 6: Bachelors or Masters degree (ISCED-11 6-7/ ISCED-97 5a); 7: Doctorate degree (ISCED-11 8/ ISCED-97 6); 88: Other, including foreign)	PISA, EUROSTUDENT, PIAAC, TREE,
Current educational attainment Completed educational attainment (1: No higher education (ISCED 4 or less), 2: Short-cycle tertiary, Bachelors, Masters or Doctorate degree (ISCED 5-8))	Above datasets and GUI

3.2.3 Sense of Belonging in Education/School

Sense of belonging in school is a common variable in many of the student surveys. PIRLS, TIMSS, PISA and EUROSTUDENT include sense of belonging in their surveys and record the responses using a Likert scale (a scale that measures agreement with a given statement). There are three major differences between the variables in the surveys. The first difference among the surveys is that PISA, PIRLS and TIMSS ask the respondent's level of agreement to an affirmative sentence about belonging (e.g. I belong in school) where the EUROSTUDENT survey uses a negative sentence (e.g. I don't belong in higher education). The second difference is the type of scale. Numerical scales, meaning the response options are a list of numbers between two extremes, exists in EUROSTUDENT. A bipolar scale, meaning it comprises two opposite continua (Menod and Bogner 2016), exists in PISA, PIRLS and TIMSS. The third difference amongst the surveys is the size of the scale used for responses, in PISA, PIRLS and TIMSS the responses are reported on a four-point scale, where EUROSTUDENT uses a five-point scale.

Table 32: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
PIRLS	2016, 2011	Belong at School (1: Agree a lot; 2: Agree a little; 3: Disagree a little; 4: Disagree a lot)
PIRLS	2006, 2001	N/A
TIMSS G4 TIMSS G8	2019, 2015, 2011	Belong at School (1: Agree a lot; 2: Agree a little; 3: Disagree a little; 4: Disagree a lot)
TIMSS G4 TIMSS G8	2007, 2003	N/A
PISA	2018, 2015, 2012, 2003	Belong at School (1: strongly agree; 2: agree; 3: disagree; 4: strongly disagree)
PISA	2009, 2006	N/A
PISA	2000	Belong at School (1: strongly disagree; 2: disagree; 3: agree; 4: strongly agree)
EUROSTUDENT	2019	Don’t really belong in higher education (Strongly agree - Do not agree at all, 5 point scale)
PIAAC	All	N/A
EU SILC	All	N/A

Table 33: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI ‘98	All	N/A
NEPS SC4	All	N/A
TREE2	All	N/A
Épstan	All	N/A

We propose the following harmonisation: *Belonging (Likert scale values between 1-5, by linear transformation)*

In PISA 2010 the value scheme is reversed starting with strongly disagree instead of strongly agree. A reversal of the value scheme is sufficient to harmonise the responses with the rest of

the PISA, PIRLS and TIMSS. In order to harmonise PISA, PIRLS and TIMSS with the EUROSTUDENT's 5-point scale, we reverse the EUROSTUDENT scale (to adjust for the difference in the original statements) and then we apply a linear transformation to the 4-point scales in PISA, PIRLS and TIMSS to adjust the metric.

Applying a linear transformation to the 4-point scales will increase the comparability between the different scales sizes. In order to maintain a midpoint we have opted to transform all scales to a Likert scale with values between 1 and 5. The linear transformation of the four-point scale will be applied as follows:

$$New\ Rating = \frac{4}{3} Old\ Rating - \frac{1}{3}$$

The only survey that differs in scale length is the EUROSTUDENT belonging variable. For researchers who do not wish to use the EUROSTUDENT data, it is sufficient to keep the original 4-point Likert scale, with the appropriate transformations for those scales, which have been reversed.

Additionally, as is the case for all harmonised variables, we create a new target variable and retain the original variable and categories in the datasets. This leaves the option for researchers to define the scales differently. We alternatively propose defining belonging in terms of a **binary extreme** (for example including only those students who feel they do not belong in school, or only those students who feel they do belong). A binary variable for belonging in school offers an alternative to linear scaling, and avoids complications that arise from the EUROSTUDENT belonging variable being the only one with a midpoint. It can also be easily derived from the original variables in the datasets to best fit the needs of the researchers.

Proposed harmonisation:

Operationalization	Available datasets
Belonging (Likert scale values between 1-5)	PIRLS, TIMSS, PISA, EUROSTUDENT

3.2.4 Educational Aspiration of Student

Educational aspiration defines the student's desired future educational attainment. Most surveys report educational aspiration, with the exception of PIRLS and TIMSS-grade 4 where the students are most likely too young to answer such a question, PIAAC where the observed sample has already completed their educational careers and ÉpStan. In general attainment is reported using the ISCED levels or is easily transformed into an ISCED level making the harmonisation of this variable straightforward. *Table 34* and *Table 35* summarize the variables available in the data.

Table 34: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
PIRLS	All	N/A
TIMSS G4	All	N/A
TIMSS G8	2019, 2015, 2011	How far in education do you expect to go (1: Finish Lower secondary; 2: Finish Upper secondary; 3: Finish Post-secondary, non-tertiary; 4: Finish Short-cycle tertiary; 5: Finish Bachelor's or equivalent; 6: Finish Postgraduate degree)
TIMSS G8	2011	How far in education do you expect to go (1: Finish Lower secondary; 2: Finish Upper secondary; 3: Finish Post-secondary, non-tertiary; 4: Finish Short-cycle tertiary; 5: Finish Bachelor's or equivalent; 6: Finish Postgraduate degree; 7: Don't know)
TIMSS G8	2007, 2003	How far in education do you expect to go (1: Finish ISCED 3; 2: Finish ISCED 4; 3: Finish 5b; 4: Finish 5a, First degree; 5: Finish Beyond ISCED 5a, First degree; 6: Don't know)
PISA	2018, 2009	-Do you expect to complete? ISCED level 2 -Do you expect to complete? ISCED level 3B or C -Do you expect to complete? ISCED level 3A -Do you expect to complete? ISCED level 4 -Do you expect to complete? ISCED level 5B -Do you expect to complete? ISCED level 5A or 6 (0: not checked; 1: checked)
PISA	2015	Which of the following do you expect to complete? (1: ISCED level 2; 2: ISCED level 3B or C; 3: ISCED level 3A; 4: ISCED level 4; 5: ISCED level 5B; 6: ISCED level 5A or 6)
PISA	2003	Which of the following do you expect to complete? (0: ISCED level 1; 1: ISCED level 2; 2: ISCED level 3B or C; 3: ISCED level 3A or 4; 4: ISCED level 5B; 5: ISCED level 5A or 6)
PISA	2012, 2006, 2000	N/A
EUROSTUDENT	2019	N/A
PIAAC	All	N/A

Table 35: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	(wave 2)	-Child's educational expectation (1: Junior Cert; 2: Leaving Cert; 3: Certificate or Diploma (including plc., apprenticeship); 4: Degree or higher degree)
NEPS SC4	(waves 5, 7-11)	-Idealistic educational aspiration - vocational qualification - higher education -Idealistic educational aspiration - vocational qualification - voc. training -Idealistic educational aspiration - vocational qualification - no voc. training (0: Not specified, 1: Specified)
TREE2	(waves 0, 1, 2)	Realistic educational aspirations (1: 2-year VET certificate at upper secondary level (EBA); 2: 3–4-year VET certificate at upper secondary level (EFZ, FMS, business or computer science diploma); 3: Vocational or specialized baccalaureate; 4: General education baccalaureate; 5: Non-university tertiary level VET (e.g. Eidg. Fachausweis, Meisterdiplom); 6: Education at a University of Applied Sciences / University of Teacher Education (completed with a bachelor's, master's or a doctoral degree etc.); 7: Education at a University or a Swiss Federal Institute of Technology (ETH) (completed with a bachelor's, master's or a doctoral degree etc.); 18: No educational certificate)
ÉpStan	All	N/A

We propose the following harmonisation: *Aspiration (1: Finish upper-secondary or below (ISCED-11 level 3 or less); 2: Finish post-secondary, non-tertiary (ISCED-11 level 4); 3: Finish short cycle tertiary (ISCED-11 level 5); 4: Finish tertiary –Bachelors, Masters or PhD (ISCED-11 6-8)).*

The PISA 2003 data is the main outlier when it comes to harmonising educational aspirations. In the original study, it groups together ISCED-97 levels 3A and 4, and thus we are unable to separate these two levels. Only for the PISA 2003 data will the ISCED-97 level 3A aspiration be included in Group 2: Post-secondary non-tertiary, due to the data constraints. Otherwise, in all the remaining surveys anyone who aspired to complete ISCED-97 level 3A will be placed in Group 1: Finish upper-secondary.

Due to the nature of the data in the national longitudinal studies, which often group together all higher education aspirations under one response option, in these studies we will group ISCED-11 levels 5-8 together into one final category.

The harmonisation of the TREE2 variable has two important notes when classifying the national programs into ISCED-11 levels. First, in Switzerland there are two types of vocational baccalaureates, one is an ISCED-11 level 3 and the other is an ISCED-11 level 4. The ISCED-11 level 3 baccalaureate is by far more common, thus, we will include the vocational

baccalaureates in Group 1: Finish upper-secondary or below (ISCED-11 level 3 or less). Secondly, the Switzerland ISCED mapping defines non-university tertiary level VET (e.g., Eidg. Fachausweis, Meisterdiplom) programmes as ISCED-97 level 5b (UNESCO Institute for Statistics 2022). This is somewhat uncommon compared to other countries with similar programmes, which define this type of education as ISCED-11 level 4. We will nonetheless stick to the ISCED mappings and include this program in Group 3: Finish tertiary. It is important to note that in TREE2 there are no observations in Group 2: Finish Post-Secondary (non-tertiary), due to the nature and ISCED designations of the national programs.

Proposed harmonisation:

Operationalization	**Available datasets**
Aspiration (1: Finish upper-secondary or below (ISCED level 3 or less); 2: Finish post-secondary, non-tertiary (ISCED level 4); 3: Finish short cycle tertiary (ISCED level 5); 4: Finish tertiary –Bachelors, Masters or PhD (ISCED 6-8))	TIMSS (grade 8), PISA
Aspiration (1: Finish upper-secondary or below (ISCED level 3 or less); 2: Finish post-secondary, non-tertiary (ISCED level 4); 3: Finish tertiary (ISCED levels 5-8))	Above datasets and NEPS, GUI, TREE2

3.2.5 Out of School Lessons / Tutoring

Out of school lessons/tutoring is of interest in shadow education research and the datasets relevant to this research topic within our project include, GUI '98, NEPS SC4, TREE2, PISA 2012 and TIMSS 2019.

The participation in private tuition variable is quite similar across the three longitudinal surveys. Focusing on the surveys administered to the students, the main difference is in the timing of the question. GUI asks about private tutoring in the final year of school (when students are between 17 and 18 years old), NEPS asks about tutoring received in Grade 10 (at about 15 years old) and TREE2 asks about tutoring in Grades 8 or 9 (around 14 and 15 years old).

The out-of-school lessons variable in the cross-sectional surveys differs in some of the response options. Where TIMSS specifies the reason for taking tutoring in the response options, PISA specifies the quantity of lessons.

Table 36: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
TIMSS G4 G8	2019	-Extra lessons, mathematics -Extra lessons, science *(1: Yes, to excel in class; 2: Yes, to keep up in class; 3: No)*
PISA	2012	-Out of school lessons, mathematics -Out of school lessons, science *(1: I do not attend <out-of-school time lessons> in this subject; 2: Less than 2 hours a week; 3: 2 or more but less than 4 hours a week ; 4: 4 or more but less than 6 hours a week; 5: 6 or more hours a week)*

Table 37: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition *(Value scheme)*
GUI '98	(wave 3) (wave 4)	In this/ final school year, have/had any grinds or private tuition in any of school subjects *(1: Yes; 2: No)*
NEPS SC4	(waves 1, 5, 7)	-Parent: Private tutoring - panel questions - occurrence *(1: Yes; 2: No)* - Parent: Private tutoring - panel questions - Subjects: Mathematics - Parent: Private tutoring - panel questions - Subjects: German - Parent: Private tutoring - panel questions - Subjects: English - Parent: Private tutoring - panel questions - Subjects: French - Parent: Private tutoring - panel questions - Subjects: Latin - Parent: Private tutoring - panel questions - Subjects: Physics - Parent: Private tutoring - panel questions - Subjects: Chemistry - Parent: Private tutoring - panel questions - Subjects: Biology

		- Parent: Private tutoring - panel questions - Subjects: other subjects *(0: not specified; 1: specified)*
NEPS SC4	(wave 3)	-Student: Tutoring yes/no *(1: Yes; 2: No)* -Student: Tutoring subject 1st specification -Student: Tutoring subject 2nd specification -Student: Tutoring subject 3rd specification *(1: Mathematics; 2: German; 3: English; 4: French; 5: Latin; 6: Physics; 7: Chemistry; 8: Biology; 9: other subject)*
TREE2	(wave 0)	-Paid after-school tutoring *(0: Respondent has not received paid after-school tutoring in the 8th or 9th grade; 1: Respondent has received paid after-school tutoring in the 8th or 9th grade)* -Private tutoring: frequency, test language -Private tutoring: frequency, mathematics -Private tutoring: frequency, foreign language -Private tutoring: frequency, other subjects *(1: No, never; 2: Yes, sometimes; 3: Yes, regularly)* -Private tutoring: reason - Prepare for next level -Private tutoring: reason - Pass entrance exam -Private tutoring: reason - Achieve better results -Private tutoring: reason - Catch up on material -Private tutoring: reason - Other reasons *(0: not applicable; 1: applicable)*

We propose the following harmonisation for the tutoring variable: Participate in tutoring *(1: Yes; 2: No).* In the cross-sectional datasets, we further distinguish between the two different types of tutoring since math tutoring and science tutoring are distinct in both TIMSS and PISA.

An additional note for researchers that wish to look at the different types of shadow education (compensatory vs enrichment) the TIMSS 2019 and TREE2 original variables contain information about the reason for taking private tutoring.

Proposed harmonisation:

Operationalization	**Available datasets**
Participate in tutoring *(1: Yes/ 2: No)*	GUI, NEPS, TREE
Out-of-school lesson: Math *(1: Yes/ 2: No)* Out-of-school lesson: Science *(1: Yes/ 2: No)*	TIMSS, PISA

3.2.6 Extra-Curricular Activities (Clubs, Sports, Religious Groups)

Extra-curricular activities are of interest in shadow education research and the datasets relevant to this research topic in our project include, GUI '98, NEPS SC4 and TREE2.

We define extra-curricular activities as participation in informal education activities. This includes after or before school activities at a club, association or volunteer centre where some sort of organized activity takes place. Some common examples taken from the surveys include participation in sports clubs, culture clubs, religious clubs, political associations, music lessons or volunteer aid organizations. Some of the surveys additionally include participation in leisure activities, such as pick-up sports without a coach or instructor, playing music, acting without an instructor, or attending religious services.

The longitudinal national surveys all include multiple binary questions about participation in specific clubs and afterschool activities. The activities specified in each of the datasets differ slightly with the only activity that is clearly stated in all three longitudinal studies namely participation in sports/fitness clubs. In order to best harmonise this variable for the specific needs of the researcher and since this variable is only of interest in Task 4.3 of Work Package 4, the researchers involved will themselves collaborate on a harmonisation that best suits the available data. See *Table 38* below for a full overview of the available variables on participation in informal education activities.

Table 38: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	(wave 1)	Does child participate outside school in: - Sports/fitness club - Cultural activities - Youth club - Scouts/Guides/Boys Brigade/ Girls Brigade - Homework Club - Other (1: Yes; 2: No)
GUI '98	(wave 2)	How often do you: -(a) Play sports or undertake physical activities without a coach or instructor - (b) Play sports with a coach or instructor, or as part of a organised team - (c) Take part in dance, drama or music lessons - (d) Take part in a homework club (either school or elsewhere) - (e) Take part in clubs, or groups such as Guides or Scouts, youth club, community (1: Never; 2/ Less than once a week; 3: 1 to 3 times a week; 4: Four or more times a week)
GUI '98	(wave 3)	In the past year took part in: - Sports clubs/teams - School/student councils

		- Breakfast club or after school club - Computer clubs/groups - Art, drama, dance or music clubs/groups/rehearsals - Religious groups or organizations - Youth clubs where you can hang out with other people - Games/hobbies clubs - Other (specify) (1: Yes)
NEPS SC4	(waves 1, 5)	- Participation organization/group: voluntary aid organizations - Participation organization/ group: sports club - Participation organization/ group: religious youth groups - Participation organization/ group: fan club - Participation organization/ group: culture club - Participation organization/ group: political association (1: Yes; 2: No) - Participation organization/ group: other (0: Not specified, 1: Specified)
TREE2	(waves 1, 2)	- Leisure activities - Dancing, acting, or playing music - Leisure activities - Doing sports - Leisure activities - Participation in the activities of a club or association (e.g., scouts, sports club, rifle association) - Leisure activities - Volunteer work in a association or social organisation - Leisure activities - Going to church, religious events - Leisure activities - Involvement in a political group (e.g., political party or some interest group) (1: (almost) daily; :2 several times a week; 3: about once a week; 4: about one to three times a month; 5: once a month or less; 6: never)

We propose the following harmonisation: Extra-curricular activities (1: Participation in sports clubs or teams; 2: Participation in art, drama, dance, music or cultural clubs/groups/rehearsals)

We focus on two main categories for the harmonisation of the participation in informal education activities that can be distinguished in all the three national surveys. The first category includes any participation in a sports club or team, the second category is broader and includes any activities related to the arts (drama, music, dance, ect.) or culture groups. All other activities were excluded from the harmonisation, as they were less relevant for the research questions at hand.

Alternatively, those who wish to include all the different extra-curricular activities, can code the remaining activities into a third category labelled “3: Other extra-curricular activities”.

Proposed harmonisation:

Operationalization	Available datasets
Extra-curricular activities (1: Participation in sports clubs or teams; 2: Participation in art, drama, dance, music or cultural clubs/groups/rehearsals)	GUI '98, NEPS SC4, TREE2

3.2.7 Educational Track/Programme

School tracking is a valuable variable in inequality research, especially in highly-stratified education systems. Additionally, there is a lot of variation in educational tracking variables, since they are based on national education systems. PISA, GUI '98, NEPS SC4, TREE2 and ÉpStan contain information about programme tracking of the students. These indices vary in detail and information depending on the country in question.

PISA is the only international survey that records the study programme indices. They derive their indices from the student tracking form of the student questionnaire. They include each countries national classification, as well as two standardized measures recording the ISCED-97 orientation and designation of the student's study programme. The ISCED-97 designation refers to level A, which includes general programs that grant access to the next level, level B, which includes any program that grants access the next level vocational studies, level C, which gives direct access to the labour market and level M which is a combination of the characteristics of the other levels (OECD 2018a). The orientation describes the curricular content of a programme, distinguishing between general, pre-vocational and vocational.

The national longitudinal surveys also collect information on programme indices. Each country uses their own national classification of school type or programme. Harmonising the national classifications required the collaboration and expertise of the respective national partners. Below in *Table 39* and *Table 40* you will find the definitions for school tracks in each of the datasets.

Table 39: Cross sectional datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
PIRLS	All	N/A
TIMSS G4	All	N/A
PISA	2018, 2015, 2012, 2009, 2006	-ISCED Orientation *(1: General; 2: Pre-Vocational; 3: Vocational; 4: Modular)* -ISCED Designation *(1: A; 2:B; 3:C; 4:M)* -Unique national study programme code *(VALUE, 7-digit code for each country specific programme)*
PISA	2003	-ISCED Orientation *(1: General; 2: Pre-Vocational; 3: Vocational)* -ISCED Designation *(1: A; 2:B; 3:C; 4:M)* -Unique national study programme code *(VALUE, 7-digit code for each country specific programme)*
PISA	2000	N/A
EUROSTUDENT	2019	N/A
PIAAC	All	N/A

Table 40: Longitudinal datasets

Dataset	Year (wave)	Variable label/definition (Value scheme)
GUI '98	(waves 3, 4)	Current programme or final year programme (1: Regular (Established) Leaving Certificate; 2: Leaving Certificate

		Applied (LCA); 3: Leaving Certificate Vocational (LCVP); 4: Something else (please specify))
NEPS SC4	(waves 1, 2, 3, 5, 7, 8)	Current type of school *(1: Elementary school; 2: Orientation stage; 3: Hauptschule; 4: Realschule; 5: School with multiple educational programs: No separation in school branches (yet); 6: School with multiple educational programs: School branch of the Hauptschule; 7: School with multiple educational programs: School branch of the Realschule; 8: School with multiple educational programs: Unclear; 9: Gymnasium; 10: Special needs school; 11: Comprehensive school: no separation of school branches (yet); 12: Comprehensive school: school branch of the Hauptschule; 13: Comprehensive school: school branch of the Realschule; 14: Comprehensive school: school branch of the Gymnasium; 15: Comprehensive school: unclear; 16: Waldorf school; 17: other school)*
NEPS SC4	(waves 5, 7) Not enough observations in data	Type of school *(1: Gymnasium; 2: Comprehensive school; 3: Higher vocational school; 4: Berufsfachschule; 5: Vocational college; 6: Berufsoberschule (BOS, BOB); 7: Fachakademie; 8: Fachoberschule (FOS); 9: Fachschule; 10: Gemeinschaftsschule; 11: Higher commercial school; 12: Integrated secondary school; 13: Oberschule; 14: Stadtteilschule (not Waldorf school); 15: Waldorf school; 16: other kind of school)*
TREE2	(waves 1, 2)	-Type of educational programme ISCED 2011 *(34: General/academic upper secondary education; 35: Vocational/professional upper secondary education; 341: General/academic upper secondary education (insufficient for level completion))* -Type of educational programme *(10: Pre-apprenticeship ; 11: 10th school year (12th year HarmoS) / one-year preparatory programme for choice of occupation / VET, integration year (for recently immigrated youths with insufficient command of the native language), preparation for upper-secondary school etc.; 12: 10th language school year / school exchange year (in a foreign-language region); 13: Au pair / language stay; 14: Housekeeping training year; 15: Social year (Sozialjahr); 16: Motivation semester; 17: Preparatory course; 29: Other bridge-year programmes / interim solutions; 30: 3–4-year firm-based vocational education and training (VET diploma with / without FVB) (training contract with a firm or training network); 31: 3–4-year full-time school-based vocational education and training (VET diploma with / without FVB) at (another) vocational school / vocational training centre; 32: 2-year firm-based vocational education and training (training contract with a firm or training network); 33: 2-year full-time school-based vocational education and training at a vocational school / vocational training centre ; 34: Upper-*

		secondary commercial / business school (VET diploma with / without FVB); 35: Upper-secondary IT school (VET diploma with / without FVB); 39: Baccalaureate (high-school, college, etc.); 40: Upper-secondary specialised school - specialised school diploma ; 42: German-speaking Switzerland: Integrative upper-secondary school Rudolf Steiner (Waldorf school); 43: Ticino: Cantonal baccalaureate in commerce and business + federal VET diploma (cantonal high school of commerce) ; 95: Other upper-secondary education ; 1000: Office administration diploma / commercial diploma (Swiss Association of Commercial Schools [VSH]); 1040: Language course ; 9000: Other education / further education and training / course [TEXT])
ÉpStan Grade 3	2013	N/A (no school tracks in grade 3)
ÉpStan Grade 9	2019	Educational tracking *("ES": Classical track (classique), "EST": technical track (général); "PRE": preparatory track (préparation))*

We propose the following harmonisation: *Educational track (1: General and Comprehensive, ISCED A designation, 2: Technical, ISCED B designation, 3: Vocational, ISCED C designation and 4: Other (including non-certifying programs and special needs education)).*

The harmonisation is based on the ISCED designation and orientation defined in PISA, with some deviations from these definitions where necessary. The ISCED orientation defines three distinct categories, including general, pre-vocational and vocational, where each category provides general information about the content of curriculum of a given education track (OECD 2018a).

The ISCED designation, which is defined using letters also has three categories, where designation A refers to tracks that lead to higher education (for example to university studies), designation B refers to tracks that lead to higher education vocation training and designation C refers to tracks that lead directly to the labour market (OECD 2018a). Our harmonisation will focus more heavily on the ISCED designation, as it provides more information about where the educational tracks lead and their impact on educational attainment. However, we will deviate slightly from the original ISCED designation used in PISA and defined in the Manual for ISCED-97 Implementation in OECD Countries - 1999 Edition (OECD 1999b), because for certain countries the strict definition of ISCED designations removes a lot of valuable information about the tracks from the data by grouping many different tracks together. Instead of applying the definition literally, we will apply it with caution. Meaning that in education systems with mobility between tracks, where many different tracks “could” lead to higher education, we will consider the educational trajectories of the majority of the students in those tracks in order to determine a designation. This is where our harmonisation will deviate from the one proposed in PISA, since PISA determines the designation based on any potential mobility between the tracks, our definition will focus on the most likely transitions to upper-secondary and post-secondary education of a given track.

We focus this harmonisation on the countries with planned empirical studies for Work Package 4 (Germany, Hungary, Ireland, Lithuania, Luxembourg, and Switzerland), and countries with known early tracking (where tracking begins before the age of 13). The ISCED designation defined by the OECD is generally quite reliable for countries with comprehensive education systems or with tracking at later ISCED levels. On the contrary, countries with early tracks generally have, at least theoretically speaking, more mobility between tracks and the OECD ISCED designation for these countries is less informative and hence less reliable. For this reason, we additionally include the following countries with known early tracking: Austria, Belgium, the Czech Republic, the Netherlands and the Slovak Republic (Woessmann 2009). For any remaining countries with educational tracks not included in this guide, we suggest the researchers refer to the ISCED designation provided by PISA.

Additionally, we will only define educational tracking for national programs at ISCED level 2 and 3 (lower and upper secondary), as this is where educational tracking is most prominent. The datasets included in the harmonisation of education track are: PISA, Growing up in Ireland '98, TREE2, NEPS-SC4 and ÉpStan.

It is important to note that within certain countries the tracking variable may not include all the different track options (especially for education systems that are comprehensive in nature). We also recognise that between countries, there may exist some conflicting understandings of the definitions for technical and vocational tracks that result in similar national programs being classified differently. For an enhanced international comparability of the educational tracks that avoids any confusion between technical and vocational classifications, we would suggest grouping the 2: Technical and 3: Vocational programs together.

Alternatively, we would suggest that researchers use the original definitions for ISCED designation and orientation proposed by the OECD and PISA, which uses a stricter definition of track designation and orientation and thus has enhanced international comparability.

However, we will leave this decision up to the researchers as we acknowledge that for some research questions, the benefits of keeping more information in the track variable may outweigh the need for enhanced international comparability. Overall, we emphasise that this harmonisation is a proposition and that researchers may use their discretion when deciding how they choose to define educational tracks in the context of their research questions.

Proposed harmonisation:

Operationalization	Available datasets
Educational Track (1: General and Comprehensive; 2: Technical; 3: Vocational 4: Other)	PISA, GUI '98, TREE2, NEPS-SC4, ÉpStan

The precise details for each country and dataset included in this proposed harmonisation can be found in the Appendix 6.2 (Tables 41-56).

4 Conclusion

The deliverable “Data harmonisation guidelines” pursue two principal goals: to identify and review data sources that are best suited for the objectives of our empirical work program; and to provide harmonisation solutions for key variables that enable the comparability of results across the countries, educational stages, and across different datasets in research that will be undertaken within PIONEERED.

The first part of the deliverable reviews 14 national and 11 international surveys and studies that offer potentially valuable content for research on educational inequalities. While many of the datasets were already identified during the project preparation stage, several new data sources were identified with the help of national partners during the preparation of the current deliverable. The main objective was to offer a comprehensive overview of data sources both across the countries, as well as cover various educational stages, from early childhood and care to tertiary education and beyond. Additionally, surveys that address topics relevant to education themes, such as students’ well-being and health, were also included. Finally, some surveys on adult population that can serve as an empirical basis for understanding the reproduction of inequalities within societies are also listed.

The second part of the deliverable provides harmonisation guidelines for key variables: one section of variables contains information pertinent to the axes of socio-economic inequality, the other section of variables is used for measuring achievement and related outcomes in education. Altogether, there are 16 variables for measuring intersectional inequalities, and 7 variables that can be treated as education-related outcomes. As anticipated from the start, the surveys and datasets vary substantially in the level of detail, consistency, and the availability of information, including considerable variation in defining socio-demographic variables. Importantly, there are significant differences in the structure of the national education systems and reporting of educational outcomes. Given the large scale of the deliverable and rather limited timeline, we pursued a pragmatic data-driven approach where a comparability across databases was the main priority. The syntax codes provided for Stata programmes are available upon request, and are accompanied with detailed comments to assist the national partners.

The principal challenge faced during the preparation of this deliverable has been related to the access restriction for some of the datasets and the technical documentation, such as codebooks and questionnaires. This led to the Lithuanian and Hungarian data sources not being integrated in the harmonisation guidelines, while Luxembourgish assessment data was only partially included. However, this does not pose a barrier to researchers that will utilise these datasets: provided guidelines and syntax code will serve them as an example and will enable them to prepare remaining data in a similar format.

5 References

Altinok, N., Angrist, N., and Patrinos, H.A., 2018. *Global Data Set on Education Quality (1965-2015)*. Washington, DC: World Bank, No. 8314.

Ammermueller, A., 2013. Institutional Features of Schooling Systems and Educational Inequality: Cross-Country Evidence from PIRLS and PISA. *German Economic Review*, 14 (2), 190–213.

Aparicio, J., Cordero, J.M., and Ortiz, L., 2021. Efficiency Analysis with Educational Data: How to Deal with Plausible Values from International Large-Scale Assessments. *Mathematics*, 9 (13), 1579.

Avvisati, F. and Keslair, F., 2014. *REPEST: Stata module to run estimations with weighted replicate samples and plausible values*. Boston College Department of Economics.

Buchmann, M., Bayard, S., Kriesi, I., Stoll, M., Marti, R., Käslin, R.S., Staffelbach, M., and Kappeler, S., 2013. *Data from: Kohorte 3: Befragung zu Lebensverhältnissen, Lebenserfahrungen und psychosozialer Entwicklung - 2006* [dataset]. Universität Zürich. Distributed by FORS, Lausanne.

Buchmann, M. and Igel, C., 2019. *COCON: Project Documentation 2005-2018*. Universität Zürich - Jacobs Center for Productive Youth Development, Lausanne.

Buchmann, M., Paroz, A.-L., Basler, A., and Steinhoff, A., 2019. *Data from: Cocon: Kohorte 1: Befragung zu Lebensverhältnissen, Lebenserfahrungen und psychosozialer Entwicklung - 2006-2018* [dataset]. Universität Zürich - Jacobs Center for Productive Youth Development. Distributed by FORS, Lausanne.

Buchmann, M., Paroz, A.-L., Steinhoff, A., and Basler, A., 2018. *Data from: Cocon: Kohorte 2: Befragung zu Lebensverhältnissen, Lebenserfahrungen und psychosozialer Entwicklung - 2006-2012* [dataset]. Universität Zürich -Jacobs Center for Productive Youth Development. Distributed by FORS, Lausanne.

Children's World, 2013. ISCWeB project – 10 parts of the questionnaire.

Chmielewski, A. and Dhuey, E., 2017. The Analysis of International Large-Scale Assessments to Address Causal Questions in Education Policy. *National Academy of Education's Workshop Series on Methods and Policy Uses of International Large-Scale Assessment (ILSA)*.

Chmielewski, A.K. and Reardon, S.F., 2016. Patterns of Cross-National Variation in the Association between Income and Academic Achievement. *AERA Open*, 2 (3), 233285841664959.

Cuppen, J., Muja, A., Hauschildt, K., Buck, D., and Daniel, A., 2021. *Eurostudent VII. Micro Data. Data and Methods Report*. Nijmegen, Netherlands: ResearchNed.

Cuppen, J., Muja, A., Hauschildt, K., Daniel, A., Buck, D., Mandl, S., and Unger, M., 2021. *Data from: Eurostudent VII* [dataset].

DAB Panel Study, 2021. *Documentation of the DAB panel study (Data-Release Version 5.0)*. University of Bern, Institut für Erziehungswissenschaft, Abteilung Bildungssoziologie.

DAB Panel Study: W1 - W8 [dataset]. Universität Bern, Institut für Erziehungswissenschaft, Abteilung Bildungssoziologie. Distributed by SWISSUbase, Lausanne, 2022. DOI: https://doi.org/10.48573/cf4p-6s72

Dijkstra, L. and Poelman, H., 2014. *A Harmonised Definition of Cities and Rural Areas: The New Degree of Urbanisation*. European Commission, Working Paper.

Economic and Social Research Institute, 2021. About Growing Up in Ireland [online]. *Growing Up in Ireland National Longitudinal Study of Children*. Available from: https://www.growingup.ie/about-growing-up-in-ireland/ [Accessed 15 Nov 2021].

Ehmke, T., van den Ham, A.-K., Sälzer, C., Heine, J., and Prenzel, M., 2020. Measuring Mathematics Competence in International and National Large Scale Assessments: Linking PISA and the National Educational Panel Study in Germany. *Studies in Educational Evaluation*, 65, 100847.

Erzinger, A.B., Hauser, M., Dutrevis, M., Hascher, T., Keller, R., Lenz, P., and Soucis, A., 2019. Erläuterungen zu den Skalen des Kontextfragebogens der ÜGK Sprachen 2017: Theoretischer Hintergrund, Inhalte und Konstrukte.

European Commission, 2020. Standardised Key Social Variables: Implementing Guidelines.

Eurostat, 2013. *Draft AES Manual: Version 9*. European Commission, Eurostat, Directorate F: Social Statistics, Unit F-3: Education, Health and Social Protection Statistics.

Eurostat, 2017. *2016 ASE Manual: Version 3*. European Commission, Eurostat, Directorate F: Social Statistics, Unit F-3: Labour Market and Lifelong Learning.

Eurostat, 2020. *Methodological Guidelines and Description of EU-SILC Target Variables: 2020 Operation*. European Commission, Eurostat, Directorate F: Social Statistics, Unit F-4: Quality of Life.

FDZ-LIfBi, 2022a. *Data Manual NEPS Starting Cohort 1– Newborns, Education from the Very Beginning, Scientific Use File Version 9.0.0*. Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.

FDZ-LIfBi, 2022b. *Data Manual NEPS Starting Cohort 5– First-Year Students, From Higher Educa- tion to the Labor Market, Scientific Use File Version 16.0.0*. Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.

FDZ-LIfBi, 2022c. *Data Manual NEPS Starting Cohort 6– Adults, Adult Education and Lifelong Learning, Scientific Use File Version 13.0.0*. Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study.

Foy, P., 2017. *TIMSS 2019 User Guide for the International Database*. Boston College, TIMSS & PIRLS International Study Center.

Foy, P., 2018. *PIRLS 2016 User Guide for the International Database*. Boston College, TIMSS & PIRLS International Study Center.

Foy, P., Arora, A., and Stanco, G.M., 2013. *TIMSS 2011 User Guide for the International Database*. Boston College, TIMSS & PIRLS International Study Center.

Foy, P. and Drucker, K.T., 2013. *PIRLS 2011 User Guide for the International Database*. Boston College, TIMSS & PIRLS International Study Center.

Foy, P. and Kennedy, A.M., 2008. *PIRLS 2006 User Guide for the International Database*. Boston College, TIMSS & PIRLS International Study Center.

Foy, P. and Olson, J.F., 2007. *TIMSS 2007 User Guide for the International Database*. Boston College, TIMSS & PIRLS International Study Center.

Frøyland, L.R., 2017. Ungdata – Lokale ungdomsundersøkelser Dokumentasjon av variablene i spørreskjemaet.

García-Pérez, J.I. and Hidalgo, M.H., 2014. Evaluación de PROA: su efecto sobre el rendimiento de los estudiantes.

Gonzalez, E.J. and Kennedy, A.M., 2003. *PIRLS 2001 User Guide for the International Database*. Boston College, TIMSS & PIRLS International Study Center.

Growing Up in Ireland, n.d. *Growing Up in Ireland Cohort '98 at 9 Years Codebook for Wave 1 of the Child Cohort*. Dublin: Economic and Social Research Institute (ESRI) and Department of Children and Youth Affairs and Trinity College Dublin.

Growing Up in Ireland, n.d. *Growing Up in Ireland Cohort '98 at 13 Years Codebook for Wave 2 of the Child Cohort*. Dublin: Economic and Social Research Institute (ESRI) and Department of Children and Youth Affairs and Trinity College Dublin.

Growing Up in Ireland, n.d. *Growing Up in Ireland Cohort '98 at 17/18 Years Codebook for Wave 3 of the Child Cohort*. Dublin: Economic and Social Research Institute (ESRI) and Department of Children and Youth Affairs and Trinity College Dublin.

Growing Up in Ireland, n.d. *Growing Up in Ireland Cohort '98 at 20 Years of Age Codebook for Wave 4 of the Child Cohort*. Dublin: Economic and Social Research Institute (ESRI) and Department of Children and Youth Affairs and Trinity College Dublin.

Hascher, T., Brühwiler, C., and Girnat, B., 2019. *Erläuterungen zu den Skalen des Kontextfragebogens der ÜGK 2016 Mathematikteil: Theoretischer Hintergrund und Forschungsinteressen*. Bern: Universität Bern, Pädagogische Hochschule St. Gallen (PHSG) und Pädagogische Hochschule FNHW.

Health Behaviour in School-aged Children, 2016a. *HBSC-2001/02, ed.1.0, Study Documentation*.

Health Behaviour in School-aged Children, 2016b. *HBSC-2005/06, ed.1.0, Study Documentation*.

Health Behaviour in School-aged Children, 2016c. *HBSC-2009/10, ed.1.0, Study Documentation*.

Health Behaviour in School-aged Children, 2018. *HBSC-2013/14, ed.1.0, Study Documentation*.

Hoffmeyer-Zlotnik, J.H.P., 2016. *Standardisation and Harmonisation of Socio-Demographic Variables (Version 2.0)*. Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften.

Hupka-Brunner, S., Jann, B., Koomen, M., Krebs-Oesch, D., Meyer, T., Müller, B., Von Rotz, C., Sacchi, S., and Wilhelmi, B., 2021. *TREE2 Study Design*. Bern: TREE.

Hupka-Brunner, S., Jann, B., Meyer, T., Imdorf, C., Sacchi, S., Müller, B., and et al., 2016. Erläuterungen zum Kontextfragebogen der ÜGK 2016: Allgemeiner Teil.

International Association for the Evaluation of Educational Achievement (IEA), 2003a. *PIRLS 2001 Assessment: Codebooks (Zip File)*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/pirls2001i/PIRLS2001_Pubs_UG.html.

International Association for the Evaluation of Educational Achievement (IEA), 2003. *Data from: PIRLS 2001 International Database* [dataset]. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

International Association for the Evaluation of Educational Achievement (IEA), 2005a. *TIMSS 2003 Assessment: Codebooks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/timss2003i/userguide.html.

International Association for the Evaluation of Educational Achievement (IEA), 2005b. TIMSS 2003 International Database. *TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College*.

International Association for the Evaluation of Educational Achievement (IEA), 2007a. *PIRLS 2006 Assessment: Codebooks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/pirls2006/user_guide.html#.

International Association for the Evaluation of Educational Achievement (IEA), 2007. *Data from: PIRLS 2006 International Database* [dataset]. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

International Association for the Evaluation of Educational Achievement (IEA), 2009a. *TIMSS 2007 Assessment: Codebooks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/TIMSS2007/idb_ug.html.

International Association for the Evaluation of Educational Achievement (IEA), 2009b. TIMSS 2007 International Database. *TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College*.

International Association for the Evaluation of Educational Achievement (IEA), 2013a. *PIRLS 2011 Assessment: PIRLS 2011 Codebooks (2MB)*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/pirls2011/international-database.html.

International Association for the Evaluation of Educational Achievement (IEA), 2013b. *TIMSS 2011 Assessment: TIMSS 2011 Fourth Grade Codebooks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/timss2011/international-database.html.

International Association for the Evaluation of Educational Achievement (IEA), 2013c. *TIMSS 2011 Assessment: TIMSS 2011 Eighth Grade Codebooks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/timss2011/international-database.html.

International Association for the Evaluation of Educational Achievement (IEA), 2013d. TIMSS 2011 International Database. *TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College*.

International Association for the Evaluation of Educational Achievement (IEA), 2017a. *TIMSS 2015 Assessment Frameworks: TIMSS 2015 Fourth Grade Codebook*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/timss2015/international-database/.

International Association for the Evaluation of Educational Achievement (IEA), 2017b. *TIMSS 2015 Assessment Frameworks: TIMSS 2015 Eighth Grade Codebook*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/timss2015/international-database/.

International Association for the Evaluation of Educational Achievement (IEA), 2017. *Data from: TIMSS 2015 International Database* [dataset]. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

International Association for the Evaluation of Educational Achievement (IEA), 2018. *Data from: PIRLS 2016 International Database* [dataset]. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

International Association for the Evaluation of Educational Achievement (IEA), 2019. *PIRLS 2016 Assessment Frameworks: Codebooks PIRLS 2016*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timssandpirls.bc.edu/pirls2016/international-database/index.html.

International Association for the Evaluation of Educational Achievement (IEA), 2021a. *IEA's Trends in International Mathematics and Science Study – TIMSS 2019: TIMSS 2019 Codebooks - Grade 4*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timss2019.org/international-database/.

International Association for the Evaluation of Educational Achievement (IEA), 2021b. *IEA's Trends in International Mathematics and Science Study – TIMSS 2019: TIMSS 2019 Codebooks - Grade 8*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Available from: https://timss2019.org/international-database/.

International Association for the Evaluation of Educational Achievement (IEA), 2021. *Data from: TIMSS 2019 International Database* [dataset]. TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College.

International Labour Organization, 2012. *International Standard Classification of Occupations 2008 (ISCO-08): Structure, group definitions and correspondence tables*. Geneva: International Labour Office.

Kalter, F., Kogan, I., and Dollmann, J., 2021. *Data from: Children of Immigrants Longitudinal Survey in Four European Countries - Germany (CILS4EU-DE) - Reduced versio*n [dataset]. Cologne: GESIS Data Archive.

Kentaro Yamamoto, Lale Khorramdel, Lale Khorramdel, and Usama S. Ali, 2019. Chapter 18: Scaling Outcomes. *In*: *Survey of Adult Skills Technical Report (3rd Edition).*

Konsortium ÜGK, 2019a. *Überprüfung der Grundkompetenzen. Nationaler Bericht der ÜGK 2016: Mathematik 11. Schuljahr*. Bern & Genf: EDK & SRED.

Konsortium ÜGK, 2019b. *Überprüfung der Grundkompetenzen. Nationaler Bericht der ÜGK 2017: Sprachen 8. Schuljahr*. Bern & Genf: EDK & SRED.

KRTK Adatbank, 2020. National Assessment of Basic Competencies [online]. *Databank of Centre for Economic and Regional Studies*. Available from: https://adatbank.krtk.mta.hu/en/adatbazisok/elerheto-adatbazisok/ [Accessed 25 Aug 2022].

Luxembourg Center for Educational Testing, 2021. ÉpStan Dashboard [online]. Available from: https://dashboard.epstan.lu/?lang=EN [Accessed 18 Nov 2021].

Luxembourg Income Study, 2019. *The LIS User Guide: 2019 Template*.

Luxembourg Income Study (LIS), 2019. *Generic Codebook LIS Database 2019 Template*.

Macdonald, K., 2008. *PV: Stata module to perform estimation with plausible values*. Boston College Department of Economics.

Martin, M.O., 2005. *TIMSS 2003 User Guide for the International Database*. 1st edition. 140 Commonwealth Ave., Chestnut Hill, MA: TIMSS & PIRLS International Study Center Lynch School of Education Boston College.

Menod, N. and Bogner, K., 2016. *Design of Rating Scales in Questionnaires*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences, No. Version 2.

Ministry of Education, 2010. *Evaluación general de diagnóstico 2009. Educación Primaria. Cuarto curso. Informe de resultados*. Madrid.

Ministry of Education, 2011. *Evaluación general de diagnóstico 2010. Educación Secundaria Obligatoria. Segundo curso. Informe de resultados*. Madrid.

Mullis, I.V.S., Martin, M.O., Foy, P., and Hooper, M., 2017. PIRLS 2016 International Results in Reading [online]. *Boston College, TIMSS & PIRLS International Study Center website*. Available from: http://timssandpirls.bc.edu/pirls2016/international-results/.

Mullis, I.V.S., Martin, M.O., Ruddock, G., O'Sullivan, C.Y., and Preuschoff, C., 2020. TIMSS 2019 International Results in Mathematics and Science [online]. *Boston College, TIMSS & PIRLS International Study Center*. Available from: https://timssandpirls.bc.edu/timss2019/international-results/.

Murray, A., Quail, A., McCrory, C., and Williams, J., 2013. *A Summary Guide to Wave 2 of the Infant Cohort (at 3 Years) of Growing Up in Irelan*. Irish Social Science Data Archive.

Murray, A., Williams, J., Quail, A., Neary, M., and Thornton, M., 2015. *A Summary Guide to Wave 3 of the Infant Cohort (at 3 Years) of Growing Up in Ireland*. Irish Social Science Data Archive.

National Education Panel Study (NEPS), 2020. *Codebook NEPS Starting Cohort 4—Grade 9 School and Vocational Training Scientific Use File Version 11.0.0*. Bamberg: Leibniz Institute for Educational Trajectories (LIfBi).

Nidegger, C., 2019. *Data from: ÜGK / COFO / VECOF 2016: Competencies of Swiss pupils in mathematics* [dataset]. Schweizerische Konferenz der kant. Erziehungsdirektoren - EDK, Universität Bern, Fachhochschule Nordwestschweiz - FHNW, Pädagogische Hochschule Schwyz - PHSZ, Université de Genève, Pädagogische Hochschule St. Gallen - PHSG, Canton de Genève, Scuola universitaria professionale della Svizzera italiana - SUPSI, FH Graubünden - FHGR, Universität Zürich, FORS - Centre de compétences suisse en sciences sociales, Deutsches Institut für Internationale Pädagogische Forschung - DIPF. Distributed by SWISSUbase, Lausanne. DOI: https://doi.org/10.23662/FORS-DS-1004-1

Nidegger, C., 2021. *Data from: ÜGK / COFO / VECOF 2017: Competencies of Swiss pupils in languages* [dataset]. Schweizerische Konferenz der kant. Erziehungsdirektoren - EDK, Universität Bern, Fachhochschule Nordwestschweiz - FHNW, Pädagogische Hochschule Schwyz - PHSZ, Pädagogische Hochschule St. Gallen - PHSG, Canton de Genève, Scuola universitaria professionale della Svizzera italiana - SUPSI, Pädagogische Hochschule Zürich - PHZH, PH Zug, Haute Ecole Pédagogique - HEP-VD, Institut de plurilinguisme, FH Graubünden - FHGR, Universität Zürich, FORS - Centre de compétences suisse en sciences sociales, Deutsches Institut für Internationale Pädagogische Forschung - DIPF. Distributed by SWISSUbase, Lausanne, 2021. DOI: https://doi.org/10.23662/FORS-DS-876-2

Norwegian Mother and Child Cohort Study, 2020. *Documentation of MoBa Instruments (as used in Q1 to Q-8 year).*
OECD, 1999a. *Measuring Student Knowledge and Skills: A New Framework for Assessment*. Paris: PISA, OECD Publishing.
OECD, 1999b. *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries: 1999 Edition*. Paris: OECD Publishing.
OECD, 2001a. *Database - PISA 2000: Codebook for Student Questionnaire Data File (120kB)*. Available from: https://www.oecd.org/pisa/data/database-pisa2000.htm.
OECD, 2001b. *Database - PISA 2000: Codebook for School Questionnaire Data File (52kB)*. Available from: https://www.oecd.org/pisa/data/database-pisa2000.htm.
OECD, 2002. *PISA 2000 Technical Report*. Paris: OECD Publishing.
OECD, 2003. *The PISA 2003 Assessment Framework: Mathematics, Reading, Science, Problem Solving, Knowledge and Skills*. Paris: PISA, OECD Publishing.
OECD, 2004a. *Database - PISA 2003: Codebook for Student Questionnaire*. Available from: https://www.oecd.org/pisa/data/database-pisa2003.htm.
OECD, 2004b. *Database - PISA 2003: Codebook for School Questionnaire*. Available from: oecd.org/pisa/data/database-pisa2003.htm.
OECD, 2005. *PISA 2003 Technical Report*. Paris: OECD Publishing.
OECD, 2006. *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*. Paris: PISA, OECD Publishing.
OECD, 2007a. *Database - PISA 2006: Codebook for Student Questionnaire Data File*. Available from: https://www.oecd.org/pisa/data/database-pisa2006.htm.
OECD, 2007b. *Database - PISA 2006: Codebook for School Questionnaire Data File*. Available from: https://www.oecd.org/pisa/data/database-pisa2006.htm.
OECD, 2009a. *PISA 2009 Assessment and Analytical Framework: Key Competencies in Reading, Mathematics and Science*. Paris: OECD Publishing.
OECD, 2009b. *PISA 2006 Technical Report*. Paris: OECD Publishing.
OECD, 2010a. *TALIS 2008 Technical Report*. Paris: OECD Publishing.
OECD, 2010b. *Data base PISA 2009: Codebook for Student Questionnaire Data File*. Available from: https://www.oecd.org/pisa/data/pisa2009database-downloadabledata.htm.
OECD, 2010c. *Data base PISA 2009: Codebook for School Questionnaire Data File*. Available from: https://www.oecd.org/pisa/data/pisa2009database-downloadabledata.htm.
OECD, 2012. *PISA 2009 Technical Report*. Paris: OECD Publishing.
OECD, 2013a. *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. Paris: OECD Publishing.
OECD, 2013b. *Database - PISA 2012 : Codebook for Student Questionnaire Data File*. Available from: https://www.oecd.org/pisa/data/pisa2012database-downloadabledata.htm.
OECD, 2013c. *Database - PISA 2012 : Codebook for School Questionnaire Data File*. Available from: https://www.oecd.org/pisa/data/pisa2012database-downloadabledata.htm.
OECD, 2013d. *OECD Skills Outlook 2013 First Results from the Survey of Adult Skills*. Paris: OECD Publishing.
OECD, 2014a. *PISA 2012 Technical Report*. Paris: OECD Publishing.

OECD, 2014b. *TALIS 2013 Technical Report*. Paris: OECD Publishing.
OECD, 2014c. *PISA 2012 Technical Report*. Paris: OECD Publishing.
OECD, 2016. *PISA 2015 Database: Codebooks for the Main Files*. Available from: https://www.oecd.org/pisa/data/2015database/.
OECD, 2017a. *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*. Paris: PISA, OECD Publishing.
OECD, 2017b. *PISA 2015 Technical Report*. Paris: OECD Publishing.
OECD, 2018a. *Chapter 16 Scaling Procedures and Construct Validation of Context Questionnaire Data*.
OECD, 2018b. *Chapter 15 Proficiency Scale Construction*. OECD Publishing.
OECD, 2019a. *PISA 2018 Assessment and Analytical Framework*. Paris: PISA, OECD Publishing.
OECD, 2019b. *TALIS 2018 and TALIS Starting Strong 2018 User Guide*. Paris: OECD Publishing.
OECD, 2019c. *TALIS 2018 Technical Report*. Paris: OECD Publishing.
OECD, 2019d. *Survey of Adult Skills Technical Report (3rd Edition)*. Paris: OECD Publishing.
OECD, 2019e. *PISA 2018 Database: Code Book*. Available from: https://www.oecd.org/pisa/data/2018database/.
OECD, 2019e. Where All Students Can Succeed. Paris: OECD Publishing.
OECD, n.d. PIAAC International Codebook Public Use File [online]. *OECD*. Available from: http://www.oecd.org/site/piaac/International%20Codebook_PIAAC%20Public-use%20File%20(PUF)%20Variables%20and%20Values.xlsx.
OECD, n.d. *Data from: PIAAC Data Public Use Files* [dataset].
Pham, G., Helbling, L., Verner, M., and Ambrosetti, A., 2019. *ÜGK – COFO – VeCoF 2017 results: Technical appendices*. St.Gallen & Genève: Pädagogische Hochschule St.Gallen (PHSG) & Service de la recherche en éducation (SRED).
Pham, G., Helbling, L., Verner, M., Petrucci, F., Angelone, D., and Ambrosetti, A., 2019. *ÜGK– COFO – VeCoF 2016 results: Technical appendices*. St.Gallen & Genf: Pädagogische Hochschule St.Gallen (PHSG) & Service de la recherche en éducation (SRED).
Quail, A., Williams, J., McCrory, C., Murray, A., and Thornton, M., 2011. *A Summary Guide to Wave 1 of the Infant Cohort (at 9 Months) of Growing up in Ireland*. Irish Social Science Data Archive.
Quail, A., Williams, J., Thornton, M., and Murray, A., 2014. *A Summary Guide to Wave 2 of the Child Cohort (at 13 Years) of Growing up in Ireland*. Irish Social Science Data Archive.
Schneider, S. and Ortmanns, V., 2019. *Database of Educational Attainment, with Explanatory Note*. European Union's Horizon 2020 research and innovation programme under grant agreement No 654221, Deliverable D8.8.
Skopek, J., Pink, S., and Bela, D., 2012a. *Data Manual. Starting Cohort 2 – From Kindergarten to Elementary School. NEPS SC2 1.0.0*. NEPS Research Data Paper, University of Bamberg.
Skopek, J., Pink, S., and Bela, D., 2012b. *Data Manual. Starting Cohort 3 – From Lower to Upper Secondary School. NEPS SC3 1.0.0*. NEPS Research Data Paper, University of Bamberg.

Skopek, J., Pink, S., and Bela, D., 2013. *Starting Cohort 4: 9th Grade (SC4). SUF Version 1.1.0. Data Manual. NEPS Research Data Paper*. National Educational Panel Study (NEPS), University of Bamberg.

Smyth, E., 2008. The Irish Educational System: A Note on Classification. *In*: *The International Standard Classification of Education (ISCED-97): An Evaluation of Content and Criterion Validity for 15 European Countries*. Mannheim: Mannheimer Zentrum für Europäische Sozialforschung (MZES).

Strand, N.P. and von Soest, T., 2008. *Young in Norway – Longitudinal. Documentation of design, variables, and scales. Revised version*. Oslo: NOVA.

TARKI Foundation, 2006. *Data from: TDATA-H81: Life Course - Sociological Survey on Students Dropout from Secondary School 1* [dataset]. Version: 2019-05-29. Hungary.

TARKI Foundation, 2007. *Data from: TDATA-H82: Life Course - Sociological Survey on Students Dropout from Secondary School 2* [dataset]. Version: 2019-05-29. Hungary.

TARKI Foundation, 2008. *Data from: TDATA-H83: Life Course - Sociological Survey on Students Dropout from Secondary School 3* [dataset]. Version: 2019-05-29. Hungary.

TARKI Foundation, 2009. *Data from: TDATA-H84: Life Course - Sociological Survey on Students Dropout from Secondary School 4* [dataset]. Version: 2019-05-29. Hungary.

TARKI Foundation, 2011. *Data from: TDATA-H85: Life Course - Sociological Survey on Students Dropout from Secondary School 5* [dataset]. Version: 2019-05-29. Hungary.

TARKI Foundation, 2012. *Data from: TDATA-H86: Life Course - Sociological Survey on Students Dropout from Secondary School 6. Wave* [dataset]. Version: 2019-05-23.Hungary.

Thornton, M., Williams, J., McCrory, C., Murray, A., and Quail, A., 2010. *Guide to the Datasets Wave 1 of the Nine-Year Cohort of Growing up in Ireland*. ESRI & Office of the Minister for Children and Youth Affairs.

TREE, 2016a. *Documentation on the first TREE cohort (TREE1), 2000–2016*. Bern: TREE.

TREE, 2016b. *Transitions from Education to Employment Cohort 2 (TREE2) Panel waves 0-2 (2016-2018) Codebook for data of panel wave 0 (baseline survey, 2016)*. Bern University.

TREE, 2017. *Transitions from Education to Employment Cohort 2 (TREE2) Panel waves 0-2 (2016-2018) Codebook for data of panel wave 1 (2017)*. Bern University.

TREE, 2018. *Transitions from Education to Employment Cohort 2 (TREE2) Panel waves 0-2 (2016-2018) Codebook for data of panel wave 2 (2018)*. Bern University.

TREE, 2019. *Data from: Transitions from Education to Employment, Cohort 1 – 2000-2014* [dataset]. University of Bern. Distributed by FORS, Lausanne. DOI: https://doi.org/10.23662/FORS-DS-816-7

TREE, 2021. *Data from: Transitions from Education to Employment, Cohort 2 (TREE2), Panel waves 0-2 (2016-2018)* [dataset]. University of Bern. Distributed by FORS, Lausanne. DOI: https://doi.org/10.23662/FORS-DS-1255-1

UNESCO Institute for Statistics, 2021. SDG 4 Ensure Inclusive and Equitable Quality Education and Promote Lifelong Learning Opportunities for All.

UNESCO Institute for Statistics, 2022. ISCED Mappings [online]. *UNESCO Institute for Statistics*. Available from: http://uis.unesco.org/en/isced-mappings.

Ursin, J., Hyytinen, H., and Silvennoinen, K., 2021. *Assessment of Undergraduate Students' Generic Skills in Finland Findings of the Kappas! Project*. Helsinki: Ministry of Education and Culture.

Woessmann, L., 2009. *International Evidence on School Tracking: A Review*. Research Report No. 1.

Wu, M., 2005. The Role of Plausible Values in Large-Scale Surveys. *Studies in Educational Evaluation*, 31 (2–3), 114–128.

Wu, M., 2009. A Comparison of PISA and TIMSS 2003 Achievement Results in Mathematics. *PROSPECTS*, 39 (1), 33–46.

6 Appendices

6.1 Benchmarks and Proficiency Levels in PIRLS, TIMSS, PISA, PIAAC

6.1.1 Reading Benchmarks and Proficiency Scores in PIRLS and PISA

PIRLS 2016 Reading Benchmark Summary	PISA 2018 Reading Proficiency level descriptions
Low (400) When reading predominantly simpler Literary Texts, students can: • Locate and retrieve explicitly stated information, actions, or ideas • Make straightforward inferences about events and reasons for actions • Begin to interpret story events and central ideas When reading predominantly simpler Informational Texts, students can: • Locate and reproduce explicitly stated information from text and other formats (e.g., charts, diagrams) • Begin to make straightforward inferences about explanations, actions, and descriptions	**Level 1c** Readers at Level 1c can understand and affirm the meaning of short, syntactically simple sentences on a literal level, and read for a clear and simple purpose within a limited amount of time. Tasks at this level involve simple vocabulary and syntactic structures.
Intermediate (475) When reading a mix of simpler and relatively complex Literary Texts, students can: • Independently locate, recognize, and reproduce explicitly stated actions, events, and feelings • Make straightforward inferences about the attributes, feelings, and motivations of main characters	**Level 1b** Readers at Level 1b can evaluate the literal meaning of simple sentences. They can also interpret the literal meaning of texts by making simple connections between adjacent pieces of information in the question and/or the text. Readers at this level can scan for and locate a single piece of prominently placed, explicitly stated information in a single sentence, a short text or a simple list. They can access a relevant page from a small set based on simple prompts when explicit cues are present. Tasks at Level 1b explicitly direct readers to consider relevant factors in the task and in the text. Texts at this level are short and typically provide support to the reader, such as through repetition of information, pictures or familiar symbols. There is minimal competing information.
	Level 1a Readers at Level 1a can understand the literal meaning of sentences or short passages. Readers at this level can also recognise the main theme or the author's purpose in a piece of text about a familiar topic, and make a simple connection between several adjacent pieces of information, or between the given information and their own prior knowledge. They can select a relevant page from a small set based on simple prompts, and locate one or more independent pieces of information within short texts. Level 1a readers can reflect on the overall purpose, and on the relative importance of information (e.g., main idea vs.non-essential detail) in simple texts containing explicit cues. Most tasks at this level contain explicit cues as regards what needs to be done, how to do it, and where in the text(s) readers should focus their attention.
	Level 2 Readers at Level 2 can identify the main idea in a piece of text of moderate length. They can understand relationships or construe meaning within a limited part of the text when the information is not

• Interpret obvious reasons and causes, recognize evidence, and give examples • Begin to recognize language choices When reading a mix of simpler and relatively complex Informational Texts, students can: • Locate and reproduce two or three pieces of information from text • Make straightforward inferences to provide factual explanations • Begin to interpret and integrate information to order events	prominent by producing basic inferences, and/or when the text(s) include some distracting information. They can select and access a page in a set based on explicit though sometimes complex prompts, and locate one or more pieces of information based on multiple, partly implicit criteria. Readers at Level 2 can, when explicitly cued, reflect on the overall purpose, or on the purpose of specific details, in texts of moderate length. They can reflect on simple visual or typographical features. They can compare claims and evaluate the reasons supporting them based on short, explicit statements. Tasks at Level 2 may involve comparisons or contrasts based on a single feature in the text. Typical reflective tasks at this level require readers to make a comparison or several connections between the text and outside knowledge by drawing on personal experience and attitudes.
High (550) When reading relatively complex Literary Texts, students can: • Locate and distinguish significant actions and details embedded across the text • Make inferences to explain relationships between intentions, actions, events, and feelings, and give text-based support • Interpret and integrate story events and character actions, traits, and feelings as they develop across the text • Recognize the use of some language features (e.g., metaphor, tone, imagery) When reading relatively complex Informational Texts, students can: • Locate and distinguish relevant information within a dense text or a complex table • Make inferences about logical connections to provide explanations and reasons • Integrate textual and visual information to interpret the relationship between ideas • Evaluate and make generalizations about content and textual elements	**Level 3** Readers at Level 3 can represent the literal meaning of single or multiple texts in the absence of explicit content or organisational clues. Readers can integrate content and generate both basic and more advanced inferences. They can also integrate several parts of a piece of text in order to identify the main idea, understand a relationship or construe the meaning of a word or phrase when the required information is featured on a single page. They can search for information based on indirect prompts, and locate target information that is not in a prominent position and/or is in the presence of distractors. In some cases, readers at this level recognise the relationship between several pieces of information based on multiple criteria. Level 3 readers can reflect on a piece of text or a small set of texts, and compare and contrast several authors' viewpoints based on explicit information. Reflective tasks at this level may require the reader to perform comparisons, generate explanations or evaluate a feature of the text. Some reflective tasks require readers to demonstrate a detailed understanding of a piece of text dealing with a familiar topic, whereas others require a basic understanding of less-familiar content. Tasks at Level 3 require the reader to take many features into account when comparing, contrasting or categorising information. The required information is often not prominent or there might be a fair amount of competing information. Texts typical of this level may include other obstacles, such as ideas that are contrary to expectation or negatively worded.
	Level 4 At Level 4, readers can comprehend extended passages in single or multiple-text settings. They interpret the meaning of nuances of language in a section of text by taking into account the text as a whole. In other interpretative tasks, students demonstrate understanding and application of ad hoc categories. They can compare perspectives and draw inferences based on multiple sources. Readers can search, locate and integrate several pieces of embedded information in the presence of plausible distractors. They are able to generate inferences based on the task statement in order to assess the relevance of target information. They can handle tasks that require them to memorise prior task context. In addition, students at this level can evaluate the relationship between specific statements and a person's overall stance or conclusion about a topic. They can reflect on the strategies that authors use to convey their

Advanced (625) When reading relatively complex Literary Texts, students can: • Interpret story events and character actions to describe reasons, motivations, feelings, and character development with full text-based support • Begin to evaluate the effect on the reader of the author's language and style choices When reading relatively complex Informational Texts, students can: • Distinguish and interpret complex information from different parts of text, and provide full textbased support • Integrate information across a text to explain relationships and sequence activities • Begin to evaluate visual and textual elements to consider the author's point of view	points, based on salient features of texts such as titles and illustrations. They can compare and contrast claims explicitly made in several texts and assess the reliability of a source based on salient criteria. Texts at Level 4 are often long or complex, and their content or form may not be standard. Many of the tasks are situated in multiple-text settings. The texts and the tasks contain indirect or implicit cues.
	Level 5 Readers at Level 5 can comprehend lengthy texts, inferring which information in the text is relevant even though the information of interest may be easily overlooked. They can perform causal or other forms of reasoning based on a deep understanding of extended pieces of text. They can also answer indirect questions by inferring the relationship between the question and one or several pieces of information distributed within or across multiple texts and sources. Reflective tasks require the production or critical evaluation of hypotheses, drawing on specific information. Readers can establish distinctions between content and purpose, and between fact and opinion as applied to complex or abstract statements. They can assess neutrality and bias based on explicit or implicit cues pertaining to both the content and/or source of the information. They can also draw conclusions regarding the reliability of the claims or conclusions offered in a piece of text. For all aspects of reading, tasks at Level 5 typically involve dealing with concepts that are abstract or counterintuitive, and going through several steps until the goal is reached. In addition, tasks at this level may require the reader to handle several long texts, switching back and forth across texts in order to compare and contrast information.
	Level 6 Readers at Level 6 can comprehend lengthy and abstract texts in which the information of interest is deeply embedded and only indirectly related to the task. They can compare, contrast and integrate information representing multiple and potentially conflicting perspectives, using multiple criteria and generating inferences across distant pieces of information to determine how the information may be used. Readers at Level 6 can reflect deeply on the text's source in relation to its content, using criteria external to the text. They can compare and contrast information across texts, identifying and resolving inter-textual discrepancies and conflicts through inferences about the sources of information, their explicit or vested interests, and other cues as to the validity of the information. Tasks at Level 6 typically require the reader to set up elaborate plans, combining multiple criteria and generating inferences to relate the task and the text(s). Materials at this level include one or several complex and abstract text(s), involving multiple and possibly discrepant perspectives. Target information may take the form of details that are deeply embedded within or across texts and potentially obscured by competing information.
Information retrieved from "PIRLS 2016 International Results in Reading" **(Mullis *et al.* 2017)**	Information retrieved from the "PISA 2018 Technical Report, Chapter 15" **(OECD 2018b)**.

6.1.2 Mathematics Benchmarks and Proficiency Scores in TIMSS and PISA

TIMSS Grade 4 Math Benchmark Summary
Low (400) Students have some basic mathematical knowledge. They can add, subtract, multiply, and divide one- and two-digit whole numbers. They can solve simple word problems. They have some knowledge of simple fractions and common geometric shapes. Students can read and complete simple bar graphs and tables.
Intermediate (475) Students can apply basic mathematical knowledge in simple situations. They can compute with three- and four-digit whole numbers in a variety of situations. They have some understanding of decimals and fractions. Students can identify and draw shapes with simple properties. They can read, label, and interpret information in graphs and tables.
High (550) Students apply conceptual understanding to solve problems. They can apply conceptual understanding of whole numbers to solve two-step word problems. They show understanding of the number line, multiples, factors, and rounding

TIMSS Grade 8 Math Benchmark Summary
Low (400) Students have some knowledge of whole numbers and basic graphs.
Intermediate (475) Students can apply basic mathematical knowledge in a variety of situations. They can solve problems involving whole numbers, negative numbers, fractions, decimals, and ratios. Students have some basic knowledge about properties of two-dimensional shapes. They can read and interpret data in graphs and have some rudimentary knowledge of probability.
High (550) Students can apply their understanding and knowledge in a variety of relatively complex situations. They can solve problems with fractions, decimals, ratios, and proportions. Students at this level show basic procedural knowledge related to algebraic expressions and equations. They can solve a variety of problems with angles, including problems involving triangles, parallel lines, rectangles, and congruent and similar figures. Students can interpret data in a variety of graphs and solve

PISA Math 2012 Proficiency level descriptions
Level 1 At Level 1 students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are almost always obvious and follow immediately from the given stimuli.
Level 2 At Level 2 students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures, or conventions to solve problems involving whole numbers. They are capable of making literal interpretations of the results.
Level 3 At Level 3 students can execute clearly described procedures, including those that require sequential decisions. Their interpretations are sufficiently sound to be a base for building a simple model or for selecting and applying simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They typically show some ability to handle percentages, fractions and decimal numbers, and to work with proportional relationships. Their solutions reflect that they have engaged in basic interpretation and reasoning.
Level 4 At Level 4 students can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic, linking them directly to aspects

numbers, and operations with fractions and decimals. Students can solve simple measurement problems. They demonstrate understanding of geometric properties of shapes and angles. Students can interpret and use data in tables and a variety of graphs to solve problems.	simple problems involving outcomes and probabilities.	of real-world situations. Students at this level can utilise their limited range of skills and can reason with some insight, in straightforward contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments, and actions.
Advanced (625) Students can apply their understanding and knowledge in a variety of relatively complex situations and explain their reasoning. Students can solve a variety of multistep word problems involving whole numbers and show an understanding of fractions and decimals. They can apply knowledge of two- and three-dimensional shapes in a variety of situations. Students can interpret and represent data to solve multistep problems.	**Advanced (625)** Students can apply and reason in a variety of problem situations, solve linear equations, and make generalizations. They can solve a variety of fraction, proportion, and percent problems and justify their conclusions. They can understand linear functions and algebraic expressions. Students can use their knowledge of geometric figures to solve a wide range of problems involving angles, area, and surface area. They can calculate means and medians, and understand how changing data points can impact the mean. Students can interpret a wide variety of data displays to draw and justify conclusions, and solve multistep problems. They can solve problems involving expected values.	**Level 5** At Level 5 students can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare, and evaluate appropriate problem-solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations. They begin to reflect on their work and can formulate and communicate their interpretations and reasoning.
		Level 6 At Level 6, students can conceptualise, generalise and utilise information based on their investigations and modelling of complex problem situations, and can use their knowledge in relatively non-standard contexts. They can link different information sources and representations and flexibly translate among them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply this insight and understanding, along with a mastery of symbolic and formal mathematical operations and relationships, to develop new approaches and strategies for attacking novel situations. Students at this level can reflect on their actions, and can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situation.
Information retrieved from "*TIMSS 2019 International Results in Mathematics and Science"* **(Mullis *et al.* 2020)**.	Information retrieved from "*TIMSS 2019 International Results in Mathematics and Science"* **(Mullis *et al.* 2020)**.	Information retrieved from the "PISA 2012 Technical Report, Chapter 15" **(OECD 2014a)**.

6.1.3 Science Benchmarks and Proficiency Scores in TIMSS and PISA

TIMSS Grade 4 2019 Science Benchmark summary	TIMSS Grade 8 2019 Science Benchmark summary	PISA 2015 Science Proficiency level descriptions
Low (400) Students show limited understanding of scientific concepts and limited knowledge of foundational science facts.	**Low (400)** Students show limited understanding of scientific principles and concepts and limited knowledge of science facts.	**Level 1b** At Level 1b, students can use everyday content knowledge to recognize aspects of simple scientific phenomenon. They are able to identify simple patterns in data, recognize basic scientific terms and follow explicit instructions to carry out a scientific procedure.
Intermediate (475) Students show knowledge and understanding of some aspects of science. Students demonstrate some basic knowledge of plants and animals. They demonstrate knowledge about some properties of matter and some facts related to electricity, and can apply elementary knowledge of forces and motion. They show some understanding of Earth's physical characteristics.	**Intermediate (475)** Students show and apply some knowledge of biology and the physical sciences. Students demonstrate some knowledge of characteristics of animals and apply knowledge of ecosystems. They show some knowledge of the properties of matter, chemical changes, and a few physics concepts.	**Level 1a** At Level 1a, students are able to use everyday content and procedural knowledge to recognize or identify explanations of simple scientific phenomenon. With support, they can undertake structured scientific enquiries with no more than two variables. They are able to identify simple causal or correlational relationships and interpret graphical and visual data that require a low level of cognitive demand. Level 1a students can select the best scientific explanation for given data in familiar personal, local and global contexts.
High (550) Students communicate and apply knowledge of life, physical, and Earth sciences. Students communicate knowledge of characteristics of plants, animals, and their life cycles, and apply knowledge of ecosystems and of humans' and organisms' interactions with their environment. Students demonstrate knowledge of states and properties of matter and of energy transfer in practical	**High (550)** Students apply understanding of concepts from biology, chemistry, physics, and Earth science. Students can apply knowledge of the characteristics of groups of animals, life processes in humans, cells and their functions, genetic inheritance, ecosystems, and nutrition. Students show some knowledge and understanding of the composition and properties of matter and chemical reactions. They can apply basic knowledge of energy transformation and transfer, electrical circuits, properties of magnets, light, sound, and forces. They can apply knowledge of Earth's physical features, processes, cycles, and history, and show some understanding of Earth's resources and their use.	**Level 2** At Level 2, students are able to draw on everyday content knowledge and basic procedural knowledge to identify an appropriate scientific explanation, interpret data, and identify the question being addressed in a simple experimental design. They can use everyday scientific knowledge to identify a valid conclusion from a simple data set. Level 2 students demonstrate basic epistemic knowledge by being able to identify questions that could be investigated scientifically.
		Level 3 At Level 3, students can draw upon moderately complex content knowledge to identify or construct explanations of familiar phenomena. In less familiar or more complex situations, they can construct explanations with relevant cueing or support. They can draw on elements of procedural or epistemic knowledge to carry out a simple experiment in a constrained context. Level 3 students are able to distinguish between scientific and non-scientific issues and identify the evidence supporting a scientific claim.
		Level 4

contexts, and show some understanding of forces and motion. Students know various facts about the Earth's physical characteristics and show basic understanding of the Earth-Moon-Sun system.	**Advanced (625)** Students communicate understanding of concepts related to biology, chemistry, physics, and Earth science in a variety of contexts. Students can classify animals into taxonomic groups. They can apply knowledge of cell structures and their functions. Students show some understanding of diversity, adaptation, and natural selection. They also recognize the interdependence of populations of organisms in an ecosystem. Students demonstrate knowledge of the composition of matter and the periodic table of the elements. Students use physical properties of matter to sort, classify, and compare substances and materials. They also recognize evidence that a chemical reaction has occurred. Students communicate understanding of particle spacing and motion in different physical states. Students apply knowledge of energy transfer and electrical circuits, can relate the properties of light and sound to common phenomena, and demonstrate understanding of forces in everyday contexts. Students communicate understanding of Earth's structure, physical features, and processes. They demonstrate knowledge of the Earth's resources and their conservation.	At Level 4, students can use more sophisticated content knowledge, which is either provided or recalled, to construct explanations of more complex or less familiar events and processes. They can conduct experiments involving two or more independent variables in a constrained context. They are able to justify an experimental design, drawing on elements of procedural and epistemic knowledge. Level 4 students can interpret data drawn from a moderately complex data set or less familiar contexts and draw appropriate conclusions that go beyond the data and provide justifications for their choices.
Advanced (625) Students communicate their understanding of life, physical, and Earth sciences and demonstrate some knowledge of the process of scientific inquiry. Students demonstrate knowledge of characteristics and life processes of a variety of organisms. They can communicate understanding of relationships in ecosystems and interactions between organisms and their environment. They communicate understanding of properties and states of matter and physical and chemical changes. Students communicate understanding of Earth's physical characteristics, processes, and history and show knowledge of Earth's revolution and rotation.		**Level 5** At Level 5, students can use abstract scientific ideas or concepts to explain unfamiliar and more complex phenomena, events and processes. They are able to apply more sophisticated epistemic knowledge to evaluate alternative experimental designs and justify their choices and use theoretical knowledge to interpret information or make predictions. Level 5 students can evaluate ways of exploring a given question scientifically and identify limitations in interpretations of data sets including sources and the effects of uncertainty in scientific data.
		Level 6 At Level 6, students can draw on a range of interrelated scientific ideas and concepts from the physical, life and earth and space sciences and use procedural and epistemic knowledge in order to offer explanatory hypotheses of novel scientific phenomena, events and processes that require multiple steps or to make predictions. In interpreting data and evidence, they are able to discriminate between relevant and irrelevant information and can draw on knowledge external to the normal school curriculum. They can distinguish between arguments that are based on scientific evidence and theory and those based on other considerations. Level 6 students can evaluate competing designs of complex experiments, field studies or simulations and justify their choices.
Information retrieved from "*TIMSS 2019 International Results in Mathematics and Science"* **(Mullis *et al.* 2020)**.	Information retrieved from "*TIMSS 2019 International Results in Mathematics and Science"* **(Mullis *et al.* 2020)**.	Information retrieved from the "PISA 2015 Technical Report Chapter 15" **(OECD 2017b)**.

6.1.4 Proficiency Levels in PIAAC Numeracy and Literacy

PIAAC Numeracy Proficiency level descriptions	PIAAC Literacy Proficiency level descriptions
Below Level 1 Tasks at this level require the respondents to carry out simple processes such as counting, sorting, performing basic arithmetic operations with whole numbers or money, or recognising common spatial representations in concrete, familiar contexts where the mathematical content is explicit with little or no text or distractors.	**Below level 1** The tasks at this level require the respondent to read brief texts on familiar topics to locate a single piece of specific information. There is seldom any competing information in the text and the requested information is identical in form to information in the question or directive. The respondent may be required to locate information in short continuous texts. However, in this case, the information can be located as if the text were non-continuous in format. Only basic vocabulary knowledge is required, and the reader is not required to understand the structure of sentences or paragraphs or make use of other text features. Tasks below Level 1 do not make use of any features specific to digital texts.
Level 1 Tasks at this level require the respondent to carry out basic mathematical processes in common, concrete contexts where the mathematical content is explicit with little text and minimal distractors. Tasks usually require one-step or simple processes involving counting, sorting, performing basic arithmetic operations, understanding simple percents such as 50%, and locating and identifying elements of simple or common graphical or spatial representations.	**Level 1** Most of the tasks at this level require the respondent to read relatively short digital or print continuous, non-continuous, or mixed texts to locate a single piece of information that is identical to or synonymous with the information given in the question or directive. Some tasks, such as those involving non-continuous texts, may require the respondent to enter personal information onto a document. Little, if any, competing information is present. Some tasks may require simple cycling through more than one piece of information. Knowledge and skill in recognising basic vocabulary determining the meaning of sentences, and reading paragraphs of text is expected.
Level 2 Tasks at this level require the respondent to identify and act on mathematical information and ideas embedded in a range of common contexts where the mathematical content is fairly explicit or visual with relatively few distractors. Tasks tend to require the application of two or more steps or processes involving calculation with whole numbers and common decimals, percents and fractions; simple measurement and spatial representation; estimation; and interpretation of relatively simple data and statistics in texts, tables and graphs.	**Level 2** At this level, the medium of texts may be digital or printed, and texts may comprise continuous, non-continuous, or mixed types. Tasks at this level require respondents to make matches between the text and information, and may require paraphrasing or low-level inferences. Some competing pieces of information may be present. Some tasks require the respondent to • cycle through or integrate two or more pieces of information based on criteria; • compare and contrast or reason about information requested in the question; or • navigate within digital texts to access and identify information from various parts of a document
Level 3 Tasks at this level require the respondent to understand mathematical information that may be less explicit, embedded in contexts that are not always familiar and represented in more complex ways. Tasks require several steps and may involve the choice of problem-solving strategies and relevant processes. Tasks tend to require the application of number	**Level 3** Texts at this level are often dense or lengthy, and include continuous, non-continuous, mixed, or multiple pages of text. Understanding text and rhetorical structures become more central to successfully completing tasks, especially navigating complex digital

sense and spatial sense; recognising and working with mathematical relationships, patterns, and proportions expressed in verbal or numerical form; and interpretation and basic analysis of data and statistics in texts, tables and graphs.	texts. Tasks require the respondent to identify, interpret, or evaluate one or more pieces of information, and often require varying levels of inference. Many tasks require the respondent to construct meaning across larger chunks of text or perform multi-step operations in order to identify and formulate responses. Often tasks also demand that the respondent disregard irrelevant or inappropriate content to answer accurately. Competing information is often present, but it is not more prominent than the correct information.
Level 4 Tasks at this level require the respondent to understand a broad range of mathematical information that may be complex, abstract or embedded in unfamiliar contexts. These tasks involve undertaking multiple steps and choosing relevant problemsolving strategies and processes. Tasks tend to require analysis and more complex reasoning about quantities and data; statistics and chance; spatial relationships; and change, proportions and formulas. Tasks at this level may also require understanding arguments or communicating well-reasoned explanations for answers or choices.	**Level 4** Tasks at this level often require respondents to perform multiple-step operations to integrate, interpret, or synthesise information from complex or lengthy continuous, non-continuous, mixed, or multiple type texts. Complex inferences and application of background knowledge may be needed to perform the task successfully. Many tasks require identifying and understanding one or more specific, non-central idea(s) in the text in order to interpret or evaluate subtle evidence-claim or persuasive discourse relationships. Conditional information is frequently present in tasks at this level and must be taken into consideration by the respondent. Competing information is present and sometimes seemingly as prominent as correct information.
Level 5 Tasks at this level require the respondent to understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts. Respondents may have to integrate multiple types of mathematical information where considerable translation or interpretation is required; draw inferences; develop or work with mathematical arguments or models; and justify, evaluate and critically reflect upon solutions or choices.	**Level 5** At this level, tasks may require the respondent to search for and integrate information across multiple, dense texts; construct syntheses of similar and contrasting ideas or points of view; or evaluate evidence based arguments. Application and evaluation of logical and conceptual models of ideas may be required to accomplish tasks. Evaluating reliability of evidentiary sources and selecting key information is frequently a requirement. Tasks often require respondents to be aware of subtle, rhetorical cues and to make high-level inferences or use specialised background knowledge.
Information retrieved from the "OECD Skills Outlook 2013 First Results from the Survey of Adult Skills" **(OECD 2013d)**.	Information retrieved from the "OECD Skills Outlook 2013 First Results from the Survey of Adult Skills" **(OECD 2013d)**.

6.2 Tables for Educational Track/Programme Harmonisation

The sections are organised first by country then by dataset. The datasets included in the tables below are, PISA[18], GUI '98[19], TREE2[20], NEPS-SC4[21], ÉpStan[22]. In each table, the first column(s) identify the different national programs represented in the data (including the value scheme as coded in the given data set), followed by a column that identifies the official OECD definition for the ISCED level, designation and orientation (as reported in PISA) and the last column provides our proposed harmonisation, which has been approved by the respective national partners, for the national program.

6.2.1 Germany

An important clarification for the PISA 2018-2006 national programs in Germany: basic general education refers to a Hauptschule, extensive general education refers to a Realschule and academic education refers to Gymnasium.

Table 41: PISA national programs of Germany

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
02760001 Germany: Lower secondary, some with access to upper secondary (special education)	02760001 Germany: Lower secondary, some with access to upper secondary (special education)	02760001 Germany: Lower secondary with access to upper secondary (comprehensive)	2760001 DEU: Lower secondary with access to upper secondary (comprehensive)	2760001 DEU: LOWER SECONDARY WITH ACCESS TO UPPER SECONDARY (COMPREHENSIVE)	276001 DEU: Lower sec. access to upper sec. (compr., special educ.)	2A General	4: Other
02760002 Germany: Lower secondary, some with access to upper secondary; basic general education	02760002 Germany: Lower secondary, some with access to upper secondary; basic gen educ (exclusively students of the same track	02760002 Germany: Lower secondary without access to upper secondary, providing a basic general education (grade 5 - 10)	2760002 DEU: Lower secondary without access to upper secondary, providing a basic genera	2760002 DEU: LOWER SEC., NO ACCESS TO UPPER SECONDARY (HAUPTSCHULE)	276002 DEU: Lower sec. no access to upper sec. (Hauptschule)	2A General	3: Vocational

[18] PISA codebooks: (OECD 2001b, 2001a, 2004b, 2004a, 2007b, 2007a, 2010c, 2010b, 2013c, 2013b, 2016, 2019e)
[19] Growing Up in Ireland codebooks: (Growing Up in Ireland n.d., n.d., n.d., n.d.)
[20] Transistions from Education to Employment 2 codebooks: (TREE 2016b, 2017, 2018)
[21] NEPS Starting Cohort 4 codebook: (National Education Panel Study (NEPS) 2020)
[22] No available codebook for ÉpStan, information about the data was provided by the Luxembourg Center for Education Testing (LUCET).

02760003 Germany: Lower secondary, access to upper secondary; extensive general education	02760003 Germany: Lower secondary, access to upper secondary; extensive gen educ (exclusively students of the same track)	02760003 Germany: Lower secondary without access to upper secondary, providing an extensive general education (grades 5	2760003 DEU: Lower secondary without access to upper secondary, providing an extensive g	2760003 DEU: LOWER SEC., NO ACCESS TO UPPER SECONDARY (REALSCHULE)	276003 DEU: Lower sec. no access to upper sec. (Realschule)	2A General	2: Technical
02760004 Germany: Lower secondary, access to upper secondary; academic education (exclusively students of the same track)	02760004 Germany: Lower secondary, access to upper secondary; academic education (exclusively students of the same track)	02760004 Germany: Lower secondary with access to upper secondary	2760004 DEU: Lower secondary with access to upper secondary	2760004 DEU: LOWER SEC. WITH ACCESS TO UPPER SECONDARY (GYMNASIUM)	276004 DEU: Lower sec. access to upper sec. (Gymnasium)	2A General	1: General and Comprehensive
02760005 Germany: Upper secondary (exclusively students of the same track [cf. key 4])	02760005 Germany: Upper secondary (exclusively students of the same track [cf. key 4])	02760005 Germany: Upper secondary level of education	2760005 DEU: Upper secondary level of education	2760005 DEU: UPPER SEC. LEVEL (GYMNASIUM)	276017 DEU: Upper sec. (Gymnasium)	3A General	1: General and Comprehensive
02760006 Germany: Lower secondary comprehensive, achievement-based access to upper secondary (within school streaming)	02760006 Germany: Lower secondary comprehensive, achievement-based access to upper secondary (within school streaming)	02760006 Germany: Lower secondary with access to upper secondary (comprehensive)	2760006 DEU: Lower secondary with access to upper secondary (comprehensive)	2760006 DEU: COMPREHENSIVE LOWER SEC. WITH ACCESS TO UPPER SEC. (GESAMTSCHULE)	276005 DEU: Lower sec. access to upper sec. (comprehensive)	2A General	1: General and Comprehensive
02760007 Germany: Upper secondary comprehensive	02760007 Germany: Upper secondary comprehensive (taught together with students of other tracks at lower secondary [cf. key 6])	/	2760007 DEU: Upper secondary level of education	/	276018 DEU: Upper sec. (comprehensive)	3A General	1: General and Comprehensive
02760008 Germany: Lower secondary comprehensive, no access to upper; basic general education (different tracks taught separately)	02760008 Germany: Lower secondary comprehensive, no access to upper; basic general education (different tracks taught separate	/	2760008 DEU: Lower secondary without access to upper secondary, providing a basic genera	2760008 DEU: LOWER SEC., NO ACCESS TO UPPER SEC. (KOOP. GESAMTSCHULE HS)	276006 DEU: Lower sec. no access to upper sec. (Koop. Gesamtschule)	2A General	3: Vocational

02760009 Germany: Lower secondary comprehensive, access to upper; extensive general education	02760009 Germany: Lower secondary comprehensive, access to upper; extensive gen educ (different tracks taught separately)	/	2760009 DEU: Lower secondary with/without access to upper secondary, providing an extens	2760009 DEU: LOWER SEC., WITH OR WITHOUT ACCESS TO UPPER SEC. (KOOP. GS, RS)	/	2A General	2: Technical
02760010 Germany: Lower secondary comprehensive, access to upper secondary; academic education	02760010 Germany: Lower secondary comprehensive, access to upper secondary; academic educ (different tracks taught separately)	/	2760010 DEU: Lower secondary with access to upper secondary (comprehensive)	2760010 DEU: LOWER SEC., WITH ACCESS TO UPPER SEC. (KOOP. GS, GYMN.)	276013 DEU: Lower sec. with access to upper sec. (comprehensive)	2A General	1: General and Comprehensive
02760011 Germany: Upper secondary comprehensive (after lower secondary with different tracks taught separately [cf. key 10])	02760011 Germany: Upper secondary comprehensive (after lower secondary with different tracks taught separately [cf. key 10])	/	2760011 DEU: Upper secondary level of education	/	/	3A General	1: General and Comprehensive
02760012 Germany: Lower secondary, expectedly no access to upper; basic general education	02760012 Germany: Lower secondary, expectedly no access to upper; basic gen educ (students of different tracks taught together	02760012 Germany: Lower secondary without access to upper secondary, providing a basic general education (grades 5	2760012 DEU: Lower secondary without access to upper secondary, providing a basic genera	2760012 DEU: LOWER SEC., NO ACCESS TO UPPER SECONDARY (HAUPTSCHULE INTEGRATED)	276009 DEU: Lower sec. no access to upper sec.	2A General	3: Vocational
02760013 Germany: Lower secondary, expectedly access to upper; extensive general education	02760013 Germany: Lower secondary, expectedly access to upper; extensive gen educ (students of different tracks taught togethe	02760013 Germany: Lower secondary without access to upper secondary, providing an extensive general education (grades 5	2760013 DEU: Lower secondary without access to upper secondary, providing an extensive g	2760013 DEU: LOWER SEC., NO ACCESS TO UPPER SECONDARY (REALSCHULE INTEGRATED)	276010 DEU: Lower sec. no access to upper sec.	2A General	2: Technical
02760014 Germany: Lower secondary, no access to upper; basic general education (students of different	02760014 Germany: Lower secondary, no access to upper; basic general education (students of different	02760014 Germany: Lower secondary without access to upper secondary, providing a basic	2760014 DEU: Lower secondary without access to upper secondary, providing a basic genera	2760014 DEU: LOWER SEC., NO ACCESS TO UPPER SECONDARY (HAUPTSCHULKLASSE)	276011 DEU: Lower sec. no access to upper sec.	2A General	3: Vocational

tracks taught separately)	tracks taught separately	general education (grades 5					
02760015 Germany: Lower secondary, access to upper; extensive general education (students of different tracks taught separately)	02760015 Germany: Lower secondary, access to upper; extensive general education (students of different tracks taught separatel	02760015 Germany: Lower secondary without access to upper secondary, providing an extensive general education (grades 5	2760015 DEU: Lower secondary without access to upper secondary, providing an extensive g	2760015 DEU: LOWER SEC., NO ACCESS TO UPPER SECONDARY (REALSCHULKLASSE)	276012 DEU: Lower sec. no access to upper sec.	2A General	2: Technical
02760016 Germany: Lower secondary with access to upper (Waldorf school)	02760016 Germany: Lower secondary with access to upper (Waldorf school)	/	2760016 DEU: Lower secondary with access to upper secondary (comprehensive)	2760016 DEU: LOWER SECONDARY WITH ACCESS TO UPPER SECONDARY (WALDORF)	/	2A General	1: General and Comprehensive
/	/	/	/	2760017 DEU: UPPER SECONDARY LEVEL OF EDUCATION (WALDORF)	/	3A Genral	1: General and Comprehensive
02760018 Germany: Pre-vocational training year upper secondary level	02760018 Germany: Pre-vocational training year upper secondary level	02760018 Germany: Pre-vocational training year upper secondary level	2760018 DEU: pre-vocational training year upper secondary level	2760018 DEU: PRE-VOCATIONAL TRAINING YEAR	276014 DEU: pre-vocational training year	2A Pre-vocational (2018-2009, 2003) 2A General (2006)	2: Technical
02760019 Germany: Vocational school upper secondary level	02760019 Germany: Vocational school upper secondary level	02760019 Germany: Vocational school upper secondary level	2760019 DEU: Vocational school upper secondary level	2760019 DEU: VOCATIONAL SCHOOL (BERUFSSCHULE)	276015 DEU: Vocational school (Berufsschule)	3B Vocational	2: Technical
02760020 Germany: Vocational school	02760020 Germany: Vocational school	02760020 Germany: Vocational school	2760020 DEU: Vocational school	2760020 DEU: VOCATIONAL SCHOOL (BERUFSFACHSCHULE)	276016 DEU: Vocational school (Berufsfachschule)	3B Vocational	2: Technical
02760021 Germany: Upper secondary (vocational), qualifying for subject-specific tertiary education	/	/	/	/	/	3A General	1: General and Comprehensive

Table 42: NEPS-SC4 national programs of Germany (for waves 1-3, ISCED level 2)

School type (t723080_g1) waves 1,2,3 (grades 9-10)	German Labels (t723080_g1)	Follow-up school type (tx80232) wave 3 for students who were followed up with individually	Descriptive notes about the program	ISCED level, designation, orientation	Proposed harmonisation
1: Elementary school;	Grundschule	1: Elementary school, primary stage	Elementary	1 (and 2A General in some German states)[23]	1: General and Comprehensive
2: Orientation stage;	Orientierungsstufe	2: Orientation stage, trial and remedial stage	Lower-secondary, before track is chosen	2A General	1: General and Comprehensive
3: Hauptschule;	Hauptschule	4: Hauptschule	Low level of lower-secodnary	2A General	3: Vocational
4: Realschule;	Realschule	5: Realschule	Intermediate level of lower-secondary	2A General	2: Technical
5: School with multiple educational programs: No separation in school branches (yet);	Schule mit mehreren Bildungsgängen: (noch) keine Trennung in Schulzweige	/	Comprehensive school	2A General	1: General and Comprehensive
6: School with multiple educational programs: School branch of the Hauptschule;	Schule mit mehreren Bildungsgängen: Hauptschulzweig	/	Comprehensive school with tracking (low level of lower-secodnary)	2A General	3: Vocational
7: School with multiple educational programs: School branch of the Realschule;	Schule mit mehreren Bildungsgängen: Realschulzweig	/	Comprehensive school with tracking (intermediate level of lower-secodnary)	2A General	2: Technical
8: School with multiple educational programs: Unclear;	Schule mit mehreren Bildungsgängen: unklar	6: (integrated) secondary school, regular school, Mittelschule, Oberschule and Wirtschaftsschule, Regionale Schule, extended Realschule, Realschule plus, Gemeinschaftsschule, Werkrealschule, Stadtteilschule, Mittelstufenschule, verbundene Haupt- and Realschule, Aufbaugymnasium	Comprehensive school	2A General	1: General and Comprehensive
9: Gymnasium;	Gymnasium	7: Gymnasium	Upper level of lower-secondary	2A General	1: General and Comprehensive
10: Special needs school;	Förderschule	9: Special needs school or special needs center	Special needs school	2A General	4: Other

[23] In some German states, the elementary school includes additional grades similar to the orientation stage. We propose the harmonisation code of "1: General and Comprehensive" as to not systematically exclude students attending these prolonged elementary schools.

11: Comprehensive school: no separation of school branches (yet)	Gesamtschule: (noch) keine Trennung in Schulzweige	/	Comprehensive school	2A General	1: General and Comprehensive
12: Comprehensive school: school branch of the Hauptschule;	Gesamtschule: Hauptschulzweig	/	Comprehensive school with tracking (low level of lower-secodnary)	2A General	3: Vocational
13: Comprehensive school: school branch of the Realschule;	Gesamtschule: Realschulzweig	/	Comprehensive school with tracking (intermediate level of lower-secodnary)	2A General	2: Technical
14: Comprehensive school: school branch of the Gymnasium;	Gesamtschule: Gymnasialzweig	/	Comprehensive school	2A General	1: General and Comprehensive
15: Comprehensive school: unclear;	Gesamtschule: unklar	3: Comprehensive school	Comprehensive school	2A General	1: General and Comprehensive
16: Waldorf school;	Waldorfschule	8: Waldorf school	Waldorf school, specialised teaching style	2A General	1: General and Comprehensive
17: Other school	andere Schule	10: Other school		N/A	4: Other

Table 43: NEPS-SC4 national programs of Germany (for waves 5-8, ISCED level 3)

Kind of school (te00002_v1) waves 5, 7, 8 (grades 11-13)	**German label (te00002_v1)**	**Follow-up school type (tx80233) waves 5 and 7 for students who were followed up with individually**	**Descriptive notes about the program**	**ISCED level, designation, orientation**	**Proposed harmonisation**
1: Gymnasium	Gymnasium	1: Gymnasium	Upper secondary general	3A General	1: General and Comprehensive
2: Comprehensive school	Gesamtschule	2: Comprehensive school	Comprehensive school	3A General	1: General and Comprehensive
3: Higher vocational school	Höhere Berufsfachschule	3: Higher vocational school	Vocational school	3B Vocational	2: Technical
4: Berufsfachschule	Berufsfachschule	4: Berufsfachschule	Vocational school	3B Vocational	2: Technical
5: Vocational college	Berufskolleg	5: Vocational college	Vocational college	3B Vocational	2: Technical
6: Berufsoberschule (BOS, BOB)	Berufsoberschule (BOS, BOB)	6: Berufsoberschule (BOS, BOB)	Entry requirement: vocational certificate and completed grade 10 (grants access to 5A)	4A General	N/A
7: Fachakademie	Fachakademie	7: Fachakademie	Post-secondary vocational training	5B	N/A

8: Fachoberschule (FOS)	Fachoberschule (FOS)	8: Fachoberschule (FOS)	Specialized vocational high school, 2 years (grants access to 5A)	3A General	1: General and Comprehensive
9: Fachschule	Fachschule	9: Fachschule	Post-secondary vocational training	5B	N/A (ISCED level 5)
10: Gemeinschaftsschule	Gemeinschaftsschule	10: Gemeinschaftsschule	Comprehensive school	3A General	1: General and Comprehensive
11: Higher commercial school	Höhere Handelsschule	11: Higher commercial school	Vocational college	3B Vocational	2: Technical
12: Integrated secondary school	Integrierte Sekundarschule	12: Integrated secondary school	Comprehensive school	3A General	1: General and Comprehensive
13: Oberschule	Oberschule	13: Oberschule	Comprehensive school	3A General	1: General and Comprehensive
14: Stadtteilschule (not Waldorf school)	Stadtteilschule (keine Waldorfschule)	14: Stadtteilschule (not Waldorf school)	Comprehensive school	3A General	1: General and Comprehensive
15: Waldorf school	Waldorfschule	15: Waldorf school	Waldorf school, specialised teaching style	3A General	1: General and Comprehensive
16: Other kind of school	Sonstige Schulform	16: Other kind of school	Other	N/A	4: Other

6.2.2 Hungary

Table 44: PISA national programs of Hungary

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
03480001 Hungary: Primary school	03480001 Hungary: Primary school	03480001 Hungary: Primary school	3480001 HUN: Primary school	3480001 HUN: PRIMARY SCHOOL	348001 HUN: Primary school	2A General	1: General and Comprehensive
03480003 Hungary: Language Preparatory 9th grade (9Ny grade)	/	/	/	/	/	3A General	1: General and Comprehensive
03480004 Hungary: Grammar school 1 (4 years long)	03480002 Hungary: Grammar school 1 (4-5 years long)	03480002 Hungary: Grammar school 1 (4-5 years long)	03480002 HUN: Grammar school	3480004 HUN: GRAMMAR SCHOOL	348002 HUN: Grammar school	3A General	1: General and Comprehensive
03480005 Hungary: Grammar school 1 after Language Preparatory (4 years long)	/	/	/	/	/	3A General	1: General and Comprehensive
03480006 Hungary: Grammar school 2 (6 years long, grade 7-8)	/	/	/	/	/	2A General	1: General and Comprehensive
03480007 Hungary: Grammar school 2 (6 years long, grade 9-12)	/	/	/	/	/	3A General	1: General and Comprehensive
03480008 Hungary: Grammar school 2 after Language Preparatory (6 years long, grade 7-8)	03480003 Hungary: Grammar school 2 (6 years long, ISCED2)	03480003 Hungary: Grammar school 2 (6 years long, ISCED2)	/	/	/	2A General	1: General and Comprehensive
03480009 Hungary: Grammar school 2 after Language Preparatory (6 years long, grade 9-12)	03480004 Hungary: Grammar school 2 (6 years long, ISCED3)	03480004 Hungary: Grammar school 2 (6 years long, ISCED3)	/	/	/	3A General	1: General and Comprehensive
03480010 Hungary: Grammar school 3 (8 years long, grade 7-8)	03480005 Hungary: Grammar school 3 (8 years long, ISCED2)	03480005 Hungary: Grammar school 3 (8 years long, ISCED2)	/	/	/	2A General	1: General and Comprehensive
03480011 Hungary: Grammar school 3 (8 years long, grade 9-12)	03480006 Hungary: Grammar school 3 (8 years long, ISCED3)	03480006 Hungary: Grammar school 3 (8 years long, ISCED3)	/	/		3A General	1: General and Comprehensive
03480014 Hungary: Vocational secondary school	03480007 Hungary: Vocational secondary school	03480007 Hungary: Vocational secondary school	03480003 HUN: Vocational secondary school	3480003 HUN: VOCATIONAL SECONDARY SCHOOL	348003 HUN: Vocational sec. school	3A General (2018-2009) 3A Pre-vocational (2006-2003)	1: General and Comprehensive

03480015 Hungary: Vocational secondary school after Language Preparatory	/	/	/	/	/	3A General	1: General and Comprehensive
03480016 Hungary: Vocational school	03480008 Hungary: Vocational school	03480008 Hungary: Vocational school	03480004 HUN: Vocational school	3480002 HUN: VOCATIONAL SCHOOL	348004 HUN: Vocational school	3C Vocational	3: Vocational

6.2.3 Ireland

Table 45: PISA national programs of Ireland

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
03720001 Ireland: Junior Cert.	03720001 Ireland: Junior Cert.	03720001 Ireland: Junior Cert.	3720001 Ireland: Junior Cert.	3720001 Ireland: Junior Cert.	372001 Ireland: Junior Cert.	2A General	1: General and Comprehensive
03720002 Ireland: Transition Year Programme	03720002 Ireland: Transition Year Programme	03720002 Ireland: Transition Year Programme	3720002 Ireland: Transition Year Programme	3720002 Ireland: Transition Year Programme	372002 Ireland: Transition Year Programme	3C General	3: Vocational
03720003 Ireland: Leaving Cert. Applied	03720003 Ireland: Leaving Cert. Applied	03720003 Ireland: Leaving Cert. Applied	3720003 Ireland: Leaving Cert. Applied	3720003 Ireland: Leaving Cert. Applied	372003 Ireland: Leaving Cert. Applied	3C Pre-vocational (2018-2006) 3A General (2003)	2: Technical
03720004 Ireland: Leaving Cert. Established	03720004 Ireland: Leaving Cert. Established	03720004 Ireland: Leaving Cert. Established	3720004 Ireland: Leaving Cert. Established	3720004 Ireland: Leaving Cert. Established	372004 Ireland: Leaving Cert. General	3A General (2018-2006) 3B Pre-vocational (2003)	1: General and Comprehensive
03720005 Ireland: Leaving Cert. Vocational	03720005 Ireland: Leaving Cert. Vocational	03720005 Ireland: Leaving Cert. Vocational	3720005 Ireland: Leaving Cert. Vocational	3720005 Ireland: Leaving Cert. Vocational	372005 Ireland: Leaving Cert. Vocational	3A General	1: General and Comprehensive
/	/	/	3720006 IRL: Youthreach	/	/	2C Pre-vocataional	3: Vocational

Table 46: GUI '98 national programs of Ireland (Waves 3-4)

School type (wave 3: cq3b5a/ wave 4: cq4F3)	Descriptive notes about the program	ISCED level, designation, orientation	Proposed harmonisation
1: Regular (Established) Leaving Certificate	Upper secondary general, gives acces to tertiary education	3A General	1: General and Comprehensive
2: Leaving Certificate Applied (LCA)	Upper secondary pre-vocational, gives access to post-secondary non-tertiary courses	3C Pre-vocational	2: Technical
3: Leaving Certificate Vocational (LCVP)	Upper secondary general, gives acces to tertiary education	3A General	1: General and Comprehensive
4: Something else (please specify)	Other	N/A	4: Other

6.2.4 Lithuania

Table 47: PISA national programs of Lithuania

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	PISA 2003 (Not Available)	ISCED level, designation, orientation	Proposed harmonisation
04400001 Lithuania: Basic School	04400001 Lithuania: Basic School	04400001 Lithuania: Basic School	4400001 LTU: Basic School	4400001 LTU: GENERAL BASIC EDUCATION	/	2A General	1: General and Comprehensive
04400002 Lithuania: Progymnasium	04400004 Lithuania: Progymnasium	04400007 Lithuania: Progymnasium	4400007 LTU: Progymnasium	/	/	2A General	1: General and Comprehensive
04400003 Lithuania: Gymnasium (grades 7-8)	04400002 Lithuania: Secondary School (Lower Secondary)	04400002 Lithuania: Secondary School (Lower Secondary)	4400002 LTU: Secondary School (Lower Secondary)	/	/	2A General	1: General and Comprehensive
/	04400003 Lithuania: Secondary School (Upper Secondary)	/	/	/	/	3A General	1: General and Comprehensive
04400004 Lithuania: Lower Gymnasium	04400005 Lithuania: Lower Gymnasium	04400004 Lithuania: Lower Gymnasium	4400004 LTU: Lower Gymnasium	4400003 LTU: BASIC EDUCATION (LOWER GYMNASIUM)	/	2A General	1: General and Comprehensive
/	04400006 Lithuania: Upper Gymnasium	04400005 Lithuania: Upper Gymnasium	4400005 LTU: Upper Gymnasium	4400004 LTU: SECONDARY EDUCATION (UPPER GYMNASIUM)	/	3A General	1: General and Comprehensive
04400006 Lithuania: Vocational School	04400007 Lithuania: Vocational School	04400006 Lithuania: Vocational School	4400006 LTU: Vocational School	4400005 LTU: BASIC AND VOCATIONAL EDUCATION	/	2A Vocational	3: Vocational
/	/	/	/	4400006 LTU: VOCATIONAL EDUCATION	/	2C Vocational	3: Vocational

6.2.5 Luxembourg

Table 48: PISA national programs of Luxembourg

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
04420001 Luxembourg: Preparatory courses within modular classes of the Preparatory Regime (lower secondary)	04420001 Luxembourg: Preparatory courses within modular classes of the Preparatory Regime	04420001 Luxembourg: Lower secondary education	4420001 LUX: Lower secondary education	4420001 LOWER SECONDARY EDUCATION (EST: PREPARATOIRE)	442001 Year 7 or 8 or 9	2A General	3: Vocational
04420002 Luxembourg: Lower technical secondary education	04420002 Luxembourg: Lower technical secondary education	04420002 Luxembourg: Lower secondary education	4420002 LUX: Lower secondary education	4420002 LOWER SECONDARY EDUCATION (EST: INFERIEUR)	442002 Year 7 or 8 or 9	2A General	2: Technical
04420003 Luxembourg: Lower general secondary education	04420003 Luxembourg: Lower general secondary education	04420003 Luxembourg: Lower secondary education	4420003 LUX: Lower secondary education	4420003 LOWER SECONDARY EDUCATION (ES: INFERIEUR)	/	2A General	1: General and Comprehensive
04420004 Luxembourg: Vocational regime: a 3-year vocational education	04420004 Luxembourg: A 3-year vocational education	04420004 Luxembourg: A 3-year vocational education	4420004 LUX: A 3-year vocational education	4420004 A 3-YEAR VOCATIONAL EDUCATION (EST: PROF.)	442003 Year 10 or 11, with mostly VET (vocational) subjects	3C Vocational	3: Vocational
04420005 Luxembourg: Technician's regime: a 4-year technical education	04420005 Luxembourg: A 4-year technical education	04420005 Luxembourg: A 4-year technical education	4420005 LUX: A 4-year technical education	4420005 A 4-YEAR VOCATIONAL-TECHNICAL EDUCATION (EST)	442004 Year 10-12, in a program leading to an apprenticeship	3B Vocational	2: Technical
04420006 Luxembourg: Technical regime: a 4 to 5-year technical, commercial or social education	04420006 Luxembourg: A 4 to 5-year technical, commercial or social education	04420006 Luxembourg: A 4 to 5-year technical, commercial or social education	4420006 LUX: A 4-year technical education	4420006 A 4 TO 5-YEAR TECHNICAL EDUCATION (EST)	442005 Year 10-12 in a program leading to higher education	3A Vocational	2: Technical
04420007 Luxembourg: Upper general secondary education	04420007 Luxembourg: Upper general secondary education	04420007 Luxembourg: Upper secondary education	4420007 LUX: Upper secondary education	4420007 UPPER SECONDARY EDUCATION (ES: SUPERIEUR)	442006 Year 10-12 in a program leading to university	3A General	1: General and Comprehensive
04420008 Luxembourg: Lower secondary (private and/or international) schools which do not follow the official Ministry programme	04420008 Luxembourg: Lower secondary education, international and foreign schools	04420008 Luxembourg: Lower secondary private, not subsidized	4420008 LUX: Lower secondary private, not subsidized	4420008 LOWER SECONDARY PRIVATE, NOT SUBSIDIZED	/	2A General	1: General and Comprehensive

04420009 Luxembourg: Middle & upper secondary (private and/or int'l) schools which don't follow the official Ministry programme	04420009 Luxembourg: Upper secondary education, international and foreign schools	04420009 Luxembourg: Upper secondary private, not subsidized	4420009 LUX: Upper secondary private, not subsidized	4420009 UPPER SECONDARY PRIVATE, NOT SUBSIDIZED	/	3A General	1: General and Comprehensive

Table 49: EpStan (Grade 9) national programs of Luxembourg

Track (xtyp_ens): ÉpStan 2019, Grade 9	Descriptive notes about the program	ISCED level, designation, orientation	Proposed harmonisation
“ES” Classical track	Upper level lower-secondary	2A General	1: General and Comprehensive
“EST” technical track	Intermediate level lower-secondary	2A General	2: Technical
“PRE” preparatory track	Low level lower-secondary	2A General	3: Vocational

6.2.6 Switzerland

Table 50: PISA national programs of Switzerland

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
07560001 Switzerland: Secondary education, First stage	07560001 Switzerland: Secondary education, First stage	07560001 Switzerland: Secondary education, First stage	7560001 CHE: Secondary education, First stage	7560001 CHE: SECONDARY EDUCATION, FIRST	756001 CHE: sec. education, 1st stage	2A General	1: General and Comprehensive
07560002 Switzerland: Preparatory course for vocational education	07560002 Switzerland: Preparatory course for vocational education	07560002 Switzerland: Preparatory course for vocational education	7560002 CHE: Preparatory course for vocational education	7560002 CHE: PREPARATORY COURSE FOR VOCATIONAL EDUCATION	756002 CHE: Preparatory course for vocational education	3A General (2018) 2A General (2015- 2003)	1: General and Comprehensive
07560003 Switzerland: School preparing for the university entrance certificate	07560003 Switzerland: School preparing for the university entrance certificate	07560003 Switzerland: School preparing for the university entrance certificate	7560003 CHE: School preparing for the university entrance	7560003 CHE: SCHOOL PREPARING FOR THE UNIVERSITY ENTRANCE CERTIFICATE	756003 CHE: School prep. for the university entrance certificate	3A General	1: General and Comprehensive
07560004 Switzerland: Vocational baccalaureat, dual system 3-4 years	07560004 Switzerland: Vocational baccalaureat, dual system 3-4 years	07560004 Switzerland: Vocational baccalaureat, dual system 3-4 years	7560004 CHE: Vocational baccalaureat, dual system 3-4 years	STAGE 7560004 CHE: VOCATIONAL BACCALAUREAT, DUAL SYSTEM 3-4 YEARS	756004 CHE: Vocational baccalaureat, dual system 3-4 years	3A Vocational	1: General and Comprehensive
07560005 Switzerland: Vocational education, dual system 3-4 years	07560005 Switzerland: Vocational education, dual system 3-4 years	07560005 Switzerland: Vocational education, dual system 3-4 years	certificate 7560005 CHE: Vocational education, dual system 3-4 years	7560005 CHE: VOCATIONAL EDUCATION, DUAL SYSTEM 3-4 YEARS	756005 CHE: Vocational education, dual system 3-4 years	3B Vocational	3: Vocational
07560006 Switzerland: Intermediate diploma school	07560006 Switzerland: Intermediate diploma school	07560006 Switzerland: Intermediate diploma school	7560006 CHE: Intermediate diploma school	7560006 CHE: INTERMEDIATE DIPLOMA SCHOOL	756006 CHE: Intermediate diploma school 3-4 years	3B General	3: Vocational
07560007 Switzerland: Basic vocational education (2 years)	07560007 Switzerland: Basic vocational education (2 years)	07560007 Switzerland: Basic vocational education, dual system 1-2 years	7560007 CHE: Basic vocational education, dual system 1-2 years	7560007 CHE: BASIC VOCATIONAL EDUCATION, DUAL SYSTEM 1-2 YEARS	756007 CHE: Basic vocational education, dual system 1-2 years	3C Vocational	2: Technical

/	/	/	/	/	756008 CHE: Intermediate Diploma School	3C General	2: Technical

Table 51: TREE2 national programs of Switzerland

Type of education class (educ_class*): TREE2 (Waves 2-3)	Descriptive notes about the program	ISCED level, designation, orientation	Proposed harmonisation
10: Pre-apprenticeship	Interim solutions	N/A	N/A
11: 10th school year (12th year HarmoS) / one-year preparatory programme for choice of occupation / VET, integration year (for recently immigrated youths with insufficient command of the native language), preparation for upper-secondary school etc.	Interim solutions	N/A	N/A
12: 10th language school year / school exchange year (in a foreign-language region)	Interim solutions	N/A	N/A
13: Au pair / language stay	Interim solutions	N/A	N/A
14: Housekeeping training year	Interim solutions	N/A	N/A
15: Social year (Sozialjahr)	Interim solutions	N/A	N/A
16: Motivation semester	Interim solutions	N/A	N/A
17: Preparatory course	Interim solutions	N/A	N/A
29: Other bridge-year programmes / interim solutions	Interim solutions	N/A	N/A
95: Other upper-secondary education	Non-Certifying Education	N/A	N/A
1040: Language course	Non-Certifying Education	N/A	N/A
9000: Other education / further education and training / course [TEXT]	Non-Certifying Education	N/A	N/A
32: 2-year firm-based vocational education and training (training contract with a firm or training network)	Short-term VET (EBA)	3B Vocational	2: Technical
33: 2-year full-time school-based vocational education and training at a vocational school / vocational training centre	Short-term VET (EBA)	3B Vocational	2: Technical
30: 3–4-year firm-based vocational education and training (VET diploma with / without FVB) (training contract with a firm or training network)	Regular VET (EFZ)	3C Vocational	3: Vocational
31: 3–4-year full-time school-based vocational education and training (VET diploma with / without FVB) at (another) vocational school / vocational training centre	Regular VET (EFZ)	3C Vocational	3: Vocational
34: Upper-secondary commercial / business school (VET diploma with / without FVB)	Vocational Schools, leading to Vocational Baccalaureate	3A General	1: General and Comprehensive
35: Upper-secondary IT school (VET diploma with / without FVB)	Vocational Schools, leading to Vocational Baccalaureate	3A General	1: General and Comprehensive

40: Upper-secondary specialised school - specialised school diploma	Upper Secondary Specialised Schools, leading to Baccalaureate (General)	3A General	1: General and Comprehensive
42: German-speaking Switzerland: Integrative upper-secondary school Rudolf Steiner (Waldorf school)	Upper Secondary Specialised Schools, leading to Baccalaureate (General)	3A General	1: General and Comprehensive
43: Ticino: Cantonal baccalaureate in commerce and business + federal VET diploma (cantonal high school of commerce)	Upper Secondary Specialised Schools, leading to Baccalaureate (General)	3A General	1: General and Comprehensive
39: Baccalaureate (high-school, college, etc.);	Baccalaureate School, Upper Secondary Specialised Schools, leading to Baccalaureate (General)	3A General	1: General and Comprehensive

6.2.7 Austria

Table 52: PISA national programs of Austria

National program (progn): PISA 2018 (NOT AVAILABLE IN DATA)	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
~~00400001 Austria: Compulsory school~~	00400001 Austria: Compulsory school	00400002 Austria: Compulsory school	0400002 AUT: Lower Secondary school	0400002 AUT: LOWER SECONDARY SCHOOL	040002 AUT: Hauptschule (Lower Secondary school)	2A or C General or Pre-voc (2015) 2A General (2012-2003)	3: Vocational (For 2015, use the original ISCED orientation)[24]
/	/	00400004 Austria: Compulsory school	0400003 AUT: Vocational Programme	0400003 AUT: VOCATIONAL PROGRAMME	040003 AUT: Polytechnische Schule (Vocational)	3C Pre-vocational	3: Vocational
/	/	00400005 Austria: Compulsory school (special education)	0400004 AUT: Special education school (lower secondary)	0400004 AUT: SPECIAL EDUCATION SCHOOL (LOWER SECONDARY)	040004 AUT: Sonderschule (Special school (lower sec.))	2A General	4: Other
/	/	00400006 Austria: Compulsory school	0400005 AUT: Special education school (upper secondary)	0400005 AUT: SPECIAL EDUCTION SCHOOL (UPPER SECONDARY)	040005 AUT: Sonderschul-Oberstufe (Special school (upper sec.))	3C Pre-vocational	4: Other
/	/	00400007 Austria: Academic secondary school	0400006 AUT: Gymnasium Lower Secondary	0400006 AUT: GYMNASIUM LOWER SECONDARY	040006 AUT: AHS-Unterstufe (Gymnasium Lower Secondary)	2A General	1: General and Comprehensive
~~00400002 Austria: Academic secondary school~~	00400002 Austria: Academic secondary school	00400008 Austria: Academic secondary school	0400007 AUT: Gymnasium Upper Secondary	0400007 AUT: GYMNASIUM UPPER SECONDARY	040007 AUT: AHS-Oberstufe (Gymnasium Upper Secondary)	3A General	1: General and Comprehensive
/	/	00400009 Austria: Other lower and upper secondary school	0400008 AUT: Lower Secondary school	0400008 AUT: LOWER SECONDARY SCHOOL	/	2A General	1: General and Comprehensive
/	/	00400010 Austria: Other lower and	0400009 AUT: Upper Secondary school	0400009 AUT: UPPER SECONDARY SCHOOL	/	3A General	1: General and Comprehensive

[24] In PISA 2015 the national program (PROGN) variable does not properly distinguish the programs, the ISCED orientation (ISCEDO) in the original data set should be used instead as it provides the details to match with the PISA 2012 national program codes: 00400002- 00400007.

		upper secondary school					
~~00400003 Austria: Vocational school for apprentices~~	00400003 Austria: Vocational school for apprentices	00400011 Austria: Vocational school for apprentices	0400010 AUT: Apprenticeship	0400010 AUT: APPRENTICESHIP	040010 AUT: Berufsschule (Apprenticeship)	3B Vocational	3: Vocational
~~00400004 Austria: Intermediate technical and vocational school~~	00400004 Austria: Intermediate technical and vocational school	00400012 Austria: Intermediate technical and vocational school	0400011 AUT: Middle vocational school	0400011 AUT: MIDDLE VOCATIONAL SCHOOL	040011 AUT: BMS (Medium vocational school)	3B Vocational or 3C Pre-vocational (2015) 3B Vocational (2012-2003)	2: Technical (For 2015, use the orginal ISCED orientation) [25]
/	/	00400013 Austria: Intermediate technical and vocational school	0400012 AUT: Middle vocational school	0400012 AUT: MIDDLE VOCATIONAL SCHOOL (HOME ECONOMICS, HEALTH-SOCIAL SERVICES)	040012 AUT: Haushaltungs- und Hauswirtschaftsschulen (Medium voc.)	3C Pre-vocational	3: Vocational
/	/	/	/	0400013 AUT: MIDDLE VOCATIONAL SCHOOL (AGRICULTURAL, FORESTRY)	/	3C Pre-vocational	3: Vocational
/	/	/	0400014 AUT: Higher vocational school	0400014 AUT: HIGHER VOCATIONAL SCHOOL	040014 AUT: BHS (Higher vocational school)	3A Vocational	2: Technical
~~00400005 Austria: Higher technical and vocational college~~	00400005 Austria: Higher technical and vocational college	00400015 Austria: Higher technical and vocational college	0400015 AUT: Vocational college	0400015 AUT: VOCATIONAL COLLEGE	040015 AUT: Anst. Der Kindergarten-/Sozialpadagogik (Voc. college)	3A Vocational	2: Technical
/	/	00400017 Austria: Higher technical and vocational college	/	/	/	3A Vocational	2: Technical

[25] In PISA 2015 the national program (PROGN) variable does not properly distinguish the programs, the ISCED orientation (ISCEDO) in the original data set should be used instead as it provides the details to match with the PISA 2012 national program codes: 00400012 and 00400013.

6.2.8 Belgium

Table 53: PISA national programs of Belgium

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
/	00560001 Belgium: (First year A of first stage of) secondary Education	00560001 Belgium: (First year A of first stage of) General Education	0560001 BEL: (First year A of first stage of) General Education	0560101 BEL: (FIRST YEAR A OF FIRST STAGE OF) GENERAL EDUCATION	056111 BEL: 1st year A of 1st stage of General Education (Fl.)	2A General	1: General and Comprehensive
00560002 Belgium: (First year B of first stage of secondary Education) Remedial/Prevocational, Classified as General	/	/	/	/	056112 BEL: 1st year B of 1st stage of General Education (Fl.)	2A General	4: Other
00560003 Belgium: Second year of first stage - preparing for vocational upper secundary education, Classified as General	00560003 Belgium: Second year of first stage - preparing for vocational upper secundary education, Classified as General	00560003 Belgium: Second year of first stage - preparing for vocational secondary education	0560003 BEL: Second year of first stage - preparing for vocational secondary education	0560103 BEL: SECOND YEAR OF FIRST STAGE - PREPARING FOR VOCATIONAL SEC. EDUC.	056113 BEL: 2nd year of 1st stage, prep. voc. sec. education (Fl.)	2A General	3: Vocational
00560004 Belgium: Second year of first stage - Preparing for all streams in mainstream upper secondary education	00560004 Belgium: Second year of first stage - Preparing for all streams in mainstream upper secondary education	00560004 Belgium: Second year of first stage preparing for regular secondary education	0560004 BEL: Second year of first stage preparing for regular secondary education	0560104 BEL: SECOND YEAR OF FIRST STAGE PREPARING FOR REGULAR SEC. EDUC.	056114 BEL: 2nd year of 1st stage, prep. reg. sec. education (Fl.)	2A General	1: General and Comprehensive
00560005 Belgium: Second & third stage general secondary education	00560005 Belgium: Second & third stage general secondary education	00560005 Belgium: second and third stage regular secondary education	0560005 BEL: Second & third stage regular secondary education	0560105 BEL: SECOND & THIRD STAGE REGULAR SECONDARY EDUCATION	056115 BEL: 2nd & 3rd stage regular sec. education (Fl.)	3A General	1: General and Comprehensive
00560006 Belgium: Second & third stage technical secondary education	00560006 Belgium: Second & third stage technical secondary education	00560006 Belgium: second and third stage technical secondary education	0560006 BEL: Second & third stage technical secondary education	0560106 BEL: SECOND & THIRD STAGE TECHNICAL SECONDARY EDUCATION	056116 BEL: 2nd & 3rd stage technical sec. education (Fl.)	3A Vocational	2: Technical
00560007 Belgium: Second & third stage	00560007 Belgium: Second & third stage	00560007 Belgium: second and third stage	0560007 BEL: Second & third stage artistic secondary education	0560107 BEL: SECOND & THIRD STAGE ARTISTIC	056117 BEL: 2nd & 3rd stage	3A General (2018)	1: General and Comprehensive

artistic secondary education	artistic secondary education	artistic secondary education		SECONDARY EDUCATION	artistic sec. education (Fl.)	3A Vocational (2015-2003)	
00560008 Belgium: Second & third stage vocational secondary education	00560008 Belgium: Second & third stage vocational secondary education	00560008 Belgium: second and third stage vocational secondary education	0560008 BEL: Second & third stage vocational secondary education	0560108 BEL: SECOND & THIRD STAGE VOCATIONAL SECONDARY EDUCATION	056118 BEL: 2nd & 3rd stage vocational sec. ed. (Fl.)	3C Vocational	3: Vocational
00560009 Belgium: Part-time vocational secondary education focused on the labour market	00560009 Belgium: Part-time vocational secondary education focused on the labour market	00560009 Belgium: Part-time vocational secondary education focused on the labour market	0560009 BEL: Part-time vocational secondary education focused on the labour market	0560109 BEL: PART-TIME VOCATIONAL SEC. EDUC. FOCUSED ON THE LABOUR MARKET	056119 BEL: Part-time vocational sec. ed. for labour market (Fl.)	3C Vocational	3: Vocational
00560010 Belgium: Special secondary education - lower secondary (training form 3 / first 3 years)	00560010 Belgium: Special secondary education - lower secondary (training form 3 / first 3 years)	00560010 Belgium: Special secondary education - lower secondary (training form 3 / first 3 years)	0560010 BEL: Special secondary education - lower secondary (training form 3 / first 3 ye	0560110 BEL: SPECIAL SEC. EDUC. - LOWER SEC. (TRAINING FORM 3 / FIRST 3 YEARS)	056120 BEL: Special sec. education (Fl.)	2B Vocational	2: Technical
	00560011 Belgium: Special secondary education - upper secondary (training form 3 \ years 4 and 5)	00560011 Belgium: Special secondary education - upper secondary (training form 3 / years 4 and 5)	0560011 BEL: Special secondary education - upper secondary (training form 3 / years 4 an	0560111 BEL: SPECIAL SEC. EDUC. - UPPER SEC. (TRAINING FORM 3 / YEARS 4 AND 5)		3C Vocational	3: Vocational
00560012 Belgium: 1st or 2d year of general education (=First degree of general education (Fr/Ger))	00560012 Belgium: 1st or 2d year of general education (=First degree of general education (Fr/Ger))	00560012 Belgium: First degree of general education (Fr/Ger)	0560012 BEL: First degree of general education (Fr/Ger)	0569612 BEL: FIRST DEGREE OF GENERAL EDUCATION (FR/GER)	056231 BEL: (1st grade of)General Education (Fr.)	2A General	1: General and Comprehensive
00560013 Belgium: 1st or 2d year of remedial education	00560013 Belgium: 1st or 2d year of remedial ed (1st degree of remedial reeducation(Fr) 1st or 2nd year B special needs (Ger)	00560013 Belgium: First degree of remedial education(Fr),First or second year B special needs (Ger)	0560013 BEL: First year B special needs (Fr/Ger)	0569613 BEL: FIRST YEAR B SPECIAL NEEDS (FR/GER)	056232 BEL: Special needs (Fr.)	2A Pre-vocational (2018-2012) 2A General (2009-2003)	4: Other
/	00560014 Belgium: 1st or 2d Complementary year to 1st degree (Fr only)	00560014 Belgium: Complementary year to 1st degree (Fr only)	0560015 BEL: Complementary year to 1st degree (Fr Com only)	0569615 BEL: COMPLEMENTARY YEAR TO 1ST DEGREE (FR COM ONLY)	056234 BEL: Complementary year or programme for 1st degree (Fr.)	2A General	1: General and Comprehensive

			0560014 BEL: 2nd year of vocational education (Fr/Ger)	0569614 BEL: 2ND YEAR OF VOCATIONAL EDUCATION (FR/GER)	056233 BEL: Vocational Education (Fr.)	2B Vocational	2: Technical
00560015 Belgium: 3d, 4th, 5th, 6th year of general education (=Second & third degrees of general education (Fr/Ger))	00560015 Belgium: 3d, 4th, 5th, 6th year of general education (=Second & third degrees of general education (Fr/Ger))	00560015 Belgium: second and third degrees of general education (Fr/Ger)	0560016 BEL: Second & third degrees of general education (Fr/Ger)	0569616 BEL: SECOND & THIRD DEGREES OF GENERAL EDUCATION (FR/GER)	056235 BEL: General Education (Fr.)	3A General	1: General and Comprehensive
00560016 Belgium: 3d, 4th, 5th, 6th year of technical or artistical education (transition)	00560016 Belgium: 3rd, 4th, 5th year of tech or artistical ed (tran) (2nd & 3rd degrees of tech or artistical ed) (Fr\Ger)	00560016 Belgium: second and third degrees of technical or artistical education (transition) (Fr/Ger)	0560017 BEL: Second & third degrees of technical or artistical education (transition) (F	0569617 BEL: 2ND & 3RD DEGREES OF TECHN. OR ART. EDUC. (TRANSITION) (FR/GER)	056236 BEL: Technical or Artistical Education (transition) (Fr.)	3A General	1: General and Comprehensive
00560017 Belgium: 3d, 4th, 5th, 6th year of technical or artistical education (qualification)	00560017 Belgium: 3rd, 4th, 5th year of tech or artistical ed (qual) (2nd & 3rd degrees of tech or artistical ed) (Fr\Ger)	00560017 Belgium: second and third degrees of technical or artistical education (qualification) (Fr/Ger)	0560018 BEL: Second & third degrees of technical or artistical education (qualification)	0569618 BEL: 2ND & 3RD DEGREES OF TECHN. OR ART. EDUC. (QUALIF.) (FR/GER)	056237 BEL: Technical or Artistical Education (qualif.) (Fr.)	3A Vocational	2: Technical
00560018 Belgium: 3d, 4th, 5th, 6th year of vocational education (= Second & third degrees of vocational education) (Fr/Ger)	00560018 Belgium: 3d, 4th, 5th year of vocational education (= Second & third degrees of vocational education) (Fr\Ger)	00560018 Belgium: second and third degrees of vocational education (Fr/Ger)	0560019 BEL: Second & third degrees of vocational education (Fr/Ger)	0569619 BEL: SECOND & THIRD DEGREES OF VOCATIONAL EDUCATION (FR/GER)	056238 BEL: Vocational Education (Fr.)	3C Vocational	3: Vocational
00560019 Belgium: Vocational training focused on the labour market (Fr /Ger)	00560019 Belgium: Vocational training focused on the labour market (Fr Com only)	00560019 Belgium: Vocational training focused on the labour market (Fr Com only)	0560020 BEL: Vocational training focused on the labour market (Fr Com only)	0569620 BEL: VOCATIONAL TRAINING FOCUSED ON THE LABOUR MARKET (FR COM ONLY)	056239 BEL: Vocational training for labour market (Fr.)	2C Vocational	3: Vocational
00560020 Belgium: Special secondary education (training form 3) (Fr Com /Ger)	00560020 Belgium: Special secondary education (training form 3) (Fr Com only)	00560020 Belgium: Special secondary education (training form 3) (Fr Com only)	0560022 BEL: Special secondary education (lower secondary - training form 3) (Fr Com onl	0569622 BEL: SPECIAL SEC. EDUC. (LOWER SEC. - TRAINING FORM 3) (FR COM ONLY)	/	2B Vocational	2: Technical
/	/	/	/	0569623 BEL: SPECIAL SEC. EDUC.	056242 BEL: Special sec.	3C Vocational	3: Vocational

				(UPPER SEC.-TRAINING FORM 3)) (FR. ONLY)	education (form 3 or 4 - voc.) (Fr.)		
/	/	00560022 Belgium: Special secondary education (lower secondary) (Germ Com only)	0560024 BEL: Special secondary education (lower secondary) (Germ Com only)	0569624 BEL: SPECIAL SEC. EDUC. (LOWER SEC.) (GER. ONLY)	/	2B Vocational	2: Technical
/	/	/	/	/	056244 BEL: Special sec. education form 3 (Germ.)	3C Vocational	3: Vocational
/	/	/	0560026 BEL: Part-time Vocational Education (Germ Com only)	/	056245 BEL: Part-time Vocational Education (Germ.)	2C Vocational	3: Vocational
/	/	/	/	/	056246 BEL: Vocational Education (Germ.)	2C Vocational	3: Vocational

6.2.9 Czech Republic

Table 54: PISA national programs of Czech Republic

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
02030001 Czech Republic: Basic school	02030001 Czech Republic: Basic school	02030001 Czech Republic: Basic school	2030001 CZE: Basic school	2030001 CZE: BASIC SCHOOL	203001 CZE: Basic school	2A General	1: General and Comprehensive
02030002 Czech Republic: 6, 8-year gymnasium and 8-year conservatory (lower secondary)	02030002 Czech Republic: 6, 8-year gymnasium and 8-year conservatory (lower secondary)	02030002 Czech Republic: 6, 8-year gymnasium and 8-year conservatory (lower secondary)	2030002 CZE: 6, 8-year gymnasium and 8-year conservatory (lower secondary)	2030002 CZE: 6, 8-YEAR GYMNASIUM AND 8-YEAR CONSERVATORY (LOWER SECONDARY)	203002 CZE: 6, 8-year gymnasium & 8-year conservatory (lower sec.)	2A General	1: General and Comprehensive
02030003 Czech Republic: 6, 8-year gymnasium (upper secondary)	02030003 Czech Republic: 6, 8-year gymnasium (upper secondary)	02030003 Czech Republic: 6, 8-year gymnasium (upper secondary)	02030003 CZE: 6, 8-year gymnasium (upper secondary)	2030003 CZE: 6, 8-YEAR GYMNASIUM (UPPER SECONDARY)	203003 CZE: 6, 8-year gymnasium (upper sec.)	3A General	1: General and Comprehensive
02030004 Czech Republic: 4-year gymnasium	02030004 Czech Republic: 4-year gymnasium	02030004 Czech Republic: 4- year gymnasium	2030004 CZE: 4- year gymnasium	2030004 CZE: 4- YEAR GYMNASIUM	203003 CZE: 6, 8-year gymnasium (upper sec.)	3A General	1: General and Comprehensive
02030005 Czech Republic: Voc/tech secondary school with maturate	02030005 Czech Republic: Voc/tech secondary school with maturate	02030005 Czech Republic: Voc/tech secondary school with maturate	2030005 CZE: Voc/tech secondary school with maturate	2030005 CZE: VOC/TECH SECONDARY SCHOOL WITH MATURATE	203005 CZE: Voc/tech sec. school with maturate	3A Vocational	2: Technical
02030006 Czech Republic: Conservatory (upper secondary)	/	02030006 Czech Republic: Conservatory (upper secondary)	/	/	203006 CZE: Conservatory (upper sec.)	3B Pre-vocational (2018) 3B Vocational (2012, 2003)	2: Technical
02030007 Czech Republic: Voc/tech secondary school without maturate	02030007 Czech Republic: Voc/tech secondary school without maturate	02030007 Czech Republic: Voc/tech secondary school without maturate	2030007 CZE: Voc/tech secondary school without maturate	2030007 CZE: VOC/TECH SECONDARY SCHOOL WITHOUT MATURATE	203007 CZE: Voc/tech sec. school without maturate	3C Vocational	3: Vocational
02030008 Czech Republic: Basic special/practical school	02030008 Czech Republic: Basic special schools	02030008 Czech Republic: Basic special schools	2030008 CZE: Basic special schools	2030008 CZE: SPECIAL SCHOOLS	203008 CZE: Special schools	2B General (2018) 2C General (2015-2003)	4: Other

02030009 Czech Republic: Practical school	02030009 Czech Republic: Secondary special schools - practical schools, vocational education predominantly	02030009 Czech Republic: Secondary special schools - practical schools, vocational education predominantly	2030009 CZE: Secondary special schools - practical schools, vocational education predomi	2030009 CZE: PRACTICAL SCHOOLS, VOCATIONAL EDUCATION PREDOMINANTLY	203009 CZE: Practical schools, vocational education predominantly	2C Pre-vocational (2018) 2C Vocational (2015 -2003)	3: Vocational

6.2.10 Netherlands

Table 55: PISA national programs of the Netherlands

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
05280001 Netherlands: practical preparation for labour market	05280001 Netherlands: practical preparation for labour market	05280001 Netherlands: practical preparation for labour market	5280001 NLD: practical preparation for labour market	5280001 NLD: PRACTICAL PREPARATION FOR LABOUR MARKET	528001 NLD: PRO	2C Vocational	3: Vocational
05280002 Netherlands: general vocational oriented; in next year tracks into BB, KB, GL, TL	05280002 Netherlands: general vocational oriented; in next year tracks into BB, KB, GL, TL	05280002 Netherlands: general vocational oriented; in next year tracks into BB, KB, GL, TL	5280002 NLD: general vocational oriented; in next year tracks into BB, KB, GL, TL	/	/	2B Genral	2: Technical
05280003 Netherlands: Admission to ISCED 3 = MBO level 1-3, post secondary, non tertiary	05280003 Netherlands: Admission to ISCED 3 = MBO level 1-3, post secondary, non tertiary	05280003 Netherlands: Admission to ISCED 3 = MBO level 1-3, post secondary, non tertiary	5280003 NLD: admission to ISCED 3 =MBO level 1-3, post secondary, non tertiary	/	/	2B Pre-vocational	2: Technical
05280004 Netherlands: admission to ISCED 3 = MBO level 1-3	05280004 Netherlands: admission to ISCED 3 = MBO level 1-3	05280004 Netherlands: admission to ISCED 3 = MBO level 1-3	5280004 NLD: admission to ISCED 3 = MBO level 1-3	/	/	2B Pre-vocational	2: Technical
05280005 Netherlands: admission to ISCED 3 = MBO tertiary level, preparing for upper secondary HAVO	05280005 Netherlands: admission to ISCED 3 = MBO tertiary level, preparing for upper secondary HAVO	05280005 Netherlands: admission to ISCED 3 = MBO tertiary level, preparing for upper secondary HAVO	5280005 NLD: admission to ISCED 3 = MBO tertiary level,	/	/	2B General	2: Technical
/	/	/	/	5280002 NLD: VMBO (GENERAL VOC.)	528002 NLD: VMBO	2A General (2006) 2B Pre-vocational (2003)	2: Technical
/	/	/	/	5280004 NLD: VMBO BB (3-4 YEAR)	528003 NLD: VMBO BB	2B Pre-vocational (2006) 2B Vocational (2003)	2: Technical
/	/	/	/	5280006 NLD: VMBO KB (3-4 YEAR)	528004 NLD: VMBO KB	2B Pre-vocational (2006) 2B Vocational (2003)	2: Technical

/	/	/	/	5280008 NLD: VMBO GL/ TL (3-4 YEAR)	528005 NLD: VMBO GL	2B General (2006) 2B Vocational (2003)	2: Technical
/	/	/	/	/	528006 NLD: VMBO TL	2B Pre-vocational	2: Technical
05280006 Netherlands: Lower secondary grades of HAVO, preparing for upper secondary HAVO or MBO	05280006 Netherlands: Lower secondary grades of HAVO, preparing for upper secondary HAVO or MBO	05280006 Netherlands: HAVO: advanced general secondary education/ admission to higher voc./ college, so tertiary level	5280006 NLD: HAVO: advanced general secondary education/ admission to higher voc./ colle	5280009 NLD: HAVO (YEAR 1-3)	528007 NLD: HAVO 2/3	2A General	1: General and Comprehensive
05280007 Netherlands: Upper secondary grades of HAVO, preparing for HBO (higher vocational education/college)	05280007 Netherlands: Upper secondary grades of HAVO, preparing for HBO (higher vocational education/college)	05280007 Netherlands: HAVO: advanced general secondary education/ admission to higher voc./ college, so tertiary level	5280007 NLD: HAVO: advanced general secondary education/ admission to higher voc./ colle	5280010 NLD: HAVO (SEC. YEAR 4-5)	528008 NLD: HAVO 4/5	3A General	1: General and Comprehensive
05280008 Netherlands: Lower secondary grades of VWO, preparing for upper secondary VWO	05280008 Netherlands: Lower secondary grades of VWO, preparing for upper secondary VWO	05280008 Netherlands: VWO: admission to university (ISCED2)	5280008 NLD: VWO: admission to university (ISCED2)	5280011 NLD: VWO (YEAR 1-3)	528009 NLD: VWO 2/3	2A General	1: General and Comprehensive
05280009 Netherlands: Upper secondary grades of VWO, preparing for university	05280009 Netherlands: Upper secondary grades of VWO, preparing for university	05280009 Netherlands: VWO: admission to university (ISCED3)	5280009 NLD: VWO: admission to university (ISCED3)	5280012 NLD: VWO (YEAR 4-6)	528010 NLD: VWO 4/5	3A General	1: General and Comprehensive
05280010 Netherlands: Mixed 1-2, no tracking, preparing for VMBO or HAVO or VWO, determined at the end of grade 2	/	/	/	5280007 NLD: VMBO GL/ TL (1-2 YEAR)	/	2A General	2: Technical

/	/	/	/	5280003 NLD: VMBO BB (1-2 YEAR)	/	2A General	2: Technical
/	/	/	/	5280005 NLD: VMBO KB (1-2 YEAR)	/	2A General	2: Technical

6.2.11 Slovak Republic

Table 56: PISA national programs of Slovak Republic

National program (progn): PISA 2018	National program (progn): PISA 2015	National program (progn): PISA 2012	National program (progn): PISA 2009	National program (progn): PISA 2006	National program (progn): PISA 2003	ISCED level, designation, orientation	Proposed harmonisation
07030001 Slovak Republic: Basic school	07030001 Slovak Republic: Basic school	07030001 Slovak Republic: Basic school	7030001 SVK: Basic school	7030001 SVK: BASIC SCHOOL	703002 SVK: Basic school (lower sec.)	2A General	1: General and Comprehensive
07030002 Slovak Republic: Vocational Basic school	07030002 Slovak Republic: Vocational Basic school	07030002 Slovak Republic: Vocational Basic school	7030002 SVK: Vocational Basic school	7030002 SVK: VOCATIONAL BASIC SCHOOL	703004 SVK: Vocational Basic school (lower sec.)	2B General	2: Technical
07030003 Slovak Republic: Secondary school (ISCED 2)	07030003 Slovak Republic: Secondary school (ISCED 2)	07030003 Slovak Republic: Secondary school (ISCED 2)	7030003 SVK: Secondary school (ISCED 2)	7030003 SVK: GENERAL 8-YEAR SECONDARY SCHOOL (YEARS 1-4)	703005 SVK: sec. school (lower sec.)	2A General	1: General and Comprehensive
07030004 Slovak Republic: Secondary school (ISCED 3)	07030004 Slovak Republic: Secondary school (ISCED 3)	07030004 Slovak Republic: Secondary school (ISCED 3)	7030004 SVK: Secondary school (ISCED 3)	7030004 SVK: GENERAL 8-YEAR SECONDARY SCHOOL (YEARS 5-8)	703006 SVK: sec. school (upper sec.)	3A General	1: General and Comprehensive
07030005 Slovak Republic: High School	07030005 Slovak Republic: High School	07030005 Slovak Republic: High School	7030005 SVK: High School	7030005 SVK: HIGH SCHOOL (GYMNASIUM)	703007 SVK: High School (Gymnasium)	3A General	1: General and Comprehensive
07030006 Slovak Republic: Secondary College - branch/class with a school leaving examination	07030006 Slovak Republic: Secondary College - branch/class with a school leaving examination	07030006 Slovak Republic: Secondary College - branch/class with a school leaving examination	7030006 SVK: Secondary College	7030006 SVK: SECONDARY COLLEGE	703008 SVK: sec. College	3M Modular (2018) 3A Modular (2015-2012) 3A Pre-vocational (2009-2003)	2: Technical
/	/	/	7030007 SVK: Technical College - branch/class with a school leaving examination	7030007 SVK: TECHNICAL COLLEGE, CLASS WITH A SCHOOL LEAVING EXAMINATION	703009 SVK: Technical College	3A Vocational	2: Technical
07030007 Slovak Republic: Secondary	07030007 Slovak Republic: Secondary	07030007 Slovak Republic: Secondary	7030008 SVK: Secondary College	7030008 SVK: TECHNICAL	/	3C Vocational	3: Vocational

College - branch/class without a school leaving examination	College - branch/class without a school leaving examination	College - branch/class without a school leaving examination	or Technical College - branch/class without the school le	COLLEGE, CLASS WITHOUT A SCHOOL LEAVING EXAMINATION			
/	/	/	/	7030009 SVK: VOCATIONAL COLLEGE	703010 SVK: Vocational College	3C Vocational	3: Vocational