



Mapping Longitudinal Datasets on Educational Inequalities in Primary and Secondary Education: Research Findings and Implications

WP 3 - Deliverable 3.1 Mapping Research Report

Lead authors: Emmanuele Pavolini, Gianluca Argentin, Veronica Mobilio

Authors: Ilaria Lievore, Ekaterina Kochergina, Giulietta Zanga

Main Reviewers: Martin Neugebauer, Gabriele Ballarino

2025



Funded by the European Union's Horizon Europe Framework Programme (HORIZONCL2-2023-TRANSFORMATIONS-01-06) under grant agreement No 101132455. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.



Funded by
the European Union



All LINEup resources are licensed under [CC BY 4.0 LEGAL CODE Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)

Deliverable Information

Project Information	
Project number	101132455
Project acronym	LINEup
Project title	LINEup - Longitudinal data for INequalities in Education
Start date	01/02/2024
Duration	36 months
Type of Action	HORIZON-RIA
Call Identifier	HORIZON-CL2-2023-TRANSFORMATIONS-01
Project website	https://www.lineup-project.eu/

Deliverable D3.1 mapping Research Report	
Work Package	WP3
Lead Partner	University of Milan (UNIMI)
Due date	31/07/2024
Submission date	30/07/2024
Dissemination level	PU-Public
Lead authors	Emmanuele Pavolini, Gianluca Argentin, Veronica Mobilio
Authors	Ilaria Lievore, Ekaterina Kochergina, Giulietta Zanga
Main Reviewers	Martin Neugebauer, Gabriele Ballarino
Other Reviewers	Johanna Unewisse, Luis Ortiz, Giampiero Passaretta, Kars van Oosterhout, Léonard Moulin, Priya Maurya, Ioannis Cholezas, Nikolaos Kanellopoulos, Maria Fragkiadaki Theodorouleas, Vincenzo Mauro, Daniel Santos Da Costa.
Version	0.3

History of changes

Version	Date	Change History	Organisation
0.1	04/07/2025	Initial version	FpS, UNIMI
0.2	16/07/2025	First Review	Main Reviews (PHKA-UNIMI) Partners (PHKA, UPF, INED, KEPE, PIREUS, UNIMC, UTAD)
0.3	21/07/2025	Second Review	FpS
0.4	25/07/2025	Third Review	UNIMI
1	29/08/2025	Final version	FpS, UNIMI

Disclaimer: This project is funded by the European Union's Horizon programme under grant agreement No 101132455. However, the views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Mapping Report Author Statement

Lead authors: Emmanuele Pavolini¹, Gianluca Argentin², Veronica Mobilio³

Authors: Ilaria Lievore³, Ekaterina Kochergina², Giuletta Zanga¹

Lead reviewers: Martin Neugebauer⁴, Gabriele Ballarino¹

Reviewer: Johanna Unewisse⁴, Luis Ortiz⁵, Giampiero Passaretta⁵, Kars van Oosterhout⁵, Léonard Moulin⁶, Priya Maurya⁶, Ioannis Cholezas⁷, Nikolaos Kanellopoulos⁷, Maria Fragkiadaki Theodorouleas⁸, Vincenzo Mauro¹⁰, Daniel Santos Da Costa¹¹.

Participating Institutions:

1. Università degli Studi di Milano La Statale (UNIMI - WP3 leader), Italy
2. Università degli Studi Di Milano-Bicocca (UniMiBicocca – WP3 leader), Italy
3. Fondazione per la Scuola della Compagnia di San Paolo (FpS – LINEup project coordinator), Italy
4. Karlsruhe University of Education (PHKA), Germany
5. Universidad Pompeu Fabra (UPF), Spain
6. Institut National D'Etudes Demographiques (INED), France
7. Centre of Planning and Economic Research (KEPE), Greece
8. University of Pireus (Pireus - WP2 leader), Greece
9. Computer Technology Institute and Press "Diophantus"
10. Università Degli Studi di Macerata (UNIMC), Italy
11. Universidade de Trás-os-Montes e Alto Douro (UTAD), Portugal

CRedit Authorship Contribution Statement¹:

Emmanuele Pavolini: Conceptualisation, Methodology, Validation, Formal Analysis, Writing - Original Draft, Writing - Review & Editing, Visualisation, Supervision, Project administration

Gianluca Argentin: Conceptualisation, Methodology, Validation, Formal Analysis, Writing - Original Draft, Writing - Review & Editing, Visualisation, Project administration, Supervision

Veronica Mobilio: Conceptualisation, Methodology, Validation, Formal Analysis, Writing - Original Draft, Writing - Review & Editing, Visualisation, Supervision, Project administration, Funding acquisition

Ilaria Lievore: Methodology, Validation, Investigation, Data Curation, Writing - Original Draft, Visualisation

Ekaterina Kochergina: Validation, Investigation, Data Curation, Writing - Original Draft, Visualisation

Giuletta Zanga: Validation, Investigation, Data Curation, Writing - Original Draft

Infographic Layout: Andrea Succhielli, European Grants International Academy (EGInA)

Corresponding author: Veronica Mobilio, Piazza Bernini, 5 10138 Torino - Italy, +39 011 1563 0500, veronica.mobilio@fondazione scuola.it

Acknowledgement of Contributors to Dataset Mapping: the authors gratefully acknowledge the researchers and institutional teams who contributed to the identification, collection, and documentation of the longitudinal datasets analysed in this report. Contributors include researchers from the eleven institutions listed above.

How to cite:

Pavolini E., Argentin G., Mobilio V., Lievore I., Kochergina E., Zanga G. (2025). *Mapping Longitudinal Datasets on Educational Inequalities in Primary and Secondary Education: Research Findings and Implications*. Deliverable D3.1, Horizon Europe Project LINEup.

¹ CRediT Authorship: <https://www.elsevier.com/researcher/author/policies-and-guidelines/credit-author-statement>

Abstract

Background: Longitudinal data offer valuable insights into students' learning trajectories and how educational inequalities emerge, persist, or change over time. Yet, while some datasets are well known and widely used, many others remain underused and unevenly documented, limiting their contribution to research and policymaking. To address this gap, the LINEup project conducted a comprehensive mapping of existing national and regional longitudinal datasets on student learning outcomes at the primary and secondary education levels (International Standard Classification of Education - ISCED 1, 2, 3). By making these data sources more visible, the project helps unlock their potential to inform evidence-based research and promote fairer education policies.

Objectives: This report presents the outcomes of WP3 of the LINEup project, which aimed to: (i) Identify and document existing longitudinal datasets on student learning outcomes in 32 countries; (ii) Assess their structure, content, accessibility, and relevance for studying educational inequalities; (iii) Classify and analyse the datasets based on common criteria to support future harmonisation and comparative research; (iv) Lay the groundwork for the next steps of the project.

Methodology: The mapping followed a multi-step strategy grounded in a shared theoretical and methodological framework. It built upon the systematic literature review (conducted in WP2); an expert questionnaire disseminated to over 300 stakeholders across 28 countries; structured web searches by project partners and collaboration with sister projects. Each dataset was assessed against predefined inclusion criteria and documented using a standardised online form, capturing metadata on structure, coverage, outcomes, background variables, and accessibility.

Key findings: Out of 229 datasets identified, 104 met the inclusion criteria and were included in the mapping. These span 24 of the 32 European countries covered, with Germany and the UK reporting the highest numbers. The longitudinal datasets originate from administrative data, ad hoc surveys and hybrid sources. Their span ranges from 2 to over 14 years, though more than half cover only one education level. Almost all datasets include information on literacy and numeracy skills and at least basic student background characteristics. By contrast, non-cognitive outcomes as well as teacher- and school-level information are often missing or inconsistently documented. Access to microdata is uneven. Overall, the mapping confirms wide disparities in data availability, structure, and quality across countries.

Implications: The mapping highlights both the opportunities and limitations in using longitudinal data to study patterns of educational inequality. For researchers, it offers a valuable, despite uneven, empirical foundation for analysing students' learning trajectories. The accompanying classification framework and rich metadata provide a structured basis for datasets selection and comparative use. For policymakers and education stakeholders, the mapping facilitates the identification of key sources of empirical evidence on school education, supporting the design of more effective, evidence-based interventions. More broadly, the findings underscore the need to strengthen national data infrastructures, enhance data documentation and accessibility, and guide future data collection efforts toward better capturing the multidimensional nature of equity in education.

Next Steps: The results of this mapping underpin the development of the *LINEup Education Data Explorer*, an online platform to enhance dataset discovery and use by researchers and educational stakeholders. They also provide the empirical basis for the upcoming *Feasibility Study* on data harmonisation (*Deliverable D3.2*), to be published in January 2026. This study will assess the potential for cross-country comparison and the development of shared indicators for analysing educational inequalities.

Table of Contents

Abstract	4
Executive summary.....	8
1. Introduction	14
2. Theoretical and Methodological Framework.....	17
Key Concepts and Definitions	17
Methodological Approach.....	19
Dataset Inclusion Criteria	21
Challenges and Limitations	22
3. Mapping of Longitudinal Datasets: Methods and Procedures.....	24
3.1 Preparatory Activities: Identification of Data Sources	24
Insights from the Systematic Literature Review.....	25
Expert Questionnaire: Expanding Dataset Coverage through Stakeholder Input.....	27
Sister Projects: A Call for Cooperation to Strengthen Dataset Mapping and Analysis.....	29
3.2 Guidelines for Datasets Description.....	30
Step-by-Step Mapping Protocol	30
Keywords and Practices for Systematic Web-Search.....	31
A Standard Template for In-Depth Datasets Descriptions.....	32
3.3 Mapping Results.....	33
Distribution and Geographical Coverage.....	35
Sources of Dataset Identification.....	36
Reasons for Dataset Exclusion	38
4. Dataset Analysis	41
4.1 Dataset Validation and Classification	41
Validation and Classification Framework.....	41
Data Validation Process	43
Data Validation Results	44
Data Quality Assessment Results	45
Classification Scheme.....	46
4.2 Descriptive Analysis.....	47
How: Methodology	48
What: Information Collected	51
Where: Countries Included.....	57
When: Duration and Depth of Longitudinal Data	59
Data Accessibility	62
4.3 Analytical Potential & Barriers.....	65
The Analytical Potential.....	66
Cross-cutting Challenges for Comparative Use and Research Design Reflections	67
5. Visualising Longitudinal Data: The LINEup Education Data Explorer	72

General Overview	72
Structure and Functionalities.....	72
Next Steps for Development	75
6. Discussion: Interpreting the Research Potential and Limitations of Mapped Datasets.....	77
7. Conclusions and Implications for Research, Policy, and Practice	83
Summary of Key Findings	83
Implications for Research, Policy, and Practice	84
Next Steps	84
8. References	87
9. Appendices	89
9.1 List of mapped and excluded datasets.....	89
I Part - Mapped Datasets	89
II Part - Excluded Datasets	92
9.2 Structure of the Expert Questionnaire.....	98
9.3 Structure of Dataset Description Form.....	105
9.4 Glossary of Technical Terms	113

List of Tables

Table 1: Distribution of countries among partners.....	21
Table 2: Inclusion Criteria - Literature Review vs. Dataset Mapping	26
Table 3: Suggested keywords for systematic web-search	32
Table 4: Number of datasets identified, included, and excluded by country by partner	34
Table 5: Number of datasets identified by sources	37
Table 6: Number of datasets excluded by reasons of exclusion	39
Table 7: Exclusion criteria	42
Table 8: Criteria for manual review.	43
Table 9: Type of longitudinal design	48
Table 10: Type of sample and sample size.....	49
Table 11: Information gathered through administrative sources.....	50
Table 12: Linkability of student's information with other data at the individual level.	50
Table 13: Type of education outcomes included in the datasets.	51
Table 14: Students' performance included in the datasets by specific topic.....	52
Table 15: Distribution of datasets by type of scale used for skill measurement.....	52
Table 16: Distribution of datasets by presence of info on students' characteristics	53
Table 17: Distribution of datasets by presence of info on students' nationality or country of birth ..	54
Table 18: Distribution of datasets by presence of info on student household's characteristics.....	55
Table 19: Distribution of datasets by presence of info on teachers' characteristics.....	56
Table 20: Distribution of datasets by presence of info on school's characteristics.	56
Table 21: Distribution of datasets by country.	58
Table 22: Geographical scope of the datasets.....	59

Table 23: Distribution of datasets according to countries and their geographical scope	59
Table 24: Data-Collection Duration across Mapped Datasets.....	60
Table 25: School grades covered in the datasets	61
Table 26: Classification of datasets according to the coverage of primary/secondary education ...	61
Table 27: Classification of datasets according to the year of ending of the study	62
Table 28: Classification of datasets according to microdata accessibility.....	63
Table 29: A framework to assess the availability and quality of longitudinal datasets for the study of education inequalities	78

List Figures

Figure 1: The funnel strategy of the mapping.	21
Figure 2: Number of datasets identified by sources (N=227).....	36
Figure 3: Reasons for dataset exclusion from the mapping (N=124).	39
Figure 4: Flowchart of the validation activity.	44
Figure 5: The landing page of the LINEup Education Data Explorer.....	73
Figure 6: LINEup Education Data Explorer - Map View.	74
Figure 7: LINEup Education Data Explorer - List View.....	74

Executive summary

Introduction

Longitudinal data offer valuable insights into students' learning trajectories and how educational inequalities emerge, persist, or change over time. While many European countries have invested in education data collection, longitudinal datasets specifically tracking student learning outcomes over time remain fragmented and inconsistently documented. The LINEup project addresses this gap by conducting the first comprehensive mapping of existing national and regional longitudinal datasets on student learning outcomes at the primary and secondary education levels (International Standard Classification of Education - ISCED 1, 2, 3).

The mapping covers 32 countries: the 27 EU Member States, three European Economic Area (EEA) associated countries (Iceland, Liechtenstein, and Norway), as well as Switzerland and the United Kingdom. This selection reflects the primary geographic scope of the LINEup project, which focuses on European education systems while ensuring representation of diverse policy frameworks, administrative structures, and data collection traditions within the broader European context. The inclusion of EEA countries, Switzerland, and the UK allows for a more comprehensive understanding of longitudinal data availability and use beyond the EU alone. In addition to this core group, selected practices from outside Europe - particularly from the United States, displaying leading examples in the field - were considered and reviewed during the methodological design to enrich the project's comparative lens.

This report presents the results of the mapping exercise, conducted under WP3 of the LINEup project. The mapping aimed to identify, document, and classify existing datasets; evaluate their analytical relevance for studying educational inequalities; and explore their potential for cross-country comparability. The mapping also provides the empirical foundation for two key project outputs: the *LINEup Education Data Explorer*, an online platform that supports dataset discovery and reuse by researchers and educational stakeholders; and a *Feasibility Study* on data harmonisation (Deliverable D3.2 due in January 2026), which will assess the conditions for aligning and comparing data structures and indicators across countries. Together, these outputs aim to foster a stronger and more integrated use of longitudinal data in education research and equity-oriented policy development.

Methodology

The mapping followed a structured, multi-step strategy grounded in a shared theoretical and methodological framework. It built upon three main sources:

- A **systematic literature review** (conducted in WP2), which identified 43 relevant datasets used in empirical studies on educational inequalities with a longitudinal or repeated cross-sectional research design (out of the 77 collected overall).
- An **expert questionnaire** sent to over 300 stakeholders across 28 countries, which led to 70 additional dataset suggestions.
- **Structured web searches and desk research** conducted by project partners across 32 countries, complemented by collaboration with sister projects, which added a further 116 datasets to the inventory.

All suggested datasets were reviewed using predefined inclusion criteria: they had to be longitudinal; refer to primary (ISCED 1) and/or secondary education (ISCED 2, 3); focus on student learning

outcomes (such as standardized achievement tests or school grades); originate from one of the 27 EU Member States, EEA associated countries, Switzerland, or the UK; be nationally or regionally representative (or based on a large scale sample); and be ongoing or recently collected (with the latest wave not older than 2014).

Eligible datasets were documented using a standardised online form that captured key metadata on basic dataset information (name, responsible institution, country, accessibility conditions) along with details on longitudinal structure, grade coverage, data type, skills and educational outcomes, background variables related to student, household, teacher and school characteristics.

A structured validation protocol was applied to ensure relevance, accuracy, and comparability of the documented datasets. This included an initial quality review by the mapping team, followed by peer validation by at least one additional project partner. Datasets flagged as *uncertain* or *borderline* were subject to manual re-assessment by the WP3 coordination team. Finally, all validated datasets were reviewed a second time to ensure cross-country consistency and alignment with inclusion criteria and classification guidelines. The entire validation process was grounded in a transparent, replicable methodological framework, ensuring coherence across countries and robustness in comparative dataset classification.

Key Findings

Of the 229 datasets identified, 104 met the inclusion criteria and were included in the final mapping. These datasets are distributed across 24 of the 32 countries covered by LINEup. The countries with the highest number of eligible datasets are Germany (31) and the United Kingdom (11), followed by Finland and Portugal (both 7), and France (5). A clear geographic pattern emerges: most countries in Western and Northern Europe have at least one eligible dataset, often stemming from well-established administrative systems or long-standing student assessments. In contrast, several countries in Central and Eastern Europe have no eligible datasets, reflecting more limited investments in longitudinal data infrastructure and/or lower levels of data accessibility. These patterns often mirror national policy priorities, institutional capacity, and the overall maturity of student assessment or data collection systems.

The mapped datasets reveal substantial variation in design, scope, coverage, and accessibility. Nearly half (46.6%) rely exclusively on survey data, while a smaller share (19.4%) is based solely on administrative sources. An additional 34% adopt a hybrid approach that combines both survey and administrative information. Although administrative datasets often reflect national assessment systems and official registries, typically covering larger share of the population and longer timespan, survey-based and hybrid datasets tend to provide richer contextual and background information, particularly on student characteristics, household conditions, and school environments. This is largely due to their research-oriented design.

The longitudinal span of the datasets ranges from 2 to over 14 years, reflecting diverse research objectives and data collection strategies. While many datasets follow students for five to eight years, longer-term tracking remains relatively rare. Nearly half of datasets (47.1%) cover both primary and secondary education, while the remaining focus exclusively on either primary (21.2%) or secondary education levels (31.7%). Coverage is often concentrated around key transition points - particularly grades 8, 9, and 10 - rather than providing continuous observation across all educational stages. Notably, around 30% of datasets begin in pre-primary education, enabling early tracking of inequalities before formal schooling. While several datasets include data points in upper secondary education (grades 11–13), continuous individual-level tracking across the full span of compulsory

schooling is relatively rare. The few datasets that provide this level of coverage, typically designed as birth cohorts, remain the exception.

In terms of outcomes, the vast majority of datasets include academic performance measures in the form of standardised test scores and/or teacher-assigned grades in literacy and numeracy. Fewer datasets (around 30%) cover either science, foreign languages, or other skills domains. School- and teacher-level variables are inconsistently reported: only a minority of datasets include classroom-level information or teacher-level data.

Data linkability is uneven across the mapped datasets. Only 29.1% are both linkable and accessible to external users; an additional 26.2% allow linkage but restrict access to data owners, while 44.7% does not allow linkability to other data sources at all. True inter-dataset linkage - such as connecting learning outcomes to enrolment records or external registries - is generally limited to countries with strong administrative systems, such as Germany or the Nordic countries. Legal, technical, and institutional barriers continue to hinder broader data integration and reuse, particularly where student identifiers are absent, or governance frameworks are fragmented.

Finally, access to microdata represents a major constraint. Only 10.7% of datasets are publicly accessible (open data), while 49.5 % require an ad hoc request and formal authorisation to obtain full access. A further 26.2% are either inaccessible or subject to significant restrictions. In addition, metadata on accessibility conditions are unclear or not available for a consistent share of datasets (13.6%). Overall, metadata have been often reported to be incomplete, outdated, or inconsistently documented - limiting transparency and reducing the potential for secondary analysis and evidence-based policy use. The combination of limited accessibility, documentation gaps, and divergent metadata standards reduce the usability of longitudinal education data across the mapped countries. Moreover, while the diversity of aims, structures, and target populations across datasets reflects the richness of the field, it also entails potential barriers that researchers and policymakers may face when trying to identify the most suitable sources for specific analytical or strategic purposes.

To support dataset analysis, harmonisation, and comparative use, a structured classification framework was developed. It organises the mapped datasets according to data structure, outcome types, availability of equity-related variables, and longitudinal tracking potential. This framework also underpins the visualisation and filtering functionalities of the *LINEup Education Data Explorer*.

Discussion and Conclusions

The mapping reveals both opportunities and limitations in the current landscape of longitudinal education data in Europe. On the one hand, the 104 datasets included in the mapping provide a unique, albeit uneven, empirical foundation for analysing student learning trajectories and educational inequalities. On the other hand, persistent disparities in data access, scope, and usability hinder broader analytical and comparative potential.

Several key insights emerge:

- **Data exist but are underused and frequently inaccessible:** Although numerous longitudinal datasets have been developed across Europe, they remain underutilised due to a combination of structural and procedural constraints. In several cases, even high-quality administrative data are off-limits to researchers because of the absence of clear data-sharing protocols, insufficient institutional coordination, or overly cautious interpretations of data protection regulations.
- **Significant heterogeneity in variables collected:** While most datasets focus on academic outcomes, particularly literacy and numeracy, the coverage of other dimensions - such as

non-cognitive skills, school climate, teaching practices, or equity-related indicators - varies widely. Only few datasets collect repeated measures on socio-emotional outcomes or classroom-level dynamics, which are essential for understanding broader learning conditions.

- **Limited comparability across countries:** Despite the presence of valuable data infrastructures in many countries, the lack of harmonised definitions, coding schemes, and measurement instruments limits cross-country comparison. For example, similar constructs (e.g., parental education) are measured differently across datasets - even within the same country - complicating efforts to build shared indicators or conduct joint analyses.
- **Uneven quality and granularity:** The granularity of information also varies. While some datasets enable longitudinal tracking at the individual level across multiple transitions, others are restricted to a single education stage or provide only aggregate data. This limits the scope of research questions that can be addressed, particularly those related to equity trajectories, long-term outcomes, or causal inference.
- **Technical and legal barriers to linkage:** Although linkability across datasets exists in several contexts - especially where strong administrative systems are in place - it remains underexploited. Legal, technical, and institutional barriers hinder more systematic integration, particularly in countries where centralised student identifiers are absent and governance frameworks for data management remain fragmented or underdeveloped.

Despite these limitations, the mapping lays the groundwork for improved use and potential harmonisation of education data in Europe. It enables researchers and educational stakeholders to identify promising datasets for future analysis and to draw inspiration from exemplary practices. While other countries - such as the UK or Finland - also host well-developed longitudinal datasets, the Italian INVALSI-SNV system and the German NEPS programme were selected as illustrative cases due to their comprehensive design, longitudinal depth, inclusion of contextual variables, and accessibility protocols. These features closely align with the core dimensions of the LINEup mapping framework, illustrating how coordinated investment in data infrastructure and governance can foster robust and policy-relevant research on educational inequalities. Beyond Europe, the U.S. Early Childhood Longitudinal Study (ECLS) provides an additional benchmark: by tracking the same individuals over time and systematically linking student data with information on families, teachers, and schools, it exemplifies the potential of fully integrated longitudinal systems to support developmental research and policy evaluation. The existence and use of rich longitudinal datasets in several countries can serve as a catalyst for increased awareness and commitment among national authorities in less advanced contexts, encouraging progress in data infrastructure and governance.

Implications for Research, Policy, and Practice

The results of the mapping exercise have important implications across multiple domains.

For researchers: The validated inventory of 104 datasets, supported by a structured classification framework and rich metadata, offers a valuable starting point for future comparative studies and secondary analyses and for informing the design of future data collections that leverage the strengths of existing datasets. By enabling dataset filtering by type, coverage, and variables, the mapping facilitates more informed research design and supports future harmonisation efforts. It also highlights best practices that can inform methodological innovation and cross-country collaboration.

For policymakers (including national and local authorities): The mapping enhances transparency regarding the availability and quality of longitudinal education data across Europe. It highlights significant gaps in data infrastructure - especially in Central and Eastern Europe - and identifies barriers to data accessibility, including restrictive authorisation regimes or insufficient documentation. These insights are critical for designing policies that are both evidence-based and responsive to equity challenges. The report also emphasises the need for coordinated investment in data systems as a strategic asset for education governance. Finally, it draws attention to severe limitations in the capacity to link existing datasets with administrative sources - an obstacle that hampers efforts to understand the complex mechanisms underpinning the development of educational inequalities.

For education systems and practitioners: The mapping facilitates the identification of key sources of empirical evidence on education inequalities, supporting the design of evidence-based interventions. Its findings highlight the importance of expanding data collection beyond academic outcomes to include contextual, school-level, and non-cognitive dimensions (e.g., school engagement, motivation, well-being). Without such information, it is difficult to monitor progress toward more inclusive and equitable learning environments, or to evaluate the impact of policies and practices aimed at reducing educational inequalities. The mapping also underscores the value of linking datasets with different scopes and information, within and across systems, to support a more holistic understanding of students' educational experiences and outcomes.

To design effective and equitable education policies, policymakers need access to timely, high-quality longitudinal data. Promoting a culture of data use and addressing the structural, legal, and technical barriers to access, interoperability, and documentation is essential to improve the effectiveness, equity, and accountability of education systems across Europe. Doing so will empower researchers, decision-makers, and educational stakeholders to generate actionable insights based on robust data, and co-design strategies to reduce educational inequalities.

Report Structure

Section 1 introduces the rationale, objectives, and research questions behind the mapping. Section 2 outlines the shared theoretical and methodological framework, detailing data sources, inclusion criteria, validation procedures, and the dataset classification system. Section 3 presents the results of the mapping, including the number, type, and distribution of eligible datasets. Section 4 provides a descriptive analysis of the mapped datasets, highlighting their research potential, structure, and cross-country comparability. Section 5 introduces the *LINEup Education Data Explorer*, describing its design and key functionalities for dataset discovery and use. Section 6 discusses the feasibility of data harmonisation, identifies common challenges, and lays the groundwork for the upcoming *Deliverable D3.2*. Section 7 concludes the report with key findings and recommendations. After the report bibliography (section 8), the Annexes include, in order: the full list of included and excluded datasets, the expert questionnaire, the online form for dataset metadata collection, and a glossary of technical terms.

Infographic - Visual Summary of Key Findings

The infographic below provides a visual synthesis of the key insights from the mapping exercise. It summarises the number and types of datasets identified, their scope, structure, and content, as well as main barriers to access and use. It also illustrates what existing longitudinal data in Europe can and cannot support in the study of educational inequalities.

PROJECT OVERVIEW

The LINEup project conducted **the first EU-wide mapping of existing longitudinal datasets** on student learning outcomes within primary and secondary school education.

104 Datasets Mapped*



24 European Countries (out of 32)



Focus: Student Learning Outcomes and Academic Performance

Data Types: Longitudinal (i.e. panel, repeated cross-sections, pseudo-panels)

*One U.S. dataset is included as a best-practice example, but excluded from the descriptive statistics

RESEARCH POSSIBILITIES

The potential of a dataset depends on the research questions we ask. Most support descriptive research on individual traits. But explaining inequalities through school-level or comparative analysis remains difficult. Even when access is granted, poor documentation limits reusability.



WHAT WE CAN DO

Research Possibilities

- Analyse learning trajectories (especially via test scores or grades)
- Study socio-demographic effects (presence of basic ascriptive variables)
- Track grade retention and transitions (coverage of key educational transitions)
- Compare performance by school type/track (availability of basic school-level identifiers)



WHAT WE CANNOT DO

Analytical Limitations

- Study peer/class composition (lack of class-level data)
- Assess impact of teacher quality or turnover (limited variables)
- Analyse school segregation or territorial inequalities (no geo-referencing)
- Compare non-cognitive skills cross-country (heterogeneous or missing data)

METHODOLOGY | A Three-Step Strategy

1. DATASET IDENTIFICATION

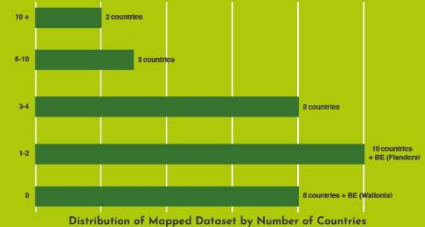
- Literature Review (**43**)
- Expert Questionnaire sent to 315 Experts from 28 Countries (**70**)
- Targeted Web-Search & Partner Knowledge (**108**)
- Collaboration with other projects (**8**)

2. DATASET SELECTION

- 229** datasets **screened**
- 104** datasets **included**
- 125** datasets **excluded**

3. DATASET ANALYSIS

Focus on Data Structure, Time Span, Indicators & Variables, Quality, Accessibility



KEY FINDINGS | By the numbers

HOW are datasets designed, linkable, and accessible to users?

- 68%** are panel data (individual tracking over time)
- 22%** are repeated cross-sectional (with individual identifiers)
- 10%** can be used as pseudo-panel
- 47%** are survey data
- 19%** are administrative data
- 34%** are hybrid data (survey + admin)
- 55%** allow linkage to other data sources
- 11%** are open access
- 50%** are accessible upon request
- 26%** are highly restricted or not accessible

WHO is included in the datasets?

- 32%** cover the entire student population
- 45%** use random samples
- 23%** use non-random samples
- 50%** with sample < 5,000 students
- 29%** with sample > 10,000

WHAT do the datasets measure?

- 77%** include standardised tests
- 69%** include school grades
- 53%** include both tests and grades
- 64%** cover both numeracy and literacy skills
- 36%** cover foreign language skills
- 30%** cover science skills
- 69%** cover other domains skills
- 95%** include information on students' gender
- 68%** include information on foreign birth country
- 74%** include information on parental education
- 61%** include information on parental occupation
- 30%** include information on teachers
- 23%** include information on classroom composition

WHERE do the datasets apply?

- 62%** have national coverage
- 38%** are sub-national.
- Most datasets come from Western & Northern Europe

WHEN were data collected and for how long?

- 23%** span 14+ years
- 20%** cover only 2 years
- 51%** are still ongoing
- 47%** cover both primary and secondary education.
- Most frequently: grades 8, 9, 10
- 30%** begin in kindergarten

The value of longitudinal data lies in its responsible, effective use.
LINEup contributes to a more transparent and visible education data landscape in Europe.



1. Introduction

Accurate, comprehensive, and **high-quality data on education and student learning outcomes** play a critical role in assessing the performance of education systems and understanding trends in student achievement and educational trajectories. By leveraging such data, educational stakeholders can gain clearer insights into which student populations are being left behind and where performance gaps are emerging. This knowledge is essential to design interventions aimed at reducing disparities and promoting equity across educational contexts - an objective increasingly reflected in European policy frameworks that call for better data to support inclusive education (Council of the European Union, 2022).

The *Basic Skills and STEM Action Plan* (European Commission, 2025) calls for improved monitoring systems, targeted support schemes for underachieving learners, and inclusive environments to strengthen foundational competences at all levels of education. Similarly, the latest version of the *Education and Training Monitor* (European Commission, 2024) highlights the urgent need for high-quality data to track equity and inform evidence-based policymaking. Complementing this, the *Key Data on Early Childhood Education and Care in Europe* (European Commission, 2025) underscores the importance of robust monitoring mechanisms to ensure inclusive and high-quality education from the early years onwards.

Data also serve as a **valuable accountability mechanism**, enabling educational stakeholders to assess whether increased public funding and other educational investments are translating into improved outcomes - a principle clearly articulated in the *Quality Investments in Education and Training* expert report (European Commission, 2022), which highlights the need to ensure that spending supports effectiveness and equity in education.

In this context, a growing number of countries have begun to systematically collect diverse data to monitor student progress, teacher effectiveness, and school performance.

International initiatives, such as the OECD's Programme for International Student Assessment (PISA), the Progress in International Reading Literacy Study (PIRLS), and the Trends in International Mathematics and Science Study (TIMSS), further enrich this landscape by offering comparative information into student learning outcomes across countries. However, these datasets are typically cross-sectional or repeated cross-sectional, capturing information from different samples of students at each wave². As such, they do not support individual-level tracking over time and are therefore limited in capturing the full longitudinal dynamics of learning or the evolution of inequalities therein. Despite their potential, existing longitudinal datasets in education remain fragmented, lack harmonisation, and are still insufficiently used to inform research and policymaking.

Longitudinal data offers unique advantages for studying educational inequality. By following the same individuals over time, they allow researchers to trace learning trajectories and analyse how students' characteristics, life events, and contextual factors interact to influence outcomes. Such datasets can reveal when and how disparities emerge, persist, or widen, offering a richer understanding of the dynamics behind social gaps in school achievement and status attainment³.

² Although TIMSS remains primarily a cross-sectional assessment, the International Association for the Evaluation of Educational Achievement (IEA) has introduced an optional longitudinal study component in TIMSS 2023, following the same student cohort one year later (from Grade 4 to Grade 5 and from Grade 8 to Grade 9). This approach is experimental and not yet part of the standard international data collection cycle.

³ In the context of this project, longitudinal data refer to datasets that follow the same students or educational units (e.g., classrooms or schools) over multiple time points, allowing for the analysis of change, causality, and developmental patterns in educational trajectories (see Annex 9.4 for further details).

In addition, longitudinal data offer key advantages for identifying causal relationships. Unlike cross-sectional data, they allow researchers to establish temporal ordering, testing whether changes in independent variables (e.g., educational support) precede changes in outcomes (e.g., academic performance). They also make it possible to examine within-person change by analysing how individuals change over time, rather than relying only on between-person differences. This strengthens the ability to draw conclusions about what influences educational outcomes.

The LINEup project aims at identifying key success factors for quality education for all. Within this framework, addressing educational inequalities is both a timely and widely acknowledged priority. Given the crucial role that longitudinal data can play in identifying trends and informing the design of effective interventions, the **LINEup mapping of existing longitudinal data** (WP3) plays a crucial role in the project. It focuses on the identification and analysis of national and regional longitudinal datasets on students' learning outcomes at the primary and secondary education levels (ISCED 1, 2, and 3), thus laying the foundation for improved data accessibility, comparability, and use in both research and policymaking.

The mapping covers **32 countries**: the 27 EU Member States, three European Economic Area (EEA) associated countries (Iceland, Liechtenstein, and Norway), as well as Switzerland and the United Kingdom. This selection reflects the primary geographic scope of the LINEup project, which focuses on European education systems while ensuring representation of diverse policy frameworks, administrative structures, and data collection traditions within the broader European context. The inclusion of EEA countries, Switzerland, and the UK allows for a more comprehensive understanding of longitudinal data availability and use beyond the EU alone. In addition to this core group, selected practices from outside Europe - particularly from the United States - were considered during the methodological design. Although the U.S. was not fully included in the mapping, its experience with large-scale longitudinal educational datasets provides valuable insights into their design, implementation, and policy use, serving as a valuable comparative reference.

The mapping is grounded in a **comprehensive desk research**, which builds on and complements the findings of the systematic literature review conducted in the first months of the LINEup project. While the literature review synthesised existing evidence on factors contributing to educational inequalities, identified relevant longitudinal studies, and reviewed the methods used to analyse this kind of data, the mapping shifts the focus to the datasets themselves: **what data are available, where they are located, and for what purposes they are collected and used**.

More specifically, the main objectives of WP3 were:

- Build on the findings of the systematic **literature review** (WP2) on longitudinal studies related to inequalities in education, with particular attention to the datasets underpinning the reviewed empirical studies. The systematic review included a dedicated section on data sources, allowing the identification of key datasets on which existing evidence is based⁴.
- Design, distribute and analyse a **targeted expert questionnaire** to gather information on the existing longitudinal datasets across the 32 countries covered by the project.
- Carry out **desk research** to identify additional national and regional longitudinal datasets not captured through either the literature review or the expert questionnaire.

⁴ For further information see: Kampylis, P., Fragkiadaki Theodoroulea M., Kandila, M., Cholezas, I., Mobilio, V., Sampson, D., Lievore, I., Mauro, V., Gunzelmann, S., Lanoë, M., Maurya, P., Moulin, L., Ortiz, L., Passaretta, G., & Moreira, P. (2024). Tracing Educational Inequalities in Primary and Secondary Schools - Insights from a Systematic Review of Longitudinal and Repeated Cross-sectional Studies. LINEup Project - Deliverable 2.1.

- **Collect and analyse structured information** on the identified datasets, using a shared online form completed by project partners to ensure standardised documentation and comparability.
- **Examine educational outcomes** and identify persistent and emerging trends of educational inequalities.
- Produce a **research report and infographic** on the key outcomes, to communicate the main findings in an accessible and engaging format.

Through this multi-step approach, WP3 aimed at strengthening the evidence concerning how educational inequalities emerge and persist over time. The insights generated through the mapping exercise feed into the development of the *LINEup Education Data Explorer*, an online platform designed to enhance the discoverability, comparability, and future use of longitudinal education data by researchers and educational stakeholders. By providing a comprehensive overview of the longitudinal data landscape in Europe, this work supports more integrated, data-informed approaches to educational research and policymaking at both national and European levels.

This report presents the outcomes of this mapping activity. It is structured as follows:

Section 1 introduces the aims and scope of the mapping and explains its relevance for understanding educational inequalities.

Section 2 outlines the theoretical and methodological framework, including key definitions, inclusion criteria, and the three-step mapping strategy.

Section 3 details the mapping process and its results, from preparatory activities and data source identification to the structured documentation of datasets and the synthesis of mapping outcomes across 32 countries.

Section 4 presents the dataset analysis, including validation procedures, classification of dataset types, descriptive statistics, and the identification of analytical potential and barriers.

Section 5 describes the structure and functionalities of the *LINEup Education Data Explorer*, the project's online mapping tool.

Section 6 discusses the research potential and limitations of the mapped datasets, offering a framework to assess their suitability for different analytical goals and highlighting key challenges to cross-country comparability.

Section 7 outlines the main conclusions and implications for research, policy, and practice, and introduces the next steps in the project, including the feasibility study on data harmonisation and the continued development of the Data Explorer.

Sections 8 and 9 include references and annexes, such as the full list of included and excluded datasets, the expert questionnaire, the dataset description form, and a glossary of technical terms.

2. Theoretical and Methodological Framework

This section presents the framework that informed the LINEup mapping of longitudinal datasets on student learning outcomes across Europe. Given the diversity of national data infrastructures and the interdisciplinary nature of the LINEup consortium, the establishment of a shared set of key concepts, inclusion criteria, and analytical procedures was a necessary precondition for ensuring rigour, transparency, and comparability across countries.

The framework integrates **theoretical constructs with a systematic and collaborative approach** for the identification, selection, and documentation of relevant datasets. It rests on three core pillars: (i) a common set of definitions that underpin the mapping, including the meaning and definition of educational inequality, the distinction between datasets and databases, and a typology of longitudinal data; (ii) a structured identification and analytical strategy that integrates insights from a systematic literature review, expert consultation, and metadata collection; (iii) a set of clear and consistent criteria for dataset inclusion, designed to ensure policy relevance and research utility.

Together, these elements provide a **transparent and rigorous foundation** for the mapping activity, allowing for consistent identification, classification, and cross-country analysis of longitudinal data sources on student learning outcomes and inequalities therein.

Key Concepts and Definitions

To ensure conceptual alignment of partners with diverse disciplinary and institutional backgrounds, the first step in the mapping activity was to establish clear and consistent definitions of the core concepts underpinning it. This was essential for aligning perspectives and creating a common foundation for collaboration.

This section outlines the key definitions agreed upon within the consortium to support the broader objectives of the mapping. It begins with a clarification of what is meant by educational inequality, then moves on to explain the distinction between a dataset and a database and concludes with a typology of longitudinal data. These conceptual clarifications served to guide the identification, classification, and analysis of relevant data sources throughout the process.

The term **educational inequality** refers to disparities in educational opportunities and outcomes *between* groups of individuals (Carnoy & Rothstein, 2013), encompassing differences in academic achievement, learning experiences, and life chances (UNICEF Innocenti Research Centre, 2012). These groups are typically defined by **social categories** such as gender, socio-economic status, migration background, ethnicity, or geographical location. However, inequalities also exist *within* these groups and often result from the interaction of multiple, overlapping factors. Educational inequality is thus a complex, multi-faceted phenomenon, present across all education systems - both in Europe and globally. It is reinforced by intersecting systems of disadvantage (Crenshaw, 1989; Walby, 2007), which produce cumulative effects on students' educational trajectories through their interaction with institutional contexts.

When discussing educational inequalities, it is crucial to distinguish between inequality of educational opportunities and inequality in academic achievements. While these concepts are closely interconnected, they address distinct aspects of educational inequality and require targeted policy interventions:

- **Inequality of educational opportunities** refers to unequal access to quality schooling, resources, and educational pathways or tracks. This concept emphasizes structural barriers

that prevent individuals from having equal chances to benefit from education (Boudon, 1974; UNESCO, 2021), highlighting the need to address systemic disparities and ensure equitable access to high-quality education for all (OECD, 2018).

- **Inequality in educational outcomes**, both in terms of achievement and attainment (conceptually different but considered together here), relates to differences in measurable educational outcomes, such as grades, standardized test scores, or graduation rates. These disparities reflect broader inequities in learning processes and environments, as well as the cumulative effects of unequal opportunities throughout a learner's educational journey.

These different types of inequality are typically analysed through indicators based on **students' learning outcomes**, which are the focus of most longitudinal datasets examined in this report. Learning outcomes refer to the knowledge, skills, attitudes, and values that students are expected to acquire and demonstrate as a result of their educational experiences (European Commission, 2017; OECD, 2003). In the context of this mapping, they are typically operationalised through standardised assessment results or performance indicators that allow for comparisons over time and between groups. These indicators serve as the primary lens through which disparities in achievement are assessed. However, while this mapping focuses on learning outcomes, longitudinal data may also include information on educational inputs (such as teacher qualifications) and processes (such as classroom composition and teaching practices), which can help explain observed variations in outcomes over time.

Longitudinal data include repeated observations on the same individuals (e.g. students) or units (e.g. classrooms, schools) over time (Singer & Willett, 2003; OECD, 2017). Such data allow researchers to 1) track the same individuals/units across multiple time points; 2) analyse within-unit changes over time; 3) control for unobserved heterogeneity, since each unit effectively serves as its own control. Longitudinal data are typically organized so that each row corresponds to an individual- or unit-time observation. Depending on the dataset, longitudinal data can follow different levels of analysis - such as students, classrooms, schools, or even regions - enabling multi-level analyses of educational dynamics over time.

Longitudinal data can be collected and organised in different ways. For the purpose of this mapping, an **inclusive classification of longitudinal data types** was adopted, encompassing not only panel studies but also repeated cross-sectional datasets (with a student identifier) - including those that can be used to construct pseudo panels enabling time-based analysis at the cohort, school, or system level:

- **Pure longitudinal data:** data repeatedly collected by following the same individuals over time (panel studies).
- **Repeated cross-sectional data:** data collected from different individuals or units (i.e. classrooms, schools, etc.) at multiple points in time that allow the construction of longitudinal data through a unique identifier for its units of analysis.
- **Pseudo-panel data:** constructed from repeated cross-sectional data by grouping individuals or units into synthetic cohorts (e.g., by birth year or education level) and tracking average outcomes over time.

These three types of longitudinal data can differ not only in structure, but also in terms of data sources and data collection methods. Some datasets are **survey-based**, typically designed for research purposes; others rely on **administrative records**, primarily developed for monitoring and governance. In some cases, these two sources are **combined in hybrid configurations**, where survey data are linked with administrative records to enrich the scope and quality of available information.

While these categories are referred to as *types* for the sake of classification, it is important to recognise that longitudinal structures can emerge from different data collection strategies. These include: (1) data purposely collected to produce panel structures, and (2) data not originally designed for longitudinal analysis - such as repeated cross-sections or administrative sources - that can nonetheless be grouped or linked into coherent temporal trajectories. Repeated cross-sectional data were included in this mapping when they enabled comparisons across multiple waves or served as the basis for constructing pseudo-panels. Conversely, international studies such as PISA, PIRLS, and TIMSS were excluded due to their visibility, extensive documentation, and wide availability.

Since the collection of longitudinal data may vary depending on the data type, it was also necessary to distinguish between datasets and databases.

A **dataset**, for the purposes of the LINEup mapping, is defined as a coherent set of data resulting from a structured process of data collection that enables longitudinal analysis across at least two points in time. Datasets may differ in scale and structure: they may originate from dedicated longitudinal studies, repeated administrative collections, or linked sources. In this mapping, each dataset is considered as **one unit of analysis**. In cases where the same data collection process has produced multiple independent longitudinal cohorts (e.g., students entering school in different years), each cohort is treated as a separate dataset - provided it allows longitudinal tracking over time.

A **database** refers to a technical or organizational infrastructure that supports the management of one or more datasets, particularly in complex or long-term data systems. It may include multiple cohorts, data linkage across sources, a structured system for ongoing updates and access. In the LINEup mapping, the term database is used descriptively to indicate the technical or organizational infrastructure of datasets but is neither a separate unit of analysis nor the object of the analysis.

Taken together, these conceptual definitions were not only necessary for analytical clarity but also provided a shared foundation for the mapping activity by guiding consistent choices during the identification and classification of datasets across different national contexts.

Methodological Approach

The overarching goal of the mapping is to gather comprehensive information on existing longitudinal datasets on student learning outcomes across Europe. The mapping approach combines insights from a systematic literature review with a structured, three-step strategy designed to progressively refine the scope of investigation.

1. Literature-informed foundation

The mapping activity builds on the systematic literature review carried out in WP2, which analysed both academic and grey literature on educational inequalities from a longitudinal perspective. In addition to identifying key drivers and patterns of inequality, the review covered empirical studies that relied on longitudinal data, thus providing an initial inventory of relevant datasets. These datasets were compiled into a shared document accessible to all partners, who were then invited to enrich the list based on their national expertise and contextual knowledge of education data sources.

2. A three-step mapping strategy

The mapping process followed a *funnel strategy*, progressively narrowing the focus based on the availability and quality of longitudinal data identified across countries (see Figure 1). It included three main steps:

- Data Sources Identification:** step one of the mapping process (see section 3.1) was based on a web-based questionnaire sent to national education specialists in countries potentially relevant to the project. The objective was to expand the initial inventory of longitudinal datasets compiled through the systematic literature review by leveraging expert knowledge of national and regional data sources related to student learning outcomes, performance, and cognitive development. All datasets identified through the questionnaire were recorded in a shared working document accessible to all partners. The document was organised by country and included key metadata such as dataset name, acronym, thematic focus, and years of data collection. To increase transparency, each dataset was flagged using a colour code to indicate its source: black for datasets identified directly by consortium members, red for those that emerged from the systematic literature review, and green for those collected through the expert questionnaire.
- Initial mapping:** Step two of the mapping process (see section 3.2) involved structured web searches conducted by consortium researchers to complement expert input and verify the existence, characteristics, and accessibility of the datasets identified in the previous step. This phase focused on checking compliance with the project's inclusion criteria and documenting the scope, content, and accessibility of all potentially relevant datasets across the 32 countries covered by the project. Each partner was responsible for a specific group of countries (see Table 1) and systematically screened the datasets accordingly. Those that did not meet the inclusion criteria were moved to a dedicated “excluded” section in the shared working document, along with a brief justification for their exclusion. This phase was critical to ensure methodological consistency, transparency, and traceability across countries and partners.
- In-depth mapping:** Step three of the mapping process (also detailed in Section 3.2) involved a detailed metadata collection exercise for all datasets retained after the screening phase. For each eligible dataset, partners completed a structured online form capturing standardised information across countries, including scope, longitudinal design, target population, learning outcomes covered, frequency of data collection, access conditions, and included variables. This phase enabled an in-depth assessment of dataset quality, the robustness of the longitudinal structure, the availability of relevant indicators, and overall accessibility. It was followed by a multi-step validation process involving peer review and manual re-assessment to ensure accuracy, completeness, and cross-country consistency. Where possible, initial reflections were added on cross-country comparability and the potential for future harmonisation of key variables. While this step applied to all eligible datasets, regardless of country, the subsequent work on data analysis and harmonisation will focus on a subset of datasets, including those from the six countries represented in the LINEup consortium (DE, EL, ES, FR, IT, PT). The resulting documentation constitutes the analytical foundation of the present report and lays the groundwork for the upcoming *Feasibility Study* foreseen in WP3, to be published as a separate deliverable in January 2026. This study will further assess the selected dataset potential for cross-country comparative research and variable harmonisation.

Figure 1: The funnel strategy of the mapping.

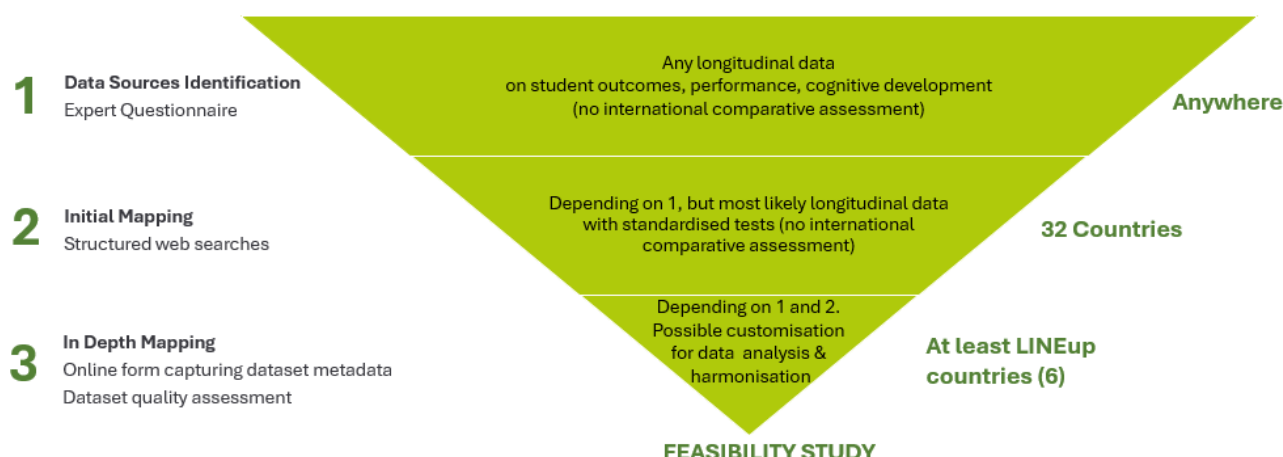


Table 1: Distribution of countries among partners.

Partners	Countries to be covered
PHKA	AT, BG, DE, LI + UK (in cooperation with UPF)
KEPE + PIREUS	CY, DK, EL, NL, HU
UPF	ES, IE, MT, U.S.* + UK (in cooperation with PHKA)
INED	BE (Wallonia), BE (Flanders), FR, LU
UTAD	EE, LT, LV, PT
FPS	IT, PL, NO, HR
UNIMI/BICOCCA	SE, CZ, SK, CH
UNIMC	FI, IC, RO, SI

* The United States (U.S.) was not part of the formal coverage of the mapping. However, it was considered selectively to inform the mapping design and provide relevant examples of best practices in longitudinal data collection and use.

Dataset Inclusion Criteria

A crucial step in the mapping activity was the definition of the criteria guiding the inclusion of longitudinal datasets. These criteria were developed through **iterative discussions** and knowledge-sharing among consortium partners during several monthly project meetings and were **progressively refined** to ensure clarity, consistency, and alignment with the project's objectives.

The final criteria for dataset inclusion are:

- Data must be **longitudinal in nature** – this includes purely longitudinal survey data, pseudo-longitudinal (repeated cross-sectional) survey data, and/or administrative data⁵.

⁵ Datasets derived from Randomised Controlled Trials (RCTs) were considered eligible only when they extended beyond a basic pre/post measurement design - specifically, when they included at least three waves of data collection and demonstrated some degree of representativeness at the national, regional, or local level.

- Data must **pertain to the primary and secondary levels of education** (excluding higher education).
- Data must focus on **students' learning outcomes** (basic skills and others) typically measured through standardized test scores; and/or on **academic performance indicators** (e.g. school grades, exam scores, drop-out, truancy, grade retention, school trajectories); and/or on **other measures of cognitive and non-cognitive development** (e.g. socio-emotional and transversal skills, students' school engagement, sustained attention, etc.).
- Data must **originate from EU Member States (27), European Economic Area Associated Countries** (Iceland, Liechtenstein, and Norway) **and/or Switzerland and the UK**.
- Data must **be representative at regional/national level** or at least based on a large-scale sample.
- **Data collection** must be **ongoing or recent** - if concluded, the latest wave must have taken place no earlier than 2014.

International comparative assessments such as PISA were considered out of scope as they are already well-known and widely used in existing research. This mapping aimed to identify and valorise national or regional longitudinal data sources less visible in the international literature but equally valuable for understanding educational inequalities from a longitudinal perspective.

These inclusion criteria ensured a consistent and policy-relevant selection of datasets across countries, and enabled a consistent cross-country comparison in the subsequent analytical phases of the project.

Challenges and Limitations

The mapping presented in this report represents the first attempt to systematically summarise and document information on existing longitudinal datasets on student learning outcomes across Europe. However, several methodological limitations should be acknowledged.

First, the identification process relied primarily on desk research and expert input, which - despite being extensive and systematic - may have **missed some datasets**, particularly in countries with limited data transparency, poor online documentation, or fragmented institutional oversight. In some cases, restricted access to official sources further constrained visibility into national data infrastructures. To mitigate this limitation, the Consortium contacted experts from 28 of the 32 countries covered through the expert questionnaire, gathering local insights to help identify less visible datasets. Moreover, the LINEup Education Data Explorer, presented in this report as a visual representation of the mapping (Section 5), is conceived as an ongoing process, where users and experts can suggest new datasets, signal mistakes, and contribute further documentation.

Second, a substantial number of datasets included in the mapping **featured incomplete, outdated, or inconsistent metadata**, which complicated efforts to assess their structure, content, indicators, and analytical potential. Although a standardised documentation template was used, the availability and quality of underlying information varied considerably across countries and sources. This limitation was acknowledged during data collection, with partners flagging unclear or missing information to enhance transparency for future users. As a result, the "Not clear" label was introduced as an explicit metadata category - highlighting gaps in documentation and offering a valuable dimension for further analysis of dataset quality and transparency.

Third, some datasets were described using **national terms, often in the local language**, and with limited clarity on key dimensions - such as the nature of follow-up, the continuity of observations, or the measurement of outcome variables. Since partners were not familiar with the languages of all

countries included in the mapping, this posed additional challenges and required a degree of interpretation during classification and validation. To address this issue, the research team applied a shared set of operational definitions systematically, ensuring consistency across partners through a common classification framework, dataset description template, and detailed guidelines.

Fourth, **classification inevitably involved subjective judgement**, particularly where dataset documentation was incomplete or unclear. Moreover, no external validation by third parties was conducted during the mapping phase, and quality checks were limited to internal review processes. Unclear cases were discussed in regular coordination meetings to enable collective resolution of uncertainties, and each dataset was reviewed by at least two researchers, with three rounds of internal cross-checking (local, central, and joint review) and consensus-based decision-making. Given that some inconsistencies or interpretative choices may persist, in the coming months data owners will be invited to review and, where necessary, suggest corrections to the classification and codification of their datasets. This feedback will be documented in a dedicated appendix of the upcoming Feasibility Study (Deliverable 3.2), providing an opportunity for iterative refinement and targeted adjustments where needed⁶.

Despite these challenges, the mapping provides a robust and transparent foundation for further analysis and highlights key areas where improved metadata and harmonisation efforts are most urgently needed.

⁶ All validated changes will also be integrated into the LINEup Education Data Explorer to maintain alignment between the reports and the online mapping platform.

3. Mapping of Longitudinal Datasets: Methods and Procedures

This section presents the methodological approach and closes with an overview of the mapping results, which are further explained in Section 4.

Building on the theoretical and methodological framework outlined in Section 2, the mapping employed a multi-pronged strategy to ensure both breadth and depth of coverage across diverse national contexts. As anticipated above, the process combined insights from a systematic literature review, expert consultation through a structured questionnaire, targeted web searches, and detailed documentation of eligible datasets collected using a standardised online form.

In total, 229 datasets were identified. Of these, 104 met the inclusion criteria and were retained for further analysis. The resulting comparative overview sheds light on the availability, scope, and accessibility of longitudinal data across Europe, highlighting both opportunities and gaps for research and policy action.

The section is organised as follows:

- Section 3.1 outlines the preparatory steps undertaken prior to the actual mapping activity, including the creation of shared documents and tools.
- Section 3.2 describes the development of shared procedures through which the mapping was developed, including the protocol outlining the main actions for assessing and describing datasets, the web-search practices, and the in-depth description template.
- Section 3.3 presents the main outcomes of the mapping, including dataset counts, country distribution, and key features. It also summarises the main reasons for dataset exclusion and discusses emerging patterns and challenges observed across countries.

3.1 Preparatory Activities: Identification of Data Sources

Preparatory activities for the mapping of existing longitudinal datasets involved multiple steps, each designed to gather comprehensive knowledge on available data sources across Europe. These steps aimed to expand the scope of the mapping by identifying new and relevant datasets enabling an analysis of educational inequalities from a longitudinal perspective.

In the initial phase, the objective was to compile an extensive list of longitudinal datasets addressing educational outcomes, without immediately applying the inclusion criteria. This open approach - based on the consortium's existing knowledge and the findings from the systematic literature review - enabled the creation of a broad inventory to be screened and refined at a later stage.

Once the preliminary list of identified datasets was shared within the consortium, the next step involved the launch of an expert questionnaire aimed at identifying additional data sources through national-level knowledge and experience. A final activity to enrich the dataset inventory was a call for cooperation with the *Sister Projects* - a group of Horizon Europe projects funded under the same call as LINEup and sharing a focus on mapping longitudinal data and assessing inequalities in education⁷. The initiative promoted exchange and mutual enrichment among projects through the sharing of preliminary findings and data sources. This section presents the data sources identified through these three complementary channels: the systematic literature review, the expert questionnaire, and the collaboration with Sister Projects.

⁷ The LINEup Sister Projects funded under Horizon Europe call HORIZON-CL2-2023-TRANSFORMATIONS-01-06 are: [LEARN](#); [MapIE](#); [REVERS-ED](#); [STRIDE](#).

Insights from the Systematic Literature Review

The LINEup systematic literature review was carried out between February and June 2024. It focused on studies addressing educational inequalities through longitudinal or repeated cross-sectional research designs in primary, lower-secondary, and upper-secondary education - both general and vocational. The process was conducted following the [PRISMA guidelines](#) (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), applying the standard workflow steps of Identification, Screening, and Inclusion to enhance transparency, replicability, and the overall reliability of the findings.

As part of the literature review, the research team systematically searched scientific databases and academic portals to identify both peer-reviewed and grey literature reporting studies conducted in the 32 countries covered by the project.

The screening and inclusion process was organized in three steps. In Step 1, the title, abstract and keywords of 1.399 publications were screened based on five inclusion/exclusion criteria. In Step 2, the full text of 843 publications was screened based on seven inclusion/exclusion criteria. Finally, 157 publications were selected for in-depth analysis in Step 3 through a review matrix. All the publications identified, screened, and analysed were stored in a Zotero library, which will be hosted by the end of the project in the EU Open Repository for EU-funded Research in Zenodo.

The scope of the systematic literature review was to i) identify studies on educational outcomes in school education with a longitudinal and/or repeated cross-sectional research design and the related datasets (RQ1), (ii) identify methods and techniques used for analysing this type of data (RQ2), and (iii) identify and categorize the variables that might be factors or predictors of educational outcomes (RQ3).

In line with these objectives, the review directly supported the mapping of existing data sources (as part of RQ1) by identifying the main datasets underpinning the selected studies. Through the in-depth analysis of 157 publications, the systematic literature review research team identified in total **77 distinct longitudinal or repeated cross-sectional datasets**. While these datasets and the associated studies provided valuable information, they covered only half of the 32 countries targeted by the LINEup project. This is an important finding, since it reflects the uneven availability and visibility of longitudinal and cross-sectional research across Europe and highlights the need to strengthen national data infrastructures to support future research on educational careers.

Appendix A of the Literature Review Report presents a summary of the datasets identified through in-depth analysis. This matrix includes: 1) name of the longitudinal dataset or study; 2) reference publication(s); 3) data collection method; 4) country; 5) geographical coverage; 6) timeframe; 7) target population; 8) demographics⁸.

It is important to note that the inclusion criteria applied in the systematic literature review differ slightly from those adopted in the mapping (see Table 2). Therefore, the detailed matrix presented in Appendix A of the Literature Review Report served as a basis for screening the identified datasets against the inclusion criteria defined for the mapping activity. Datasets were excluded if their most recent data collection occurred before 2014, which was adopted as the threshold to ensure recency

⁸ For further information see: Kampylis, P., Fragkiadaki Theodoroulea M., Kandila, M., Cholezas, I., Mobilio, V., Sampson, D., Lievore, I., Mauro, V., Gunzelmann, S., Lanoë, M., Maurya, P., Moulin, L., Ortiz, L., Passaretta, G., & Moreira, P. (2024). Tracing Educational Inequalities in Primary and Secondary Schools - Insights from a Systematic Review of Longitudinal and Repeated Cross-sectional Studies. LINEup Project - Deliverable 2.1.

and policy relevance. Regarding sample size, only datasets based on large-scale samples were retained. For this mapping, a 'large-scale sample' refers to data collections designed to be representative at the national or regional level, or to cover entire school populations or cohorts, as opposed to small-scale, ad hoc longitudinal studies involving limited or non-representative groups of participants. While the threshold for defining a 'large-scale sample' may vary by country, the mapping prioritised datasets with at least several hundred participants per wave and clear documentation of representativeness.

Due to the differences in inclusion criteria between the systematic literature review and the mapping (see Table 2), of the 77 datasets identified through the literature review, **only 43 datasets were retained in the shared inventory circulated among partners for review**. These 43 datasets were those that, at a preliminary screening stage, appeared to comply with the mapping criteria related to recency, sample size, and relevance. In cases where there were doubts regarding the inclusion of certain datasets - due to partial or uncertain compliance with the established mapping criteria - these were provisionally included and discussed with the project coordinator. Final decisions on their eligibility were deferred to the partner-level review and validation process.

Table 2: Inclusion Criteria - Literature Review vs. Dataset Mapping.

LITERATURE REVIEW	MAPPING
The publication is written in English, French, German, Greek, Italian, Portuguese, or Spanish	Not applicable
Publication is more than four pages, and/or it includes a methodology and results section	Not applicable
The publication reports findings from a study conducted in a country that is covered by LINEup	Data should come from EU Member States (27), European Economic Area Associated Countries (Iceland, Liechtenstein, and Norway), Switzerland and/or the UK.
It is a study with a longitudinal or repeated cross-sectional research design	Data should be longitudinal - e.g. purely longitudinal survey data, pseudo-longitudinal (repeated cross-sectional) survey data, and/or administrative data collected over time.
The study focuses on the education levels targeted by LINEup	Data should regard primary and secondary level of education (no higher education).
It is devoted to the skills/topics targeted by LINEup, i.e., educational inequalities, with a focus on academic achievement/performance and/or school engagement	Data should concern students' learning outcomes (basic skills and others), typically measured through standardized test scores; and/or on academic performance indicators (e.g. school grades, drop-out, truancy, grade retention, school trajectories); and/or on other measures of cognitive and non- cognitive development (e.g. socio-emotional and transversal skills, students' school engagement, sustained attention, etc.)
The methodology reported in the study is of high quality	Not applicable

Not applicable	Data should be representative at regional/national level, or at least based on a large-scale sample.
Not applicable	The data collection should be ongoing or recent (if completed, the latest round should have taken place in the last decade, year 2014 the latest).
Not applicable	International comparative assessments remain out of scope

Note: Some of the inclusion criteria for the literature review and dataset mapping are conceptually aligned but differ in wording to reflect the nature of the sources analysed. While the literature review criteria apply to published empirical studies, the mapping criteria refer to the characteristics of datasets. As such, some criteria are formulated differently, and additional criteria were introduced in the mapping phase to ensure policy relevance and data usability.

Given the consortium's expertise and the richness of the systematic literature review, a substantial number of relevant datasets were already identified at this early stage. The full list was shared with the consortium partners, who were invited to contribute additional datasets for further screening. As the mapping progressed, the shared inventory was continuously updated, to reflect new entries and modifications, thus ensuring that the documentation process remained transparent and easily traceable over time.

Expert Questionnaire: Expanding Dataset Coverage through Stakeholder Input

A questionnaire was developed to collect additional information on longitudinal datasets from the countries covered by the project, with a focus on datasets not previously identified through the consortium's internal knowledge or the systematic literature review.

To compile the list of experts to be contacted, each project partner proposed a number of individuals. Further experts were identified through their involvement in international networks such as EERA (European Educational Research Association), NESSE (Network of Experts on Social Aspects of Education and Training), EENEE (European Expert Network on Economics of Education), as well as other relevant European or international education research platforms. Overall, the activity engaged **315 experts** from 28 countries covered by the project, along with a small number of countries outside scope of the project⁹. The experts represented a variety of institutions. Universities accounted for the majority of respondents (62%), followed by research centres or institutes (20%), both public and private. An additional 15% were affiliated with public institutions outside the research sector - such as government agencies or educational authorities - while the remaining 3% represented other types of organisations, including NGOs and think tanks.

The **questionnaire** was implemented using LimeSurvey¹⁰. The introductory page provided an overview of the LINEup project, explained the purpose of the expert consultation, and outlined the

⁹ Experts were contacted in the following 28 countries covered by the LINEup mapping: Belgium (Wallonia and Flanders), Bulgaria, Switzerland, Cyprus, Czechia, Denmark, Estonia, Germany, Greece, Spain, Finland, France, Hungary, Iceland, Ireland, Italy, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, and the United Kingdom. In addition, experts based in non-mapping countries such as the United States, Moldova, Turkey, and Canada were also contacted.

¹⁰ LimeSurvey (<https://www.limesurvey.org>) is an open-source online survey tool used for designing, distributing, and managing questionnaires in a secure and GDPR-compliant environment.

inclusion criteria established by the consortium. It also included a description of the questionnaire's structure – made up of three main sections - and a privacy policy statement with a consent form.

The **first section** invited experts to suggest national, regional, or local longitudinal datasets concerning primary and/or secondary education. To streamline the process and avoid duplications, a country-specific list of datasets already identified by the consortium was made available in the form. The **second section** asked experts to recommend other individuals who could be contacted as experts for the same purpose. The **third and final section** collected additional information about the respondent's academic activity, including any relevant publications that could be considered for the systematic literature review. Experts were also asked whether they wished to subscribe to the LINEup newsletter(s) to receive project updates. The full questionnaire is available in Annex 9.2.

The expert questionnaire was sent via email in three waves. All messages copied the consortium member who had suggested the expert. In the first wave (between July 23 and July 29, 2024), emails were sent to 275 experts. This initial round yielded 104 responses, including 56 fully completed questionnaires, resulting in a response rate of approximately 20%. Among the incomplete responses, eight provided usable information, bringing the total number of questionnaires with useful input to 64. As a result of the first wave, information was collected on 55 new datasets, which were subsequently screened for relevance, compliance with inclusion criteria, and potential duplicates. In addition, 35 new experts were suggested by respondents, of whom 28 had not been previously identified. The questionnaire also yielded 48 suggested publications, which were later reviewed to assess alignment with the literature review's inclusion criteria and to eliminate duplicates.

The second wave, administered a month later (on August 29), consisted of a reminder sent to the 218 experts who had not replied to the initial invitation, along with a first-time invitation to 28 new experts - 25 of whom had been suggested by respondents to the first wave, and 3 identified by consortium members. At this stage, the total number of experts contacted reached 303. Due to the summer break, many invitations remained unanswered. However, after three further weeks the questionnaire had received 159 total responses, including 84 fully completed ones - bringing the response rate to approximately 24%. This wave yielded information on 64 additional datasets, as well as 50 expert suggestions, 34 of whom had not previously been identified. The newly suggested experts expanded the pool of respondents, and their responses led to the identification of 73 additional publications, which were later reviewed for potential inclusion in the literature review.

The third and final wave was launched a month later, on September 24th. In this round, a reminder was sent to 178 experts who had not responded in the previous waves, and the questionnaire was newly sent to 34 additional experts who had been suggested in the meantime. In total, 315 experts were contacted over the course of the activity, including 41 newly suggested contacts. The questionnaire remained open until October 4th, making the overall response window nearly **three months long**. All responses were collected and processed in accordance with GDPR regulations and the project's Data Management Plan.

Experts from 28 of the 32 countries covered by the LINEup mapping were engaged in the expert questionnaire, along with researchers from a few additional countries beyond the project's scope. However, the questionnaire did not lead to the identification of datasets in eight of the countries covered by the project (Austria, Croatia, Liechtenstein, Malta, Romania, Slovakia, Slovenia, and Sweden), despite efforts to reach out to relevant experts. For these remaining cases, further outreach was pursued, particularly through the project's Advisory Board Members¹¹ and Sister

¹¹ The Advisory Board of the LINEup project brings together experts from national authorities, international organisations, and academia. Their support was key in guiding and complementing the dataset mapping. More information at: <https://www.lineup-project.eu/project/advisory-board/>

Projects (see following section) - with the understanding that, in some of these countries, the absence of suggested datasets may reflect a limited availability or visibility of longitudinal data sources.

The questionnaire led to **193 total responses**, including 107 fully completed questionnaires, 9 partially completed ones, and 77 incomplete responses. The overall response rate was 37% when including partial responses, and 34% considering only fully completed questionnaires. This activity significantly contributed to the mapping effort, resulting in the addition of **70 suggested datasets**. Following an initial screening to remove duplicates, all were incorporated into the consolidated inventory of identified datasets (see Section 3.2 for further details)¹².

Beyond enriching the dataset inventory, the expert consultation also helped validate the consortium's mapping strategy, and brought greater contextual knowledge on national data infrastructures, which informs the subsequent phases of the project.

Sister Projects: A Call for Cooperation to Strengthen Dataset Mapping and Analysis

In addition to the systematic literature review and national-level expert consultation, further data sources were identified through a **dedicated cooperation initiative** with the Sister Projects funded under the Horizon Europe call HORIZON-CL2-2023-TRANSFORMATIONS-01-06, focused on the mapping of longitudinal data and the assessment of inequalities in education, training, and learning achievements. The goal of this initiative was to foster a collaborative environment among the five projects selected under the same topic, enabling mutual exchange, knowledge sharing, and the creation of synergies.

The first collaborative encounter took place during the LINEup Opening Event in Germany (20 September 2023), which featured a panel titled *Comparative Research on Educational Inequalities - Challenges, Key Findings, and Lessons Learned*, alongside a networking lunch designed to facilitate informal exchange between project teams. The panel provided an opportunity to reflect on key insights from previous EU-funded projects (e.g. eduLIFE, PIONEERED, LifeTrack), while the networking activities encouraged the development of collaborative actions between the so-called sister projects ([LINEup](#), [LEARN](#), [MapiE](#), [REVERS-ED](#), [STRIDE](#)) particularly in the area of dissemination and research uptake¹³.

This effort was formalised through the *LINEup Call for Cooperation: Establishing Synergies with Relevant Projects*, a document shared with the Sister Projects to articulate the objectives of the collaboration: promoting networking, drawing on previous research outcomes, and enriching shared knowledge on educational inequalities. The document also emphasized the importance of engaging a broad range of stakeholders – including schools, foundations, public authorities, and research institutions – and outlined possible cooperation actions such as joint events, exchange of data sources and findings, co-authored publications, and shared dissemination activities.

Cooperation meetings take place regularly every six months, either online or in person, and involve the project managers and scientific coordinators of all five projects. These meetings offer a valuable opportunity to build a shared understanding of key concepts, methodological challenges, and data access strategies across projects. Exchanges with Sister Projects also enabled the identification and description of a few additional data sources that had not emerged during the earlier stages of our

¹² In addition to datasets, the questionnaire also yielded 84 suggested publications - subsequently reviewed for potential inclusion in the literature repository - and 77 newsletter subscriptions, later integrated into the LINEup mailing list.

¹³ For further details on the LINEup Opening Event see: <https://www.lineup-project.eu/opening-event-germany-highlights/>

mapping process. For example, the Gothenburg educational longitudinal database (GOLD) dataset in Sweden – which provides longitudinal data on students' transitions and learning outcomes – was identified and described thanks to direct input from the MapLE team. As a result, new entries were added to the consolidated list of longitudinal datasets, marked with the flag *Sister Project Recommendation* to highlight their origin and acknowledge the collaboration.

3.2 Guidelines for Datasets Description

Following the identification of candidate datasets through multiple channels, the second step of the mapping process focused on assessing their eligibility and, where appropriate, producing a detailed description. This phase was carried out using a shared protocol, and methodological guidelines specifically designed to ensure rigour, comparability, and traceability across countries.

All partners were actively involved in this phase and worked in a coordinated manner through shared documents and repositories, supported by the coordination team. A detailed set of mapping guidelines was provided, including:

- A **step-by-step protocol** outlining the main actions for assessing and describing datasets;
- A set of keywords and recommended practices for **systematic web-search** (especially relevant for countries where no datasets were initially identified);
- A **structured template** (Word file and online form) for the in-depth description of eligible datasets.

These guidelines served not only as a project management tool but also as a scientific instrument for ensuring methodological consistency, enabling cross-country comparison, and promoting transparency in the treatment of borderline or partial cases. They also formalised fallback strategies to address information gaps or access barriers.

The process was designed to ensure both comprehensiveness and traceability, two key principles of scientific research design. All partners systematically documented their search efforts, archived relevant supporting materials (such as codebooks, technical reports, and sample publications), and reported back on the outcomes of their mapping activities. This collaborative and transparent approach enabled robust cross-country comparison and laid the groundwork for the subsequent phase of data analysis and harmonisation.

Step-by-Step Mapping Protocol

To guide the mapping of datasets in a consistent and replicable way, a detailed protocol was developed and shared with project partners. It provided clear instructions on how to validate dataset eligibility, conduct web searches when needed, and produce standardised, in-depth descriptions. The protocol included five key steps:

1. **Review the shared list of identified datasets:** partners began by examining the country-specific list of datasets previously identified through consortium knowledge, literature review and expert questionnaire. This preliminary list included both eligible and excluded datasets, with basic metadata and source references.
2. **Assess compliance with selection criteria:** for each dataset, partners verified compliance with the agreed-upon inclusion criteria (see Section 2 for further details). For those datasets whose inclusion was uncertain (e.g. partial compliance with criteria, outdated documentation), partners were encouraged to document their reasoning and consult the

coordination team for a shared decision. This ensured consistency and transparency in borderline cases.

3. **Conduct web-search and expert outreach:** for countries where no datasets had been identified, partners carried out a structured web-search following the shared guidelines. This involved using predefined keywords, searching in the national language, and consulting general search engines, academic databases, national statistical offices, and data archives. Where necessary, partners also contacted experts who had not responded to the expert questionnaire or sought advice from Advisory Board members¹⁴.
4. **Store supporting documentation:** for each dataset described in depth, partners collected and archived all relevant supporting materials (e.g. codebooks, methodological documents, links, sample publications). When possible, they also downloaded and stored microdata. All documentation was organised in a shared folder, by country and dataset name.
5. **Report mapping activities:** upon completing their assigned tasks, each partner submitted a summary of their mapping activities to the coordination team. This included a list of datasets described through the online form, excluded datasets with reasons for exclusion, and any relevant comments or issues encountered.

This structured and collaborative protocol ensured methodological transparency, reduced cross-country variability, and facilitated the scientific reuse of the mapped datasets.

Keywords and Practices for Systematic Web-Search

For countries where no datasets had been initially identified through previous activities (see step 3 of the Step-by-Step Mapping Protocol), partners were asked to carry out a systematic web-search based on a shared set of guidelines. This procedure aimed to ensure that potentially relevant datasets were not overlooked due to gaps in expert responses or limitations in prior desk research. The guidelines were designed to promote consistency across countries, while allowing for contextual adaptation based on language and national information infrastructures.

The recommended web-search procedure included the following steps:

1. **Define relevant search terms:** Identify keywords that capture the essential features of the datasets sought, including school level, type of data, and learning outcomes¹⁵. A suggested list of keywords is provided in Table 3.
2. **Localise the search queries:** Replace [Country] with the name of the assigned country and translate the search queries into the national language. This step was essential to improve the visibility of locally documented or hosted datasets.
3. **Use diversified sources:** Conduct broad-to-narrow searches using both general search engines (e.g. Google) and academic databases (e.g. Google Scholar, JSTOR). Include the country name in all search queries to narrow the scope.
4. **Consult trusted data repositories:** Focus on recognised sources known for hosting longitudinal or educational data, such as: ICPSR (Inter-university Consortium for Political and Social Research); GESIS - Leibniz Institute for the Social Sciences; CESSDA (Consortium

¹⁴ For more information, see footnote no. 11.

¹⁵ Examples may include: "Longitudinal education data [Country]" "Administrative data school competences [Country]" "Repeated cross-sectional data on basic skills [Country]" "Longitudinal survey data on student learning outcomes [Country]" "Longitudinal datasets primary secondary education [Country]" "Student academic performance data [Country]" "Educational trajectories time series data [Country]" "Standardized test scores longitudinal survey [Country]" "Socio-emotional skills longitudinal data [Country]" "Repeated cross-section education datasets [Country]"

of European Social Science Data Archives) as well as the National statistical offices of assigned countries.

If the systematic web-search did not yield any dataset meeting the inclusion criteria, partners were encouraged to consult additional channels. These included re-contacting non-responding national experts, leveraging contacts within the Advisory Board, or drawing on informal networks across the consortium. This multi-pronged approach was essential to mitigate information asymmetries and address potential language or access barriers.

Table 3: Suggested keywords for systematic web-search.

LINEup Mapping Keywords	Mapping Search Terms
Academic competences and outcomes	Competences / learning / academic / educational outcome / achievement / attainment / performance / basic skills / literacy / numeracy / reading / language skills
Education level	Primary / elementary, compulsory, lower / upper secondary
Longitudinal data, repeated cross-sectional data, administrative data	Administrative data, longitudinal data / panel data / time series data / research / survey / study / analysis / cohort study / design

Note: The search terms listed in the table represent the core terms used during the web search phase of the mapping process. While not exhaustive, they reflect the most frequent and effective search terms for identifying relevant longitudinal datasets.

A Standard Template for In-Depth Datasets Descriptions

When a dataset was deemed to comply with the selection criteria, partners were asked to collect more detailed information to produce an in-depth description. This required a targeted search to retrieve relevant documentation, such as codebooks, methodological notes, technical reports, and publications based on the dataset.

To ensure traceability and future reusability, researchers were encouraged to document all sources used, including web links and reference materials, and to upload them to a dedicated Basecamp folder, organised by country and dataset name.

The project provided detailed recommendations for conducting these in-depth descriptions¹⁶:

- Access the dataset directly, if available via open access or upon request;
- If access was not possible, consult the codebooks and other technical materials;
- Use literature-based documentation, especially from the systematic review;
- When necessary, rely on translation tools to interpret documentation not available in English;
- If documentation was still inaccessible, collaborate with national experts or consortium colleagues to validate or complete the dataset description.

To standardise data collection, researchers used a **word-based template** that mirrored the structure and content of the **final online form**: the former was intended as a working tool to facilitate

¹⁶ The design of the documentation form is consistent with the FAIR data principles (Wilkinson et al., 2016) and the current best practices in research data management.

information gathering and review, while the latter served as the final submission instrument for consolidating the dataset descriptions.

This template, available in Annex 9.3, was divided into six sections:

- Basic information on the datasets identified.
- Basic information on the skills investigated.
- Additional information available in the dataset: students' characteristics.
- Additional information available in the dataset: students' household characteristics.
- Additional information available in the dataset: teachers' characteristics.
- Additional information available in the dataset: school and class characteristics.

The structure of the template reflects the key analytical dimensions that are widely acknowledged in the literature on educational inequality, learning outcomes, and school effectiveness. In line with frameworks such as Coleman's model of school effects (Coleman et al., 1966), the OECD model of educational production (OECD, 2019), and more recent analyses of longitudinal data in education (e.g. Blossfeld et al., 2016; Bukodi & Goldthorpe, 2013), the form was designed to capture information at multiple levels: (i) individual, (ii) household/family background, (iii) teachers; as well as (iv) school/classroom context. This multi-level perspective is essential for understanding how different factors interact over time to shape students' educational trajectories. The second section, focusing on the skills assessed and their measurement, acknowledges the importance of distinguishing between different domains (e.g. literacy, numeracy, science) and types of assessment (e.g. low vs. high stakes, paper-based vs. computer-based), which significantly affect the interpretation and comparability of learning outcomes across datasets (see Mullis et al., 2017; OECD, 2021). Finally, the third section collects meta-information on dataset scope, sampling design, accessibility, and linkability - elements that are crucial for assessing the quality, representativeness, and research potential of each dataset (Scott, 2005; UNESCO-UIS, 2018). This design also aligns with best practices in data documentation and FAIR data principles (Wilkinson et al., 2016).

As explained above, the Word version of the form served as a working tool to collect and refine information. Researchers were encouraged to answer all questions as thoroughly as possible and to avoid vague responses (e.g., "not clear"), by instead providing the best possible estimation or specifying the nature of the information gap.

Once researchers gathered all the information in the Word document, the final in-depth description was submitted via an online form administered through LimeSurvey¹⁷. The form included open text fields to capture any additional comments or dataset-specific features that did not fit neatly into predefined categories.

The resulting output - an Excel file including all eligible datasets described through the online form - constitutes the core empirical product of the mapping activity and provides the foundation for the comparative analysis and visualisation work included in the following/next phases of the project.

3.3 Mapping Results

This section presents an overview of the results of the dataset mapping process carried out across the 32 LINEup countries. A total of **229 datasets** were identified through the combined efforts of literature review, expert consultation, partner knowledge, web-search, and cooperation with sister

¹⁷ LimeSurvey (<https://www.limesurvey.org>) is an open-source online survey tool used for designing, distributing, and managing questionnaires in a secure and GDPR-compliant environment.

projects. Of the 229 datasets identified, approximately **45% were included** in the final mapping and described in depth (104 datasets), while **55% were excluded** (125 datasets), mainly due to non-compliance with the inclusion criteria¹⁸. These proportions highlight the importance of applying consistent inclusion criteria and the uneven availability of eligible longitudinal data across countries.

Annex 9.1 provides detailed information on all mapped and excluded datasets, while Table 4 summarises the number of datasets identified per country, distinguishing between:

- The total number of datasets initially identified by country.
- The number of datasets included and described in depth by country.
- The number of datasets excluded by country.

The figures in Table 4 show considerable variation across countries in the number of datasets identified and retained for in-depth analysis. These differences reflect not only the diversity of national data landscapes, and the varying levels of documentation and accessibility encountered during the mapping process but also differences in inclusion or exclusion based on the specific objectives and criteria of the mapping. They further highlight the importance of combining systematic search strategies with contextual expertise to uncover relevant data sources.

Table 4: Number of datasets identified, included, and excluded by country by partner.

Country	Partner	Datasets identified	Datasets included	Datasets excluded
AT (Austria)	PHKA	4	2	2
BE (Flanders)	INED	4	2	2
BE (Wallonia)	INED	1	0	1
BG (Bulgaria)	PHKA	4	0	4
CH (Switzerland)	UNIMI	11	4	7
CY (Cyprus)	KEPE	1	1	0
CZ (Czechia)	UNIMI	1	0	1
DE (Germany)	PHKA	52	31	21
DK (Denmark)	KEPE	2	1	1
EE (Estonia)	UTAD	5	2	3
EL/GR (Greece)	KEPE	4	4	0
ES (Spain)	UPF	13	3	10
FI (Finland)	UNIMC	14	7	7
FR (France)	INED	9	5	4
HR (Croatia)	FpS	0	0	0
HU (Hungary)	KEPE	4	3	1
IC (Iceland)	UNIMC	2	0	2
IE (Ireland)	UPF	5	3	2

¹⁸ The final number of datasets included in the mapping is the output of the analytical work described in Section 4 and, more specifically, 4.1.

IT (Italy)	FpS	14	3	11
LI (Liechtenstein)	PHKA	1	0	1
LT(Lithuania)	UTAD	6	1	5
LU (Luxemburg)	INED	5	1	4
LV (Latvia)	UTAD	6	1	5
MT (Malta)	UPF	0	0	0
NL (Netherlands)	KEPE	7	4	3
NO (Norway)	FpS	8	3	5
PL (Poland)	FpS	2	1	1
PT (Portugal)	UTAD	15	7	8
RO (Romania)	UNIMC	2	1	1
SE (Sweden)	UNIMI	3	2	1
SI (Slovenia)	UNIMC	0	0	0
SK (Slovakia)	UNIMI	0	0	0
UK (England, Wales, Scotland, Northern Ireland)	UPF, PHKA	19	11	8
U.S.	UPF	5	1	4
TOTAL		229	104	125

Note: The country-partner assignment follows the internal distribution of responsibilities in the consortium. In some cases (e.g. UK), dataset identification and description were conducted jointly by two partners.

Distribution and Geographical Coverage

Overall, this comparative overview provides insight into the **geographical distribution** of existing longitudinal data on student learning outcomes in Europe, as well as the **variability in data availability, accessibility, and compliance** with the inclusion criteria. It also reflects differences in national data infrastructures and research traditions, which are explored further in Section 4 below.

The geographical coverage of the identified datasets is striking. Unsurprisingly, Germany reports the highest number of both identified and included datasets, respectively 52 and 31. The United Kingdom also stands out with 19 datasets identified in total, with 11 datasets included in the final mapping. The distribution of both identified and mapped countries clearly highlights a **concentration of eligible datasets in a limited number of countries**, particularly in Western and Northern Europe.

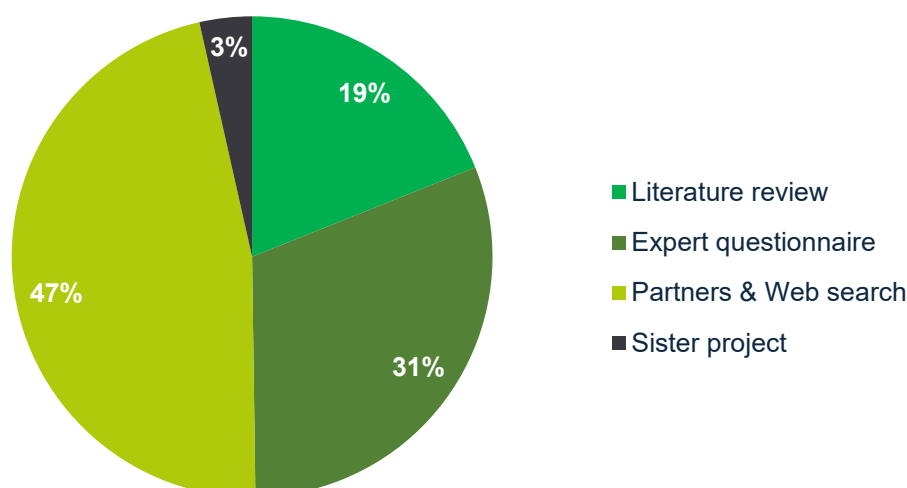
In contrast, **several countries remain entirely uncovered**: in Croatia, Malta, Slovakia, and Slovenia neither the expert questionnaires nor the systematic web-search led to the identification of any datasets. The number of countries without any longitudinal datasets included in the final mapping is even larger, encompassing Bulgaria, Czechia, Iceland, Liechtenstein, and Belgium Wallonia, where datasets were identified but subsequently excluded. More information about the geographical distribution and the countries included in the mapping is provided in section 4.2 below.

As Table 4 above shows, in addition to the 32 European countries covered by the LINEup mapping, a small number of datasets from the United States were also reviewed, to provide an international reference point. Of the five U.S.-based datasets considered, only one – the **Early Childhood Longitudinal Study (ECLS)**¹⁹ – met all the inclusion criteria and was included in the final set. ECLS stands out for its robust longitudinal design, early focus on primary education, detailed coverage of student learning outcomes, and strong documentation and accessibility. Although the other U.S. datasets examined (NAEP, HSLS:09, ELS:2002 SEDA) are widely used in research, they were excluded because they either lacked a longitudinal structure, did not include individual-level learning data, or focused on post-secondary trajectories. Despite being an outlier in the sample, ECLS represents a valuable reference for understanding what a well-documented and policy-relevant longitudinal dataset should look like. Drawing lessons from ECLS – particularly regarding early tracing of student outcomes, integration of background variables, and data accessibility – can inform efforts to strengthen longitudinal data collection and usage across European contexts (see Box in section 4.2 for further details).

Sources of Dataset Identification

Concerning the sources of dataset identification, we distinguished between the literature review, expert consultation, partner knowledge and web-search, and cooperation with sister projects. Each dataset was assigned to a single source based on its first point of identification in the mapping process, even if it was later confirmed through multiple channels. Table 5 presents the distribution of identified datasets by country and by source, while Figure 2 offers a visual synthesis of the overall shares. As shown in the figure, 19% of datasets originated from the literature review, 31% from expert questionnaire responses, 47% from partners' knowledge and systematic web-search, and 3% from collaboration with Sister Projects.

Figure 2: Number of datasets identified by sources (N=229).



¹⁹ Early Childhood Longitudinal Study (ECLS): <https://nces.ed.gov/ecls/>

These figures reflect both the fragmented nature of existing knowledge on longitudinal datasets in education and the crucial role played by the consortium partners in uncovering and systematising this information. The high proportion of datasets identified through partner knowledge and web-search highlights the **importance of research networks, institutional memory, and proactive exploration** - often beyond what is documented in academic literature or known by national experts. At the same time, the relatively limited contribution from literature and formal expert consultation suggests that **relevant data sources are frequently undocumented, underused, or invisible** to the broader research community.

While the identification phase revealed valuable data sources across Europe, the application of shared inclusion criteria resulted in the exclusion of a substantial number of datasets - highlighting persistent limitations in coverage, quality, and relevance, as explored in the next section.

Table 5: Number of datasets identified by sources.

Country	Total number of datasets identified	Literature review	Experts	Partners + web-search	Sister project
AT (Austria)	4	0	0	4	0
BE (Flanders)	4	0	1	3	0
BE (Wallonia)	1	0	1	0	0
BG (Bulgaria)	4	0	4	0	0
CH (Switzerland)	11	1	5	2	3
CY (Cyprus)	1	1	0	0	0
CZ (Czechia)	1	1	0	0	0
DE (Germany)	52	11	7	34	0
DK (Denmark)	2	0	2	0	0
EE (Estonia)	5	0	3	2	0
EL/GR (Greece)	4	2	1	1	0
ES (Spain)	13	4	7	2	0
FI (Finland)	14	7	2	3	2
FR (France)	9	3	1	5	0
HR (Croatia)	0	0	0	0	0
HU (Hungary)	4	0	2	2	0
IC (Iceland)	2	1	0	1	0
IE (Ireland)	5	0	3	2	0
IT (Italy)	14	2	8	4	0
LI (Liechtenstein)	1	0	0	1	0
LT(Lithuania)	6	0	1	5	0

LU (Luxemburg)	5	0	4	1	0
LV (Latvia)	6	0	3	3	0
MT (Malta)	0	0	0	0	0
NL (Netherlands)	7	2	4	1	0
NO (Norway)	8	4	1	3	0
PL (Poland)	2	0	1	1	0
PT (Portugal)	15	2	4	9	0
RO (Romania)	2	0	0	1	1
SE (Sweden)	3	0	0	2	1
SI (Slovenia)	0	0	0	0	0
SK (Slovakia)	0	0	0	0	0
UK (England, Wales, Scotland, Northern Ireland)	19	2	4	12	1
U.S.	5	0	1	4	0
TOTAL	229	43	70	108	8

Reasons for Dataset Exclusion

Among the 125 datasets excluded from the mapping, partners were asked to report the precise reason for exclusion. Table 6 summarizes, for each country, the number of datasets excluded and the detailed reasons for exclusion. Figure 3 presents the overall distribution of exclusion reasons as a percentage of the total.

The most frequent reason for exclusion (30%) was the **absence of information on student achievement at the individual level**. These datasets did not include measures of learning outcomes - such as standardised test scores or academic performance indicators (e.g. school grades, grade retention, truancy, or school trajectories) - nor any other dimensions of cognitive and non-cognitive development (e.g. socio-emotional skills, school engagement, or sustained attention). Many of these datasets were originally included because they were collected by means of representative longitudinal surveys but turned out to focus primarily on other domains such as health behaviours, mental well-being, or family conditions.

The second most frequent reason (28%) was the **outdated nature of the dataset**, with the latest data collection round having taken place before 2014. This highlights the extent to which current empirical research on educational achievement and inequalities still relies on data that may no longer reflect today's social and educational contexts. The finding underscores the urgent need for updated longitudinal data to ensure the continued relevance of research evidence.

A further 14% of the datasets were excluded due to insufficient representativeness, which limited their potential to support generalisable findings. Without representative data, analyses risk overlooking structural disparities and misrepresenting population-level trends.

Around 7% of the datasets were excluded because they **did not cover primary or secondary education** - for instance, panel studies focused on higher education or school-to-work transitions.

Finally, 21% fell under the category “other reasons”, which included cases of experimental designs, missing documentation, or uncertainty about the dataset’s scope.

Figure 3: Reasons for dataset exclusion from the mapping (N=125).

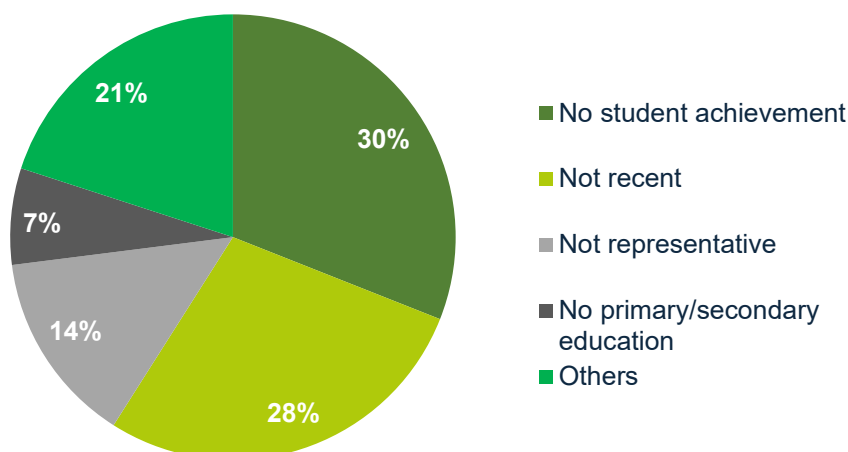


Table 6: Number of datasets excluded by reasons of exclusion.

Country	Dataset excluded	REASON FOR EXCLUSION				
		No student achievement at individual level	Not recent	Not representative	No primary/secondary education	Others ²⁰
AT (Austria)	2	0	0	0	0	2
BE (Flanders)	2	1	1	0	0	0
BE (Wallonia)	1	0	0	0	0	1
BG (Bulgaria)	4	1	0	0	0	3
CH (Switzerland)	7	2	2	0	1	2
CY (Cyprus)	0	0	0	0	0	0
CZ (Czechia)	1	0	0	0	0	1
DE (Germany)	21	2	10	7	0	2
DK (Denmark)	1	0	0	0	0	1
EE (Estonia)	3	3	0	0	0	0
EL/GR (Greece)	0	0	0	0	0	0

²⁰ This category includes reasons such as not a dataset, experimental design, no available information on the dataset, ethnic concerns, data collection incomplete or not started yet, comparative assessments.

ES (Spain)	10	2	4	2	0	2
FI (Finland)	7	0	2	1	2	2
FR (France)	4	0	2	1	1	0
HR (Croatia)	0	0	0	0	0	0
HU (Hungary)	1	0	1	0	0	0
IC (Iceland)	2	0	0	0	0	2
IE (Ireland)	2	1	1	0	0	0
IT (Italy)	11	6	0	2	1	2
LI (Liechtenstein)	1	1	0	0	0	0
LT(Lithuania)	5	3	0	0	1	1
LU (Luxemburg)	4	3	0	0	0	1
LV (Latvia)	5	5	0	0	0	0
MT (Malta)	0	0	0	0	0	0
NL (Netherlands)	3	0	3	0	0	0
NO (Norway)	5	4	1	0	0	0
PL (Poland)	1	1	0	0	0	0
PT (Portugal)	8	2	0	3	3	0
RO (Romania)	1	0	0	0	0	1
SE (Sweden)	1	0	0	0	0	1
SI (Slovenia)	0	0	0	0	0	0
SK (Slovakia)	0	0	0	0	0	0
UK (England, Wales, Scotland, Northern Ireland)	8	1	6	0	0	1
U.S.	4	0	2	1	0	1
TOTAL	125	38	35	17	9	26

The mapping confirms the existence of rich, yet unevenly distributed, data sources across Europe. It also highlights structural gaps - including in dataset documentation - that hinder the capacity to longitudinally monitor educational inequalities on a European scale. The results presented here will inform the forthcoming data analysis and harmonisation efforts and provide a foundation for strengthening national and European data infrastructures.

4. Dataset Analysis

This section presents the results of the **analytical work conducted on the 104 datasets** included in the final mapping. Building on the standardised descriptions produced during the previous phase, the analysis pursued three main objectives:

- To assess the internal consistency and completeness of the dataset entries,
- To classify datasets according to key structural and methodological features,
- To evaluate their empirical potential for comparative research on educational inequalities.

The analytical process proceeded in three main directions. First, a **validation procedure** was implemented to identify and address inconsistencies, verify compliance with the inclusion criteria, and refine the dataset inventory. Second, a **classification scheme** was developed to refine and better distinguish between different types of longitudinal data - survey-based, administrative, and hybrid. Third, **descriptive statistics** were generated to explore core dimensions such as the duration and frequency of data collection, the nature of the variables included, the use of sampling weights, and the overall quality of available documentation.

This structured analysis enhances the scientific robustness of the mapping and provides a solid basis for understanding both strengths and limitations of existing longitudinal data infrastructures across Europe.

4.1 Dataset Validation and Classification

This subsection presents the conceptual and methodological approach adopted for validating and classifying the datasets included in the final mapping. It outlines the validation framework and principles, describes the process followed, and reports the main results of the quality assurance activities.

Validation and Classification Framework

Given the diversity of collected datasets, the validation and classification activities encompassed multiple steps. The objective was to ensure that all included datasets were accurate, complete, relevant, and aligned with the conceptual and methodological standards defined in the earlier stages of the mapping.

The validation procedure was designed as a rigorous and replicable process, grounded in the conceptual framework of the project and consistent with the dataset inclusion criteria, as defined in Section 2 above. The goal was to review the available information to: (i) assess data quality (e.g. identify missing or erroneous information); (ii) evaluate internal and external validity; and (iii) ensure coherence with the project's analytical goals.

The validation framework was structured around four core objectives:

- Ensure that datasets are **accurate, complete, and relevant**.
- Verify **alignment** with the project's conceptual definitions and analytical focus.
- Assess **internal and external validity** of the data and measured constructs.
- Guarantee **transparency** and **replicability** in the validation process.

To operationalise these objectives, the following core principles guided the validation strategy:

- **Rigour** - the process follows a strict and transparent set of rules.
- **Replicability** - all validation steps are documented in detail to allow replication.
- **Coherence** - validation is consistent with the conceptual and methodological choices.
- **Fit-for-purpose** - datasets must be demonstrably suitable for the project's objectives and theoretical constructs.

The validation strategy was built around the dataset inclusion criteria established for the mapping phase (see Section 2 above). To briefly recall, eligible datasets had to meet the following conditions:

- Be longitudinal in nature.
- Refer to primary and secondary levels of education (ISCED 1, 2, or 3).
- Focus on student learning outcomes (such as standardised test scores or grades).
- Originate from the 27 EU Member States, European Economic Area Associated Countries, Switzerland, and the UK.
- Be nationally or regionally representative.
- Be ongoing or recently collected (latest wave not older than 2014).

Based on these criteria, the validation process distinguished between datasets that were fully eligible and those requiring further review. A set of exclusion rules was established to automatically exclude datasets that failed to meet one or more of the core criteria. Table 7 lists the exclusion codes applied during the validation process, along with a brief description of each exclusion criterion.

Table 7: Exclusion criteria.

Code	Exclusion Criterion	Description
E1	Non-longitudinal data	The dataset does not track variables across multiple time points.
E2	Irrelevant educational level	The dataset does not pertain to primary or secondary education (e.g., only covers tertiary or early childhood).
E3	No focus on learning outcomes	The main measured outcomes are not student learning-related (e.g., focus only on attendance, well-being, or infrastructure).
E4	Geographical mismatch	The data originates from countries outside EU-27, EEA, Switzerland, or the UK.
E5	Outdated or discontinued	The dataset does not include recent data collection and is no longer aligned with current educational contexts or policy needs.

For datasets that did not fully meet all inclusion criteria, an additional round of checks was conducted. This step involved a manual review, during which each dataset was examined in collaboration with project partners, to ensure that potentially valuable sources were not excluded due to minor issues or errors. Given the limited availability of educational longitudinal data in Europe, particular care was taken to avoid excluding relevant datasets that might still provide meaningful insights. To this end, any dataset that met one or more of the conditions listed in Table 8 was flagged for manual review, as these issues might require clarification or further investigation, rather than immediate exclusion.

Table 8: Criteria for manual review.

Code	Manual Review Trigger	Description
R1	Incomplete metadata	Key information about sampling, methodology, or variables is missing or unclear.
R2	Ambiguous educational level	Dataset partially includes primary/secondary levels, but scope is unclear or mixed with other levels.
R3	Representativeness unclear	Claims of national/regional representativeness are not supported by clear documentation or sampling method.

Data Validation Process

The validation and classification approach can be schematically represented as a series of steps (see Figure 4). The flowchart outlines a structured, step-by-step process for assessing the suitability of datasets for inclusion in a master dataset for future mapping and visualisation. Each step is designed to ensure that the data is aligned with the project's methodological standards and quality criteria.

The process begins with the collection of potential datasets from various sources – the mapping stage. A first round of **basic checks** is then performed to ensure that datasets are readable, properly formatted, and accompanied by sufficient metadata. This serves as a preliminary filter before more detailed analysis.

Each dataset is then assessed against the predefined **inclusion criteria** (e.g., longitudinal structure, focus on learning outcomes, geographic relevance, etc.). Datasets that fail to meet these standards are flagged for further evaluation. Those that pass the initial checks are subject to an assessment of **internal and external validity** (Step 3), to verify whether they measure what they are intended to and whether findings can be appropriately generalised.

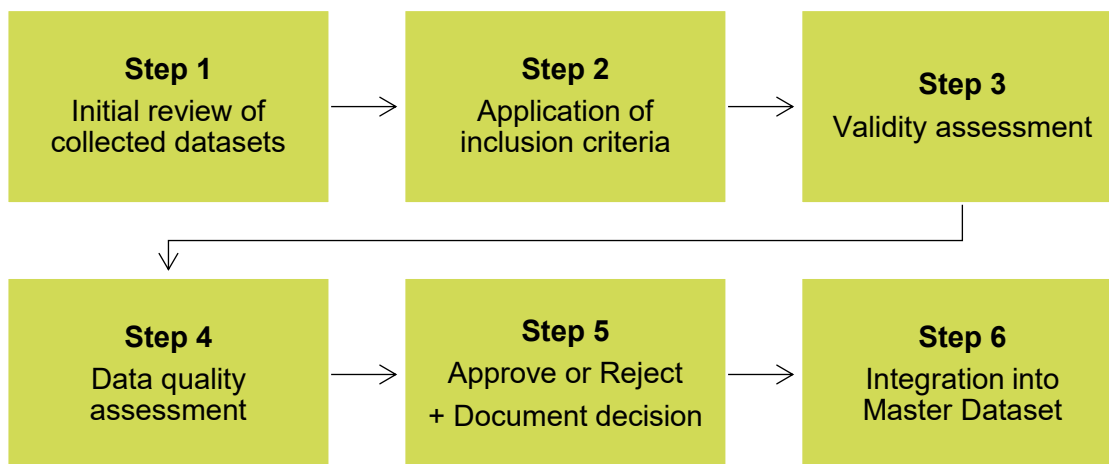
A **technical review** follows (Step 4), focusing on data completeness, internal consistency, and the quality of documentation. This step identifies potential issues such as missing values, coding errors, or lack of clarity in variable definitions.

Based on the results of the previous steps, each dataset is either approved for integration or excluded. The rationale behind each decision is systematically documented to ensure **transparency** and **replicability**.

Approved datasets are then **standardised, harmonised**, and integrated into the project's master dataset, making them ready for visualisation and further comparisons or longitudinal analysis.

The outcome of this process is a consolidated master file containing all validated, quality-assured datasets, along with a **decision log** documenting inclusion and exclusion justifications to ensure full traceability.

Figure 4: Flowchart of the validation activity.



Cross-validation was independently conducted by two researchers using a shared validation protocol. This parallel review process helped identify potential errors and inconsistencies. Validation results were systematically recorded in internal documentation and regularly discussed and updated during team meetings and monthly consortium meetings.

Data Validation Results

During the data validation phase, particular attention was given to verifying the consistency of the reported duration of observation in survey-based datasets. All surveys indicating a duration of less than two years were manually reviewed to detect potential data entry errors or deviations from inclusion criteria.

To facilitate this process, all datasets were sorted in ascending order based on the variable “Years of observation. Three datasets were found to have missing values for this variable, while two others reported only 1.5 years of observation. Each of these cases was manually checked for accuracy and documentation completeness. As a result of internal verification procedures, three datasets of the five datasets were excluded:

- The *Success for All* (UK) dataset was removed due to its limited post-treatment follow-up and, despite including standardised literacy assessments, it did not meet the inclusion criteria for longitudinal tracking of student achievement over time.
- The *ORTHOLEARN* (FR) dataset was excluded due to insufficient documentation and restricted access to data sources.
- Likewise, the *Vérification de l'atteinte des compétences fondamentales* (ÜKG / COFO / VeCoF) was excluded because it does not qualify as a longitudinal dataset.

An additional round of dataset verification was carried out in July 2025 through the LINEup Education Data Explorer (see Section 5 for further details). All project partners were asked to (1) flag any inconsistencies in the dataset records, and (2) verify the accuracy of the descriptions for

the datasets they had originally contributed. This collaborative review led to several refinements, all of which were documented in accordance with the project's standardised documentation protocol.

As a result of this process, two existing entries - *Evaluaciones Diagnósticas* (Gobierno Vasco, ES) and *Schulstatistik* (AT) - were reclassified from longitudinal to cross-sectional datasets. In addition, several studies were re-coded into multiple entries to reflect meaningful internal distinctions, particularly those related to cohort design.

In line with the LINEup methodological approach, each dataset is treated as a single unit of analysis. Accordingly, datasets involving multiple cohorts with distinct sampling, timing, or outcome structures were split to ensure analytical consistency. For instance, *Growing Up in Scotland* (UK) follows three separate pupil cohorts over a 16-year period; due to differences in sample size, measurement waves, and covered outcomes, it was divided into three distinct entries.

In total, seven studies were split into multiple entries:

1. *Growing Up in Scotland* (UK) – split into three entries by cohort.
2. *Growing Up in Ireland* (IE) – split into two entries by cohort.
3. *Transformation des Sekundarschulsystems und akademische Karrieren* (DE) – split into two entries by cohort.
4. *KRTK's linked administrative dataset* (HU) - split into two entries by cohort.
5. *Primary to Secondary Education* (NL) – split into two entries by cohort.
6. *TwinLife* (DE) – split into two entries by cohort.
7. *ReGES* (DE) – split into two entries by cohort.

Data Quality Assessment Results

The primary objective of this phase was to thoroughly review the collected information to identify potential issues, including missing, inconsistent, or erroneous data. The quality assessment process was structured around the following key steps:

1. **Cross-validation of core variables:** This included the verification of key metadata such as survey name, duration, start and end years, sample characteristics, availability of test data, and availability of student-level information.
2. **Verification of repeated measurement:** It was verified that student performance data - such as test scores or grades - were collected at least twice across different time points. This was essential to ensure that datasets captured longitudinal change, as some surveys collect various types of information over time but assess learning outcomes only once.
3. **Identification of sample type:** Sampling strategies (e.g. random, non-random, cohort-based) were reviewed and categorised to support future harmonisation efforts and inform appropriate analytical techniques, particularly the application of sampling weights.
4. **Assessment of weighting information:** For datasets based on random sampling, the availability and nature of weighting variables were documented. In the case of cohort studies, details on the timing and age of data collection were also gathered. For example, the GOLD database in Sweden follows students from grade IX (typically age 16) onward.
5. **Verification of alignment with dataset definitions:** Each dataset was assessed for compliance with the core definitions and inclusion criteria established in Section 2.

Classification Scheme

This section documents all recoding operations applied to the mapped datasets. To enhance comparability and capture the methodological diversity of the mapped data sources, a revised classification scheme was developed including the following key variables:

Dataset Structure: a categorical variable distinguishing between three types of longitudinal data, based on their structure and capacity to track change over time:

- *Pure longitudinal data:* datasets that follow the *same individuals* over multiple time points (panel data structure).
- *Repeated cross-sectional data:* datasets based on separate samples drawn from the same population at different points in time, with longitudinal tracking made possible through *unique identifiers*.
- *Pseudo-panel data:* datasets constructed from repeated cross-sections that allow grouping individuals into synthetic cohorts (e.g., by birth year, region, or other).

Administrative Data: a binary variable indicating whether the dataset is based on administrative records or not. While not analytically rich on its own, it plays a key role in supporting cross-tabulations (e.g. comparing coverage, variables or temporal span between administrative and survey-based datasets) and serves as a foundational input for the derived *Type of longitudinal Data* variable.

Type of Longitudinal Data: a categorical variable combining *Dataset Structure* and *Administrative Data* to capture both institutional origin and data collection method. This variable identifies whether data were collected through administrative records, survey instruments, or a combination of both:

- *Administrative:* based solely on administrative records/data.
- *Survey:* based solely on survey data.
- *Hybrid:* combining administrative and survey-based sources.

Additional variables were introduced based on the dataset online forms completed by project partners, or through supplementary desk research. These include:

- Whether the dataset follows students beyond their school years.
- The main thematic focus of the dataset and corresponding study (e.g. school education, school-to-work transition, household and family choices, child development).
- The presence of specific sampling criteria, particularly in non-random or non-national studies.
- Whether the study adopts an experimental design.
- Whether the dataset covers both primary and secondary education levels.

The additional variables were integrated to improve granularity and usability. The classification is not merely mechanical but reflects a scientific effort to consolidate the mapped datasets into a coherent and high-quality infrastructure for comparative research on learning outcomes and inequalities.

Rigorous validation was essential, especially for multi-source datasets harmonised *ex post*. It ensured data quality by identifying and correcting missing, inconsistent, or erroneous values. It also aligned the datasets with the project's conceptual framework, improving their reliability for further analysis. A transparent validation process enhances replicability, enabling future researchers to trace decisions and verify outcomes. Ultimately, it strengthens cross-country and longitudinal comparisons by addressing methodological heterogeneity and supporting harmonisation in definitions, indicators, and measurement frameworks.

Several challenges emerged during the validation and classification process. Metadata availability and quality varied across countries and dataset types, often requiring further desk research or partner input. Documentation standards differed between administrative and survey datasets, complicating the application of uniform criteria. In older studies, or those lacking clear methodological notes, assessing representativeness proved difficult. These limitations underscore the importance for well-documented, transparent data infrastructures and support the decision to include manual review for borderline cases.

These efforts ensure that the dataset inventory developed in WP3 is not only methodologically sound, but also ready to inform the next phases of the LINEup project - particularly comparative analysis and data harmonisation.

4.2 Descriptive Analysis

The descriptive analysis presented in this section offers a systematic overview of the main features of the 103 longitudinal datasets mapped across 32 European countries²¹. While previous sections focused on how datasets were identified, validated, and classified, the focus here shifts to what each dataset contains and how it is structured - offering crucial insights into their analytical potential for comparative research.

The analysis is organized around the following key dimensions:

- **How:** the methodology used to collect, organize, and structure longitudinal data on education-related phenomena.
- **What:** the education-related phenomena under observation in the longitudinal datasets.
- **Where:** countries and territories for which longitudinal data are available.
- **When:** the period and the time span of the information provided.
- **Access:** the degree of openness and ease of access.

This structure goes beyond a mere inventory of datasets, enabling an assessment of their suitability for examining disparities in learning trajectories and student learning outcomes. Understanding the design and content of each dataset is essential to evaluate its capacity to capture the emergence, persistence, and variation of educational inequalities, both within and between countries.

The descriptive findings reported below confirm that, despite important differences in scope, design, and accessibility, a relatively rich set of longitudinal datasets exists across Europe. Many provide detailed information on student achievement and background characteristics, and several offer the possibility to follow students across multiple educational transitions. However, the analysis also reveals structural gaps, including the limited availability of data on teachers and schools, the underrepresentation of non-cognitive outcomes, and frequent administrative barriers to data access or linkage.

Throughout this section, tables and figures present comparative statistics on the main dimensions of interest. These are accompanied by interpretative remarks that highlight both the strengths and the limitations of the current data landscape. The goal is to provide a clearer understanding of the existing potential and the current constraints for conducting longitudinal research on educational inequality in Europe.

²¹ The U.S. dataset is described separately in a dedicated box at the end of the section.

How: Methodology

The datasets included in the mapping display important methodological differences that influence both their analytical potential and their comparability. Four key aspects are particularly relevant for understanding how data are collected and structured over time:

- The type of longitudinal design, which determines whether individual learning trajectories can be fully reconstructed;
- The type and size of the student sample, which affect statistical power and external validity;
- The use of administrative and/or survey-based information, which has implications for reliability, standardisation, and scope;
- The linkability of individual-level information to other data sources, which may enable multi-level and multi-dimensional analyses.

Concerning the first dimension - the **type of longitudinal design** - most of the datasets (68.0%) were explicitly designed as pure longitudinal. In 22.3% of cases, the datasets were not originally longitudinal in design but include individual student identifiers that allow for linking data across years, thereby enabling the construction of a longitudinal dataset. Finally, around one-tenth of all datasets support the creation of pseudo-panels, based on repeated cross-sectional waves (Table 9).

These differences have important consequences. In general, the fragmentation in design types highlights the need for careful definition of analytical strategies when conducting cross-country comparisons. Pure longitudinal datasets with a panel design offer the highest potential for studying within-student changes in achievement, behaviour, or engagement, as well as for supporting reliable causal inference. In contrast, repeated cross-sectional designs typically allow for the analysis of population-level trends over time but do not support tracking individual developmental trajectories, since personal identifiers are not retained. Nevertheless, they can still yield valuable insights - especially when sample sizes are large and measurement tools remain consistent across waves.

A particularly valuable configuration occurs when repeated cross-sectional data collection is combined with the possibility of linking individuals across waves using stable student identifiers. This enables researchers to follow within-student changes while benefiting from larger sample sizes, reduced attrition, and broader time coverage. For example, in large-scale educational assessments where students are tested at regular intervals (e.g., grades 3, 6, and 9), those with consistent identifiers can be matched across waves - even when each wave involves a newly drawn sample. This allows for the reconstruction of learning trajectories, the analysis of policy impacts, and the modelling of student growth, while avoiding the high attrition rates typically associated with panel designs. Despite these benefits, this approach remains relatively rare in Europe - only 22.3% of the mapped datasets adopt such a model - likely due to a combination of practical and technical challenges. These include the need for complex and costly data infrastructures, difficulties in ensuring measurement comparability across waves, and demanding data management procedures.

Table 9: Type of longitudinal design.

	% of datasets
Dataset with a pure longitudinal design	68.0
Repeated cross-sectional dataset with the possibility to follow students over time thanks to an individual student's identifier	22.3
Pseudo-panel dataset thanks to repeated cross-sectional waves	9.7
Total	100.0

Concerning the **students' sample**, about two-thirds of the datasets are based on student samples, whereas the remaining third cover the entire student population within a specific academic year, grade, region, or country, depending on data collection purposes (Table 10).

Among the sampled datasets, two further distinctions can be made: whether the sample design is random and the size of the sample. Around 45% of the mapped datasets are based on random samples, whereas the remaining 23.6% use a non-random sampling methods, often oversampling specific social groups (e.g. students with migrant background). Random sample datasets vary in terms of sample size, but tend to be relatively small. Half of them is based on samples of fewer than 5,000 students, while the others include larger samples. Notably, 28.6% of the random sample datasets cover more than 10,000 students.

Taken together, the mapped datasets reveal a broad range of sampling strategies. While this diversity may enrich the overall data landscape, it also calls for caution when comparing findings across datasets or countries, as sampling differences can lead to systematic biases or gaps in coverage. This is particularly critical for underrepresented populations - for example, relatively small minorities or students enrolled in specific (non-mainstream) educational tracks - where sample sizes may be limited or selection mechanisms may vary. Any attempt at comparative analysis must therefore be conducted with close attention to sampling design, representativeness, and the potential effects of selection bias on observed outcomes.

Table 10: Type of sample and sample size.

	% of datasets
Population (all students)	31.7
Random sample	44.7
of whom (as a share of total datasets based on random samples):	
• datasets with a sample below 5,000 students	50.0
• datasets with a sample between 5,000 and 10,000 students	21.4
• datasets with a sample above 10,000 students	28.6
Non random sample	23.6
Total	100.0

Looking at the use of **administrative and/or survey-based information**, nearly half of the datasets included in the mapping (46.6%) are based exclusively on survey data, while 19.4% rely solely on administrative sources. The remaining 34.0% adopt a hybrid approach, linking survey and administrative data within the same dataset (Table 11).

This is relevant from both a methodological and a policy perspective. Administrative data - typically collected by public authorities as part of routine operations - offer important advantages in terms of continuity, coverage, and cost-efficiency. When well-documented and accessible, they allow for robust population-level analyses and can often be linked to other institutional records (e.g. enrolment, grades, attendance). Survey-based data, by contrast, offer greater flexibility. They can be tailored to capture variables that administrative systems often lack - such as student attitudes, motivation, or home environment. However, they tend to be collected less frequently, and are typically based on smaller samples, making them more vulnerable to non-response and attrition.

The relatively limited use of administrative data - either alone or in combination with surveys – points to a missed opportunity. Many countries maintain high-quality education registries, yet these are not consistently leveraged for research on learning trajectories, mobility, or outcome inequality. Several factors may contribute to this underutilisation:

- Institutional silos that hinder data integration across agencies;
- Legal, ethical, and procedural barriers that complicate access, especially when cooperation between institutions is required;
- Technical limitations in extracting, linking, and analysing administrative records.

Strengthening institutional frameworks for data governance, integration, and access is therefore essential. This means not only investing in infrastructure, but also fostering a policy and regulatory environment - at national level - that supports responsible data use without creating unnecessary burdens for researchers. Achieving this requires coordinated action from policymakers, data owners, researchers, and data protection authorities alike.

Table 11: Information gathered through administrative sources.

	% of datasets
Survey	46.6
Administrative	19.4
Hybrid	34.0
Total	100.0

Finally, regarding **linkability**, only a limited number of datasets offer the opportunity to connect micro-level student information to other data sources through some form of identification code (e.g. teacher or parent questionnaires, or other types of administrative data). In fact, 44.7% of the datasets of our sample do not include any identification code. Among the remaining 55.3%, linkability is technically possible; however, in some cases (26.2%) access to the identifiers is restricted to data owners. Only about one-quarter of the datasets (29.1%) are both linkable and potentially accessible to external users (usually after an ad hoc request) (Table 12).

This limited linkability constrains the analytical potential of many datasets, especially for conducting multi-level analyses or examining interactions between student characteristics and school- or teacher-level variables. It also reduces the possibility of enriching datasets with longitudinal administrative information, such as enrolment records, grade repetition, or transitions to other educational levels.

Enhancing linkability and improving access conditions - while ensuring full compliance with data protection regulations - would significantly increase the value of existing datasets for educational research and evidence-informed policymaking.

Table 12: Linkability of student's information with other data at the individual level.

	% of datasets
No	44.7
Yes, but linkability to other data available only for data owners	26.2
Yes, linkability to other data available to all users	29.1
Total	100.0

What: Information Collected

The content of the datasets can be divided in two macro-categories, both essential for understanding students' educational trajectories and the possible mechanisms behind inequalities therein: information concerning students' performance at the individual level and information on students' characteristics (individual ones and household-related ones) and on their school and teachers.

Academic performance

There are two main types of educational achievement indicators in the mapped datasets (Table 13): tests on cognitive skills (76.6%) and school grades (68.9%). Only a small share of datasets includes neither of these two indicators (7.8%)²². This means that in over 90% of the cases it is possible to measure educational achievement through at least one observable outcome.

Slightly more than half of all datasets (53.4%) provide information on both cognitive skills and school grades. The joint availability of these two indicators enhances the analytical potential of the datasets, as it enables researchers to explore possible discrepancies between objective assessments of competence and school-based evaluations. Such discrepancies often affect specific student groups and may be shaped by grading culture, teacher expectations, or institutional norms. This makes the combined presence of tests and grades particularly valuable for studies on assessment bias and differential treatment within schools.

While performance-related data are widely available, the degree of standardisation, comparability, and alignment with national curricula remains uneven across countries and datasets.

Table 13: Type of education outcomes included in the datasets.

	% of datasets
Tests on cognitive skills	76.7
School grades	68.9
Both test on cognitive skills and school grades	53.4

Literacy and numeracy are by far the most frequently assessed competence domains. Around 69.9% of all datasets include information on numeracy skills, whereas 70.9% on literacy skills (Table 14).

Notably, 64.1% of datasets provide measures for both numeracy and literacy skills, while a small share covers only one (5.8% numeracy only, 6.8% literacy only). However, nearly a quarter (23.3%) of datasets do not include either, limiting their capacity to capture foundational competences that are key to educational success (data not reported in the table).

A limited number of datasets provide information on foreign language skills (often, but not always, English) (35.9%) or on science (30.1%).

At the same time, most datasets do not include only information concerning cognitive skills, but they also investigate several other issues, underlying a high degree of heterogeneity concerning academic performance measurement. Some examples are other subject areas (e.g. history, art, social sciences, etc.) or non-cognitive constructs such as self-confidence and life satisfaction.

²² In such cases they contain, for example, information on student motivation towards learning or school engagement.

This thematic breadth is a strength of the mapped datasets, as it reflects a growing recognition that student outcomes are multi-dimensional. However, the high heterogeneity in what is measured - and how - also represents a major limitation for comparative research. Topics and constructs are often defined differently across datasets, and measurement tools vary in format, scale, and scope. As a result, while some datasets are rich and versatile for country-specific analyses, few allow for direct cross-country comparison of student performance beyond basic literacy and numeracy.

Table 14: Students' performance included in the datasets by specific topic.

	% of datasets
Numeracy	69.9
Literacy	70.9
Foreign language	35.9
Science	30.1
Other areas	68.9

Student performance is typically assessed through multi-item tests from which continuous performance scores are derived (46.7%). Some datasets, however, provide only ordinal indicators of proficiency levels (17.1%), while others offer both types of measures (19.2%). In a limited number of cases (8.8%), other measurements are used and for 8.2% of datasets, information on the scoring method is not available (Table 15).

This variation in measurement approaches has important implications. Continuous scores allow for more granular statistical analysis, while proficiency levels - easier to interpret for non-technical audiences - are often preferred for policy reporting and communication. Datasets that include both offer greater flexibility: they are better suited to track progress over time, identify threshold effects, and support comparisons across groups or systems. However, diversity in scoring and scaling methods - particularly when documentation is lacking - can significantly reduce the potential for harmonisation and comparability across datasets and countries.

Table 15: Distribution of datasets by type of scale used for skill measurement.

	% of datasets
Both proficiency levels and continuous scores	19.2
Only continuous scores	46.7
Only proficiency levels	17.1
Other	8.8
Information not available	8.2
Total	100.0

Contextual Information: Students, Households, Teachers, Schools

Beyond student performance, most longitudinal datasets mapped by LINEup provide additional information on a range of contextual dimensions that are crucial for understanding educational inequalities. These span four key levels: (i) students; (ii) student households; (iii) teachers and (iv) schools. Such multi-level information allows researchers to examine how individual, familial, and

institutional factors act and interact to shape learning trajectories. However, the extent and quality of these contextual variables vary significantly across datasets.

It is important to note that the LINEup mapping relied primarily on documentation available through datasets websites, codebooks, and published reports. In many cases, this documentation did not clearly indicate whether a specific variable was collected or accessible. As a result, each dataset variable was classified into three categories: “yes” (explicitly included), “no” (explicitly absent), and “not clear” (uncertain or undocumented). These classifications are reflected in the following tables.

Information on Student Characteristics

As shown in Table 16, the availability of student-level background variables is uneven, although some patterns can be identified. Variables can be grouped into four tiers based on how frequently they appear across datasets:

- Widely available: included in (almost) all datasets (gender and age).
- Commonly available: found in most datasets (attitudes and motivation toward school/education, region of residence, being born abroad).
- Moderately available: present in a significant minority of datasets (attendance in early childhood education and care, specific country of birth, grade repetition, citizenship, learning impairments, province/county or town of residence).
- Rarely available: included in only a few datasets (physical impairments, ethnic minority background, access to student allowance or scholarships).

Overall, the emerging picture is one of selective depth: most datasets offer basic demographic and school career information, but fewer allow for a detailed profiling of students' educational experiences or vulnerability factors. This lack of detail might be a relevant limitation, particularly for disaggregated analyses aimed at uncovering disparities across social groups. Without detailed subgroup data, researchers are constrained in their ability to identify and analyse the educational barriers faced by historically disadvantaged populations, including students with a migration background, those with special educational needs or disabilities, or those in need of financial aid. The implications are also severe for cross-country comparative research, where differences in data availability and measurement consistency can hinder the development of inclusive, evidence-informed policy responses. Moreover, the absence of detailed profiling data also limits the potential for evaluating the effectiveness of targeted interventions or support programmes.

Table 16: Distribution of datasets by presence of information on students' characteristics.

	% of datasets (*)
Gender	95.2 (1.9)
Age	90.3 (5.8)
Born in a foreign country	68.0 (17.5)
Attitude or motivation toward education/school	57.3 (16.5)
Region of residence	54.4 (22.3)
Attendance to early childhood education and care	49.5 (19.4)
Specific birth country	46.6 (26.2)
Previous grade retention	42.7 (27.2)
Citizenship	45.6 (21.4)
Learning impairments	40.8 (23.3)

Province/county of residence	39.8 (24.3)
Town of residence	36.9 (22.3)
Physical impairments	29.1 (24.3)
Belonging to a recognised ethnic minority	24.3 (25.2)
Receiving a student allowance/scholarship	14.6 (24.3)

Note: In parenthesis the percentage of datasets without clear information on the item.

Given the relevance of students' migration background in the study of educational inequalities, it is important to further examine the availability and quality of such information in the mapped datasets. Table 17 classifies datasets according to whether they contain information on either citizenship or country of birth. Overall, roughly three-quarters of the datasets include at least one of the two variables. In detail, 37.9% provide information on both, while an additional 37.9% include only one of the two. However, 10.7% of the datasets offer no information on students' migration background, and for a further 13.5%, the metadata are unclear or insufficiently documented.

Only a minority of datasets therefore allow for a detailed or triangulated analysis of this key dimension. The recurring lack of clear documentation reflects broader issues of inconsistent metadata and limited transparency - factors that hinder data reuse by researchers and policymakers. More broadly, the limited availability of robust, well-documented migration-related variables constrains the capacity to conduct cross-country and intersectional analyses of ethnic, linguistic, or migration-related inequalities - an increasingly urgent concern across many European education systems.

Table 17: Distribution of datasets by presence of information on students' nationality or country of birth.

	% of datasets
Information on both citizenship and country of birth	37.9
Information on either citizenship or country of birth	37.9
Information neither on citizenship nor on country of birth	10.7
Not clear	13.5
Total	100.0

Information on Household and Family Background

Household-level variables are essential for understanding how family resources, cultural capital, and living conditions shape educational opportunities of children. Unfortunately, unlike basic demographic information on students, no single family-related variable is consistently available across all datasets.

As shown in Table 18, the availability of these variables varies considerably and can be grouped into three broad categories:

- Frequently available: included in most datasets (parental education, working status, occupation, migration background and the number of parents in the household);
- Moderately available: found in a limited number of datasets (parental income/wealth, parents' age, presence of stepparents, presence of siblings and number of books in the household);

- Rarely available: included in only a few datasets (parental level of knowledge of host country's language, house ownership, number of digital devices in the household).

These variables would be particularly useful in analysing material deprivation, digital inequality, and the integration of migrant families - yet they are not available in most datasets. In most cases, only basic socio-economic indicators - such as parental educational or employment status - are included, often collected with limited detail. Although socio-economic background is acknowledged as a critical factor in educational research, the limited granularity, standardisation, and availability of household data remain a major obstacle to fully understanding and addressing the ways in which family circumstances shape student opportunities, experiences, and outcomes.

Table 18: Distribution of datasets by presence of information on student household's characteristics.

	% of datasets (*)
Parental education (when present, usually both parents)	73.8 (12.7)
Parental working status (when present, usually both parents)	65.1 (13.6)
Parental occupation (when present, often detailed for both parents)	61.2 (17.5)
Parental migration background (when present, usually both parents)	60.2 (13.6)
Number of parents in the household	51.5 (19.4)
Parents place of birth (when info available, mostly specific country)	51.5 (19.4)
Siblings (when present, usually precise number)	45.6 (27.2)
N° of books in the household	43.7 (17.5)
Parental income or wealth	43.7 (19.4)
Presence of stepparents	41.8 (23.3)
Parents age	39.8 (17.5)
Ownership of the apartment/house they live in	25.2 (20.4)
Parental level of knowledge of host country's language	21.4 (31.1)
N° of digital devices in the household	16.5 (29.1)
Receiving/having received a student allowance/scholarship	14.6 (24.3)

Note: In parenthesis the percentage of datasets without clear information on the item.

Information on Teachers

Compared to information on students and their households, data on teachers are relatively scarce. Only basic characteristics - such as gender, age, and seniority - are available in around one-third of all datasets, while other substantive indicators (e.g. teachers' education degrees or contract types, both relevant for assessing teacher effectiveness) are even less frequently included (Table 19).

This scarcity reflects a broader structural issue: most longitudinal datasets on student learning outcomes are not designed to capture teacher-level factors. Even when such data are available, they are often not linkable to individual student records. This limits the ability to analyse how teacher characteristics affect student achievement, or how their distribution across schools might reinforce or mitigate inequality.

An additional challenge lies in the complexity of linking students and teachers, particularly in secondary education, where students are taught by multiple subject-specific teachers. Unlike in

primary education, where a single teacher or a small team of teachers often follow the same group of students, in secondary settings it becomes technically more difficult to identify which teacher is responsible for which outcomes - making reliable student–teacher matching more complex.

All this considered, the lack of teacher-related data seems a missed opportunity. Indeed, information on the teaching workforce is essential for understanding classroom-level dynamics, analysing school resource allocation, and evaluating the impact of teacher qualifications or stability on student outcomes - particularly in disadvantaged contexts.

Table 19: Distribution of datasets by presence of information on teachers' characteristics.

	% of datasets (*)
Teacher gender	36.9 (15.5)
Teacher age	35 (15.5)
Teacher seniority	33.0 (19.4)
Teacher educational degree	28.2 (18.5)
Teacher type of contract (i.e. permanent, fixed term)	16.5 (20.4)

Note: In parenthesis the percentage of datasets without clear information on the item.

Information on Schools

While information on teachers is often lacking in longitudinal databases, some basic data on schools are often available (Table 20). The two most common piece of information available is the type of school (public or private) (58.2%) and the school track followed (57.3%). These variables are fundamental for understanding how institutional differentiation affects learning outcomes - for instance, by enabling stratified pathways or selective allocation mechanisms that may reinforce inequalities.

Any other information, from school and class composition and size to school geo-referencing, is seldom available in datasets. This limited availability significantly constrains the potential to conduct multi-level analyses or to assess the role of school context and peer effects in shaping student learning and career trajectories. In particular, the lack of geo-referenced data and class-level information makes it difficult to link educational outcomes to broader contextual variables such as neighbourhood deprivation, school segregation, or catchment area policies.

As with teacher data, the limited coverage of school characteristics reflects a **student-centred design of most longitudinal datasets**. However, given the well-documented influence of school-level factors on educational inequality, enhancing the collection and accessibility of contextual school data remains a key priority for future data infrastructure development.

Table 20: Distribution of datasets by presence of information on school's characteristics.

	% of datasets (*)
Type of school (i.e. public, private)	58.2 (14.6)
Track followed (if relevant)	57.3 (26.2)
School size	39.8 (16.5)
School composition (i.e. migration/socio-economic background)	30.1 (24.3)
Geo-referencing of the school	28.2 (26.2)

Classroom size	26.2 (21.4)
Class composition (i.e. migration/socio-economic background)	23.3 (26.2)

Note: In parenthesis the percentage of datasets without clear information on the item.

Overall, the information reported in this section shows that currently available longitudinal datasets provide information not just on students' performances but also on a broader set of contextual variables related to students, families, schools, and - to a lesser extent - teachers. However, the availability and depth of this information are highly uneven.

Variables such as age, gender, and migration background are widely available and provide basic information for equity-oriented analyses. A fair number of datasets also include indicators of parental education, occupation, and household composition, but income- and resource-related variables remain scarce. In contrast, data on schools and, especially, teachers are more limited, often restricted to big institutional differences (e.g. public/private, school track) without further contextualisation. This imbalance reflects a persistent individual-centred focus in educational data collection, which hinders the study of structural and institutional factors and of their impact on learning and inequality. To support more comprehensive and policy-relevant research on educational inequalities, future investments in longitudinal data infrastructures should prioritise:

- The inclusion of richer contextual variables.
- Better documentation of existing metadata.
- Greater data linkability across individual, school, and system levels.

Where: Countries Included

24 European countries are covered by at least one longitudinal dataset (Table 21). As previously underlined in Section 3.3 above, the countries with the highest share of longitudinal datasets are Germany (30.1%) and the UK (10.6%). Other countries where there is a relatively widespread availability of longitudinal datasets are Finland, Portugal and France. All Western European countries, together with the UK, Norway and Switzerland, have at least one dataset, whereas some countries in Central-Eastern Europe do not have any (e.g. Czechia and Slovakia).

In general, countries in Central and Eastern Europe display a structural lack of accessible longitudinal data. In many cases, relevant information is either extremely difficult to access - due to language barriers, outdated or poorly functioning websites, or lack of documentation - or simply not available. A notable exception is Hungary with three eligible longitudinal datasets. Empirical evidence on educational inequality trends is scarce or virtually absent in many of these countries. This reflects, and is exacerbated by, the absence of recent, reliable, and publicly accessible longitudinal data on educational outcomes, school trajectories, and students' academic performance. The limited availability of such data significantly constrains both the development of comparative research and the capacity of policymakers to design evidence-informed strategies.

In contrast, Germany is widely recognised as one of the leading countries in Europe in terms of producing rich and accessible survey-based data on educational inequalities, also thanks to large investment in research in the field of education inequalities. Many researchers across Europe rely on German datasets for their analyses, largely due to the volume of available data, ease of access, and clarity of documentation.

However, it is important to note that Nordic countries, particularly Finland, Sweden, and Norway, offer a different but equally valuable model: they maintain comprehensive administrative registers that enable the long-term tracking of entire student populations. These datasets can be linked across sources - such as school records, household registers, and labour market data - through personal identifiers, allowing for highly granular and longitudinal analyses without relying on repeated surveys. While the number of mapped survey datasets may be lower in these countries, the quality, coverage, and linkage potential of register data represent a major strength of their national data systems. Nonetheless, access to such register-based data is typically highly restricted, due to privacy concerns deriving from their census nature combined with the highly detailed information there recorded. In many cases, only national institutions or affiliated researchers are allowed to use them, which limits their potential for broader comparative or cross-border analyses.

These discrepancies in geographical coverage and dataset typologies do not merely reflect differences in technical capacity for data collection, but also point to deeper structural disparities in national investment in educational monitoring systems, data governance, and research infrastructure. Our findings underscore the importance of systematically mapping data availability as a prerequisite for advancing empirical research on education across Europe. Without comparable and interoperable data infrastructures, the capacity to monitor and address educational inequalities remains structurally unbalanced.

Table 21: Distribution of datasets by country²³.

	% of datasets
DE (Germany)	30.1
UK (England, Wales, Scotland, Northern Ireland)	10.6
FI (Finland)	6.8
PT (Portugal)	6.8
FR (France)	4.9
CH (Switzerland)	3.9
EL/GR (Greece)	3.9
NL (Netherlands)	3.9
ES (Spain)	2.9
HU (Hungary)	2.9
IE (Ireland)	2.9
IT (Italy)	2.9
NO (Norway)	2.9
AT (Austria)	1.9
BE (Flanders)	1.9
EE (Estonia)	1.9
SE (Sweden)	1.9
CY (Cyprus)	1.0
DK (Denmark)	1.0
LT (Lithuania)	1.0

²³ For the number of datasets identified and mapped in each country refer to Section 3.3.

LU (Luxemburg)	1.0
LV (Latvia)	1.0
RO (Romania)	1.0
PL (Poland)	1.0
Total	100.0

Often the datasets have a national scope (62.1%). However, more than one third of them (37.9%) focuses on specific regions within a country (Table 22).

Table 22: Geographical scope of the datasets.

% of datasets	
National	62.1
Sub-national	37.9
Total	100.0

Table 23 presents the distribution of countries according to the geographical scope of their datasets: those with only national-level datasets, those with only sub-national-level datasets, and those with both. Overall, a substantial number of countries have at least one national-level dataset. Combined with the high proportion of datasets based on full population coverage or random sampling, this points to a generally good level of representativeness across Europe.

Table 23: Distribution of datasets according to countries and their geographical scope.

Countries with only national level datasets	Countries with both national and sub-national datasets	Countries with only sub-national level datasets
CY, DK, EE, FR, HU, IE, IT, LT, LU, LV, NL, PL, PT, SE	AT, CH, DE, EL, FI, NO, UK	BE Flanders, ES, RO

When: Duration and Depth of Longitudinal Data

One of the key features of longitudinal datasets is their temporal structure - particularly the length of the study period and the number of time points at which observations are made. These elements are crucial for assessing the robustness of learning trajectories, capturing key educational transitions, and supporting causal inference or change-oriented analyses. When studies focus on education systems, additional aspects become particularly relevant. For example: Does the dataset span the full educational cycle, from primary to upper secondary education? Does it also include pre-primary or tertiary levels? And which specific grades are covered?

Table 24 illustrates the diversity of the mapped datasets in terms of the number of data collection waves. At one end of the spectrum, 19.7% of datasets span only two years, while at the other end, around 22.7% cover more than 13 years. Between these extremes lies a wide range of research designs, reflecting differences in duration, periodicity, and intended purpose.

This variation stems from both the heterogeneity of national data infrastructures and the diverse goals for which these datasets were created. Shorter panels may suffice for evaluating specific interventions or key educational transitions (e.g. from lower to upper secondary education), whereas longer panels are essential for capturing the cumulative nature of educational careers and disadvantage - allowing researchers to analyse how early inequalities emerge, evolve, and persist over time and across systems.

Table 24: Data-Collection Duration across Mapped Datasets.

No of years/waves	% of datasets
2	19.7
3	10.6
4	4.6
5	7.6
6	7.6
7	3.0
8	6.1
9	3.0
10	9.1
11	3.0
13	3.0
14+	22.7
Total	100.0

Table 25 shows the percentage of datasets that cover each school grade, offering several insights. First, nearly one third of all datasets (30.1%) start when students are in kindergarten, allowing for the analysis of early childhood conditions and transitions into primary - and eventually secondary - education. Such an early starting point is particularly valuable for studying the emergence of inequalities before formal schooling and the long-term effects of preschool experiences on later educational outcomes, academic performance, and engagement.

With the partial exception of grade 13, every school grade is covered by at least some datasets. Many school grades are included in over half of the datasets, indicating relatively broad coverage across the educational spectrum. However, certain grades appear more frequently: grade 9 (66.9%), grade 10 (55.4%), and grade 8 (55.3%). The prominence of grade 9 reflects its central role in many education systems, often marking the end of lower secondary education and/or the beginning of tracking into upper-secondary pathways.

The concentration of observations in particular grades mirror the structure of national school systems - especially key transition points, assessments, or policy-relevant milestones. At the same time, some stages - particularly mid-primary years and final grades like grade 13 - tend to be underrepresented.

These patterns are important to consider when selecting datasets for longitudinal or comparative research. While some datasets allow for continuous tracking across multiple grades, others focus on specific stages and thus provide only partial views of students' educational trajectories.

Table 25: School grades covered in the datasets.

	% of datasets
Kindergarten	30.1
I grade	43.7
II grade	42.7
III grade	42.7
IV grade	46.6
V grade	49.5
VI grade	52.4
VII grade	46.6
VIII grade	55.3
IX grade	66.9
X grade	55.4
XI grade	50.5
XII grade	53.4
XIII grade (when present)	28.2

Around 47.1% of datasets cover both primary education years as well as secondary education years, whereas 21.2% of them limit their focus to primary education and 31.7% to secondary education years (Table 26).

Datasets including both cycles are particularly valuable for analysing how early learning conditions, family background and/or school environments affect later outcomes - such as academic tracking, school engagement, or dropout. They also allow for the study of cumulative disadvantage, capturing how inequalities evolve across educational phases. Conversely, datasets limited to a single cycle - especially those focusing only on secondary education - may offer richer detail within that stage but provide limited insights into how prior experiences shape later outcomes. This trade-off between detail and continuity is a recurring feature in the current data landscape and should be carefully considered when selecting datasets for longitudinal or comparative research.

Table 26: Classification of datasets according to the coverage of primary and secondary education.

	% of datasets
Both primary and secondary education	47.1
Only primary education	21.2
Only secondary education	31.7
Total	100.0

An additional dimension of longitudinal dataset quality concerns the recency and continuity of data collection. As shown in Table 27, most datasets (51.4%) are still ongoing, whereas 27.2% do not reach the current decade and 21.4% were terminated in the first years of the 2020s.

While historical datasets remain valuable - particularly for long-term trend analysis - the absence of recent waves limits their policy relevance and their usefulness for studying the effects of contemporary challenges, such as COVID-19 related disruptions, digitalisation, or shifts in migration and geographical mobility.

This finding underscores the need not only to maintain the existing longitudinal studies, but also to invest in their continuity, ensuring that they evolve alongside educational systems and societal transformations. Without regular updates, even the most well-designed datasets risk becoming obsolete for present-day research and decision-making.

Table 27: Classification of datasets according to the year of ending of the study.

	% of datasets
Still ongoing (2025)	51.4
Ended between 2020-2025	21.4
Ended before 2020	27.2
Total	100.0

Data Accessibility

One of the most critical dimensions emerging from our mapping exercise concerns the accessibility of the datasets. As shown in Table 28, access remains a major barrier to the use and reuse of longitudinal data in education research.

Around 26.2% of datasets are not accessible or are subject to significant restrictions - either regarding the specific data that can be accessed or the conditions of access (e.g., some datasets can only be consulted on-site using computers of the agency managing the data). By contrast, only about 10.7% of datasets are fully open, while access to around half of the datasets can be granted upon ad hoc request. Notably, for a significant share of the datasets (13.6%), access criteria and conditions are unclear or no information has been found.

This situation of limited access undermines one of the core objectives of longitudinal data: to serve as a shared resource for advancing knowledge, supporting public accountability, and informing effective and equitable policy based on evidence.

While strict access policies and rules may be justified by issues of data protection and confidentiality requirements, the lack of transparency, standardisation, and publicly available metadata often makes even legitimate use difficult and time-consuming - especially for international or comparative research efforts. For example, access procedures may be poorly documented, metadata may be incomplete or unavailable, and the criteria for data use approval are often inconsistent or unclear. As a result, researchers may face long delays or even rejections when trying to access data. Improving access conditions - through clearer procedures, trusted research environments, and secure de-identification practices - is essential. Enhancing the availability of comprehensive, machine-readable metadata and maintaining publicly accessible data catalogues would significantly reduce entry barriers and enable researchers to more effectively identify and request relevant datasets. These improvements are not merely technical or administrative; they require a shift in institutional culture toward viewing data as a public good - essential for understanding and addressing persistent educational inequalities, and for supporting evidence-based policymaking across Europe.

These findings collectively highlight the opportunities and structural limitations that shape the empirical potential of already existing longitudinal data. The next section explores these analytical implications in greater depth.

Table 28: Classification of datasets according to microdata accessibility.

	% of datasets
Open data	10.7
Full access to data through requests	49.5
Highly restricted data / No access	26.2
Not clear/No information	13.6
Total	100.0

The evidence reported in this section of the report highlight the opportunities and structural limitations that shape the empirical potential of existing longitudinal data. While the mapping focuses on 32 European countries, a broader comparative perspective can offer valuable lessons. The box below draws on the U.S. experience to explore one of the most comprehensive and usable longitudinal education datasets, providing insights into best practices for early tracking, multi-level integration, data access and policy use.

Best practices for the U.S. - The ECLS Program

To broaden the comparative perspective and identify best practices, a small number of datasets from the United States were also reviewed. Of the five U.S.-based datasets examined - NAEP, HSLS:09, ELS:2002, SEDA, and ECLS - only one met all the inclusion criteria and was incorporated into the final mapping: the **Early Childhood Longitudinal Study (ECLS) Program**.

The ECLS programme collects **longitudinal data on different cohorts of children**, meaning that children from the same cohorts are followed over time with repeated observations, rather than being replaced by new respondents at each wave. This design enables the study of intra-individual change and developmental trajectories, which is key for identifying causal relationships and early predictors of later educational outcomes. The longitudinal follow-up begins in early childhood and continues through various stages of primary and secondary schooling, providing systematic information on children's cognitive, emotional, and social development. The ECLS program includes four longitudinal studies:

- **ECLS-B** is a sample of approximately 10,700 children born in the U.S. in 2001 and followed from birth through kindergarten entry.
- **ECLS-K** (1998-99 cohort) is a sample of approximately 21,000 children enrolled in kindergarten in 1998/99 and followed through the eighth grade.
- **ECLS-K:2011** (2010-11 cohort) is a sample of around 20,000 children enrolled in kindergarten in 2010/11 and followed through the fifth grade.
- **ECLS-K:2024** (2023/2024 cohort) is the newest program study, planning to follow a sample of about 16,500 children from kindergarten through the third grade.

The ECLS studies are based on **nationally representative samples** of children, selected through a **multistage, stratified, and clustered** sampling design. This involved selecting primary

sampling units (PSUs), then schools within the PSUs and finally children within the selected schools. This ensures the generalisability of findings to the broader population of U.S. children in the relevant cohort groups. The sample design also includes oversampling of specific subgroups - such as racial/ethnic minorities, low-income families, children with low birth weight and twins - to enable disaggregated analysis by increasing statistical power and thereby supporting equity-oriented research. The ECLS program collects rich, **multi-source data** on children's cognitive development, socio-emotional skills, health, family background, school environment, and teacher characteristics. These data are gathered through a combination of direct child assessments, parent and teacher questionnaires, school administrator surveys, and classroom observations - ensuring both breadth and depth of information. This comprehensive approach allows for detailed analysis across multiple levels - individual, family, classroom, and school - and over time. For the past 20 years, the data collected through this study has been used by educators, researchers, and policymakers with the aim of improving the quality and equity of educational experience for all children.

Within the Department's Institute of Education Sciences (IES), the ECLS program is conducted by the U.S. Department of Education and the National Centre for Education Statistics (NCES). **Multiple institutional actors** contribute to the design and implementation of the longitudinal surveys, including academic experts, policy advisors, and data collection agencies. The program adheres to rigorous quality standards in line with U.S. federal regulations on education data, including strict protocols for sampling, instrumentation, fieldwork, and data protection. This institutional framework ensures methodological consistency across waves, high reliability of measurement instruments, and robust procedures for data anonymisation and access control - making ECLS one of the most trusted sources of longitudinal education data in the United States²⁴.

What distinguishes ECLS as a best practice is its **comprehensive and policy-relevant scope**. By tracking children from the very beginning of formal learning - starting in early childhood and continuing through the early grades - the study enables researchers to identify developmental patterns, social inequalities, and early predictors of later academic and life outcomes. While its focus is limited to the early years of education, this phase is increasingly recognised by scholars as critical for shaping long-term trajectories. The integration of **individual-level data** with **contextual information** (spanning family, classroom, and school-level characteristics) allows for detailed multi-level analyses of how different environments and factors interact to influence student learning and development over time. As such, ECLS offers not only a detailed snapshot of early educational experiences but also a powerful tool for informing early intervention policies and supporting evidence-based decision-making at multiple levels. The consistency across ECLS cohorts also enhances the comparability of findings over time, offering a stable reference for longitudinal analysis. Its modular design can serve as a replicable model for future initiatives seeking to build robust, child-centred education data systems.

Furthermore, the **accessibility** and documentation of ECLS data are exemplary. Researchers worldwide benefit from extensive user guides, datasets, online tools, and well-documented

²⁴ NCES (2023). *Early Childhood Longitudinal Study (ECLS)*. U.S. Department of Education, Institute of Education Sciences. Available at: <https://nces.ed.gov/ecls>
Tourangeau, K., Nord, C., Lê, M., Sorongon, A. G., & Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Combined User's Manual for the ECLS-K Eighth-Grade Data Files and Electronic Codebook (NCES 2009-004)*.

codebooks, as well as both public-use and restricted-use data. Public-use data is freely downloadable for everyone and includes general demographic and educational outcomes information. Restricted-use data require a license and access protocol and include more detailed information and geographic identifiers that are sensitive in their nature. While public-use data is suitable for general research, restricted-use data allows analyses at a more granular level, requiring detailed information. This robust infrastructure supports **transparency, reproducibility, and broad interdisciplinary use**, from education policy to developmental psychology, sociology, and economics. Importantly, the ECLS team tailors its communication to different audiences and purposes, providing clear, well-structured materials for practitioners, policymakers, researchers, and those involved in developing data systems and infrastructures.

While ECLS is an outlier in our European-focused mapping exercise, it is worth covering it since it offers a valuable benchmark for assessing both the strengths and weaknesses of existing European longitudinal data infrastructures as to illustrate potential future developments. Its inclusion underscores key areas for improvement – most notably in terms of early tracking of learning outcomes, systematic integration of socio-demographic and institutional data, and making datasets and documentation more accessible and clearer to a broad research and policy community. By comparison, many European systems start data collection at later stages, often lack contextual depth, rely on limited or non-representative samples, or present barriers to access and usability. Drawing on the ECLS model - which demonstrates the feasibility and value of integrated, early-start longitudinal studies - can inform the development of more consistent, inclusive, and policy-relevant longitudinal data systems across Europe. Such systems are essential for understanding educational inequalities, evaluating reforms, and designing equitable interventions from early childhood onward.

As European countries strive to strengthen their data ecosystems in support of inclusive education policies, the ECLS offers a compelling reference for designing **integrated, equitable, and policy-responsive longitudinal data frameworks** from the early years onward.

Unfortunately, the ongoing data collection efforts for the latest ECLS-K:2024 cohort have been suspended due to budget cuts implemented under the Trump administration as part of the Doge Program. Although some data have already been collected, it remains unclear whether the data collection will resume in the future. Additionally, it is unclear whether the existing datasets from all cohorts will continue to be accessible to researchers in their “restricted use” format (which includes more detailed information) as was so far.

4.3 Analytical Potential & Barriers

Building on the descriptive analysis presented in Section 4.2 above, this section discusses the analytical potential of the mapped datasets and the barriers that may limit their effective use for comparative research.

The longitudinal datasets identified during the mapping phase offer a **valuable yet uneven foundation** for investigating educational inequalities over time. The mapping activity itself provided not only a detailed inventory of existing data sources, but also an opportunity to reflect on their usability for future research - particularly in light of the upcoming LINEup feasibility study.

This section is structured in two parts. It first highlights the main **analytical opportunities** emerging from the mapped datasets, focusing on features such as standardized test data, longitudinal

coverage, and the availability of contextual information. It then addresses several **cross-cutting challenges** affecting the usability of the data. These challenges - related to data access, documentation, comparability of variables and structural features - also illustrate the close connection between the mapping phase and the following data harmonisation.

Rather than diminishing the value of the mapping, these challenges help define its scope and limitations. They point to a **fundamental trade-off** between providing researchers with accessible and structured metadata and the more detailed specifications required for robust comparative longitudinal research.

The Analytical Potential

The mapping process has revealed several positive aspects regarding the availability and richness of longitudinal datasets on student achievement across Europe. First and foremost, it confirmed the existence of a substantial number of datasets, reflecting a growing investment in the longitudinal monitoring of student learning outcomes at national and regional levels.

Many of the identified datasets display key features that are particularly relevant for the LINEup project, including:

- Large sample sizes and, in many cases, national-level coverage;
- Widespread use of standardized tests to assess student achievement;
- Measurement of multiple learning domains (typically literacy and numeracy) within the same wave;
- Availability - in a significant number of cases - of more than two waves of student assessment;
- Background information on students' gender, family characteristics (such as parental education and occupation), and migration background;
- Possibility - in several cases - to link standardized test scores with teacher-assigned grades;
- In a non-negligible number of cases, information on teachers potentially linkable to their students.

Although the datasets vary in design and scope, many of them offer the necessary features to investigate how inequalities in achievement emerge, persist, or evolve across school levels and social groups. In the most favourable cases, it is even possible to trace these dynamics across multiple student cohorts. The longitudinal nature of many datasets also enables advanced statistical modelling, such as growth curve analyses of skill development or the estimation of school and peer effects on achievement inequalities.

Datasets linking students to schools and teachers opens further opportunities, particularly to explore the contribution of teacher characteristics, peer effects, and school context to the development of educational inequalities. While such analyses may initially be limited to national datasets, they represent a valuable step towards cross-country comparisons.

Overall, our mapping confirms that many countries possess datasets suitable for in-depth analyses of educational inequalities, and that conditions for cross-national comparative research are emerging - particularly in countries where datasets are methodologically robust and potentially interoperable. Germany, Italy, France, and the United Kingdom stand out in this regard, offering high-quality, nationally representative datasets with rich metadata and harmonisation potential.

These results highlight the empirical ground upon which the LINEup project can formulate its next-phase research hypotheses and comparative strategies. The feasibility study will build on this

analytical potential, identifying concrete opportunities for harmonisation and cross-country modelling, while also addressing the limitations discussed in the next section, which are key to unlocking the full comparative and policy-relevant potential of the mapped data.

Cross-cutting Challenges for Comparative Use and Research Design Reflections

As outlined above, the mapping results are promising and reveal a wide range of datasets with strong potential for the longitudinal analysis of educational inequalities. However, the process also revealed several cross-cutting challenges that must be considered - not only in preparation for the feasibility study, but also as part of a critical reflection on the limitations and conditions of the mapping exercise itself.

These challenges span multiple phases of the research cycle and are grouped into three main areas:

- Data access/documentation;
- Timing of measurement (both in terms of year and students' age/grade);
- Presence and measurement of key variables.

To illustrate the practical implications of these challenges, we refer to three hypothetical research questions that are relevant to the goals of the LINEup project and commonly addressed in the literature, such as the emergence of achievement gaps, the relationship between different forms of performance assessment, and the impact of school composition. These questions, clearly central to the educational inequality debate, are presented for illustrative purposes only, their actual feasibility will depend on the availability and comparability of data across countries. They are particularly relevant for advancing equity-oriented and comparative research in under-studied regions, including Southern Europe:

- **RQ1.** When and how do educational inequalities related to gender, socio-economic status, or migration background emerge and evolve over time? Are there differences between Southern and other European countries?
- **RQ2.** How do standardized test results and teacher-assigned grades relate to each other across different countries and student groups? Can we observe systematic patterns that affect students' future outcomes?
- **RQ3.** To what extent are schools and classrooms in Europe socially and academically mixed or segregated? How does this relate to students' academic performance and choices over time, and how do these patterns vary across education systems?

Data access

During the mapping process, three intertwined dimensions of data access emerged, each posing specific limitations for the comparative analysis of longitudinal data.

First, access to clear and structured documentation about the datasets is limited. This includes detailed information on how variables are operationalised and on the structure of the dataset. The term "structure" means not only the mapping of cohorts, school grades, and survey years, but also more granular elements such as:

- The data format (e.g. wide or long),
- The consistency of variables' definitions across different waves within the same panel,

- The availability of identifiers that reflect the hierarchical organisation of the data (e.g. region, school, class), which are particularly relevant for research questions such as RQ3 on school-level segregation and composition effects.

For all these aspects, our mapping endeavour has shown that well-documented and publicly accessible metadata are still rare. As a result, researchers often have to directly access the datasets to understand their structure and features, even at a preliminary stage, when harmonisation is only being considered. This significantly raises the threshold for planning comparative research.

Second, challenges arise concerning the practical ability to obtain datasets in a reasonable timeframe, even for exploratory purposes, when researcher are still in the process of selecting potential comparison countries or combining multiple datasets within a country, for instance to increase statistical power or to cover a longer time span. This issue emerged more acutely in countries where the mapped data already appear less robust or well-documented. In these cases, codebooks and user guides are often incomplete or ambiguous, making hands-on access to the data essential to identify critical elements such as the presence and extent of missing data on key harmonisation variables (e.g. family or migration background) or the degree of attrition across waves in longitudinal studies. Such information, while crucial to all research questions, is rarely available in advance. Furthermore, in several cases, the process of data access was slower and more uncertain than expected, with national authorities or data owners not always responding or cooperating as anticipated. This difficulty in obtaining datasets from several countries is itself a relevant finding of the mapping process.

Third, the most critical and complex barrier appears to be related to data transferability. Even when data is made available to researchers, it is often subject to restrictions that prevent sharing with other teams or across national borders. This reflects a broader absence of a data-sharing culture in education research and may create substantial operational limitations.

Two main consequences follow from this:

- Data exploration, harmonisation, and analysis may need to be carried out within each country, requiring tight coordination at the supranational level and resulting in higher transaction costs;
- The inability to merge datasets across countries, even in harmonised or anonymised form, may significantly constrain the analytical potential, especially for cross-country comparisons.

In practical terms, researchers may be forced to conduct parallel national analyses and compare results ex post, rather than building a unified supranational dataset. This approach reduces statistical power and limits the interpretability of cross-country differences.

This issue affects all three research questions. For RQ1 and RQ3, it hinders the possibility of testing country-level differences within a single statistical model and restricts the use of multilevel techniques. For RQ2, it prevents the accumulation of teacher-level data across countries, which would be particularly useful for examining the (often weak) associations between teacher characteristics and grading bias. In the case of RQ3, the lack of transferability limits the ability to compare patterns of segregation and composition effects using harmonised indicators of school or classroom-level stratification.

In summary, although the mapping revealed a promising number of longitudinal datasets, **data access remains one of the most complex and consequential barriers** to their comparative use. Anticipating these challenges, the LINEup team has already begun designing coordination protocols and distributed analysis strategies to facilitate decentralised yet aligned research efforts. Nonetheless, it is important to acknowledge that the resources and infrastructure available to the

LINEup project go well beyond what is typically accessible to individual researchers or small teams. With appropriate adaptation, the procedures developed by LINEup could be shared with the wider research community, enabling other teams to replicate and build upon them when working with the mapped datasets.

Timing of measurement

A second major challenge that emerged from the mapping concerns the identification and selection of longitudinal waves - that is, determining which datasets and survey years researchers should focus on in each case.

While the mapping provides a snapshot of available datasets and their reference years, making meaningful use of this information requires careful interpretative work. It is not sufficient to simply observe that the same set of competences was measured in different countries in the same year, nor can one assume that the measures are comparable (see further discussion in the next section).

The first layer of complexity arises from the combination of the year of data collection and the age/grade of the students involved. Researchers must assess whether this combination corresponds to a full birth cohort or whether it includes students who are ahead of or behind their expected grade level. Furthermore, institutional features - such as the presence of differentiated school tracks - can affect sample composition or even result in students taking different assessments at the same grade level, as observed in some countries.

These sampling and institutional differences affect the ability to address all three research questions. For RQ1, they influence the comparability of learning trajectories across countries; for RQ2, they constrain the possibility of aligning standardised test results with teacher-assigned grades in a coherent time frame; and for RQ3, they impact the consistency of student grouping across school levels, which is crucial for analysing segregation and composition effects.

Intersecting with these issues is the very nature of the longitudinal data. In some cases, the data come from true cohort panel designs that follow the same individuals over time. In others, longitudinality is reconstructed from a series of repeated cross-sectional assessments. Although the latter design normally precludes tracking individual students - as each wave samples a different population - in some cases individual tracking is possible when data collection covers the same groups or the population and stable personal identifiers can be matched across waves. Finally, repeated cross-sectional data may also be exploited to build synthetic panels, in which cohort-level averages are linked across waves to approximate the insights of a true longitudinal study. However, the methodological implications differ substantially, particularly in terms of sampling design, attrition, and the interpretation of change over time.

For RQ1, all these elements are essential: researchers must understand not only *what* is measured, but *who* is measured, *when*, and under *what institutional conditions*. Addressing these issues ideally precedes the harmonisation process, yet such understanding cannot be fully provided by the mapping alone.

Finally, an additional layer of complexity stems from the impact of the COVID-19 pandemic and the related lockdowns, which disrupted data collection in several countries - notably in 2020 and 2021 - and had likely long-lasting impacts on student performance. These effects appear to have varied across socio-economic groups and national contexts. For this reason, the pandemic must be taken into account when analysing learning curves (RQ1), as it may introduce structural breaks or

distortions affecting comparability. At the same time, it is likely that comparative longitudinal estimates of the pandemic's impact will become a relevant research topic in their own right.

Presence and measurement of key variables

Our mapping endeavour also highlighted a set of challenges related to the presence and measurement of key variables required to investigate the longitudinal dynamics of educational inequalities. This section discusses these issues in greater depth, as they help clarify the scope, opportunities, and limitations of the mapping exercise itself.

A first common issue concerns the difficulty of capturing the complexity and articulation of many variables by simply recording their presence or absence. Once researchers begin to reflect on actual comparability, the initially promising picture - regarding the widespread availability of information on student achievement and background factors - appears more nuanced. The same variable may be present across multiple datasets, yet differ significantly in how it is defined, collected, or reported. This comparability often weakens both across datasets and within the waves of a single longitudinal study.

In this sense, the mapping can only serve as a starting point. There is a real risk of overestimating what is truly comparable. This limitation will be addressed more directly in the feasibility and harmonisation phases, but it is important to already acknowledge it here: even highly detailed mapping remains a provisional tool that can only point toward possible research paths, whose concrete feasibility depends on further validation steps.

An example of this uncertainty concerns the gap between the mapping of student achievement variables - typically the dependent variables in longitudinal analyses - and the actual procedure involved in their construction. During the mapping phase, documentation typically stops at the subject assessed and the reporting scale. However, the informational content and comparability of achievement variables - such as mathematics - can vary considerably due to a number of underlying factors.

For example, the specific topics covered may vary, even when grade level and subject are nominally the same. Tests can be curriculum-oriented, or based on the literacy concept, meaning that they aim to assess competences considered essential for full participation in society. The purpose of the test also plays a role: evaluative assessments (e.g. for school accountability) often include easier items to ensure broad success, while diagnostic or research tests may use harder items to better distinguish performance levels. Another factor is the stakes attached to the test. In high-stakes contexts, where results carry important consequences for students and/or schools, test performance may be influenced by strategic behaviour or test-taking pressure, thus biasing the measurement of the underlying competences. Conversely, in low-stakes contexts, student motivation may be lower, affecting score reliability and increasing missing cases.

Additional differences may emerge in the scoring and validation procedures used. Thus, competence scores might be reported on more or less refined statistical scales. All of these elements shape the quality and comparability of the achievement measures and must be taken into account when analysing longitudinal trends. For instance, even when at least three waves are available in the same learning domain, allowing for growth curve modelling, biased or inconsistent measurement in one wave may compromise the interpretation of the learning trajectory. These issues directly affect our RQ1 (see above).

Similar limitations are even more relevant concerning achievement measures not based on standardised tests, such as grades. This is particularly relevant for our RQ2, where the core variable - grading bias - is calculated as the difference between teacher-assigned grades and standardised scores. Any inconsistency in how either of these measures is defined, be it across countries, schools, or subgroups, can severely compromise comparability.

The problem becomes even more pronounced for non-cognitive variables (only partially covered by our mapping), such as students' interest in a subject. In this domain, the heterogeneity of measurement instruments and constructs is even greater, often undermining comparability altogether. While this may raise concerns for some readers, it is important to acknowledge these limitations openly, as they reflect structural challenges in educational measurement that cannot be resolved through mapping alone.

A further area of concern, briefly mentioned above, relates to the hierarchical structure of the data and the detail of information available on the sampling design. These aspects are crucial to avoid taking variable comparability as merely a measurement issue. There is indeed a substantial difference between population-level data that allow for the identification of students within their classrooms and schools, and datasets that only provide individual-level information without structural links.

The mapping shows that many datasets fall somewhere between these extremes, which leads to further complications. For instance: How many classes were sampled within each school? Were all students in those classes tested? These questions directly impact the feasibility of addressing our RQ3, particularly in relation to the analysis of school- and class-level segregation and composition effects across different education systems.

Finally, it is worth noting one last issue, that remained largely implicit in our mapping, despite being highly relevant for the feasibility study: a stronger survey methodology perspective should be integrated into comparative research. This includes attending to metadata such as test administration procedures, the source of background information (e.g. administrative vs. self-reported), the impartiality of data coders, and how consistency across waves was ensured. While often considered "minor" details, these factors can substantially affect data quality and comparability - especially when dealing with datasets not originally designed for international comparison.

In sum, this section reinforces a key message already outlined above: mapping is a necessary but not sufficient step. To move from the *potential* comparability of datasets to *substantive* and *analytically meaningful* comparability, researchers must examine in depth how key variables are defined, measured, and embedded in their institutional and methodological contexts.

While the mapping has revealed promising opportunities for research, unlocking their full potential requires further work on data accessibility, harmonisation, and methodological alignment. These efforts will be central to the upcoming feasibility study, which is not only intended to implement these actions, but also to develop protocols that can be reused by the wider research community in the future.

A crucial takeaway from this exercise is that comparative research on the evolution of educational inequalities demands shared measurements and study designs across countries. Large-scale international assessments such as PISA offer strong cross-national comparability, but they lack the longitudinal dimension needed to trace individual trajectories or explore causal mechanisms. What remains missing is a combination of the cross-country comparability of studies like PISA with the analytical power of longitudinal data - a gap that future data infrastructure efforts should aim to bridge.

5. Visualising Longitudinal Data: The LINEup Education Data Explorer

To complement the mapping of longitudinal datasets and maximize its usability, the LINEup project developed an open-access, interactive system to visualize and explore the results. The LINEup Education Data Explorer transforms structured metadata of mapped datasets into a navigable online tool designed for researchers, policymakers, and educational stakeholders. This section describes the system's objectives, core functionalities, target users, and next development steps.

General Overview

The LINEup Education Data Explorer was developed as a direct continuation of the mapping activities. It provides a public-facing, interactive platform that enables users to navigate and explore longitudinal datasets on schooling across 32 European countries²⁵.

The Explorer combines a map-based and list-based interface with layered information and dynamic filters, offering users multiple paths for identifying their datasets of interest. All datasets included in the platform were described by project partners and comply with the mapping inclusion criteria reported above.

The Explorer is designed to meet the needs of two main user groups:

- **Educational stakeholders** (e.g. policymakers, NGOs, ministries) who benefit from clear summaries, comparative visualisations, and intuitive access to key dataset characteristics.
- **Researchers and analysts** seeking more detailed metadata, advanced filtering options, and information about dataset accessibility for academic or policy-oriented analyses.

To support these diverse needs, the platform incorporates the following key features:

- **Open access** with no mandatory registration, ensuring inclusivity and transparency.
- **Integrated analytics tools** for monitoring usage and interaction patterns.
- **User feedback options**, enabling contributions such as suggesting new datasets, reporting errors, or requesting clarification.
- **A dedicated back end** to support ongoing system maintenance and future updates.

The LINEup Education Data Explorer represents a core output of the mapping process and an evolving infrastructure that can be updated, expanded, and enriched over time.

Structure and Functionalities

The Explorer builds directly on the structured metadata collected during the mapping process (as described in Section 3 above), making the information accessible through visual and textual

²⁵ Although the Early Childhood Longitudinal Study (ECLS) from the United States is described in the mapping as an international best practice due to its quality and policy relevance, it is not included in the LINEup Education Data Explorer. The current version of the Explorer is limited to datasets from 32 European countries. Moreover, ECLS is not representative of the broader longitudinal data landscape outside Europe and was therefore not considered suitable for comparative visualisation purposes.

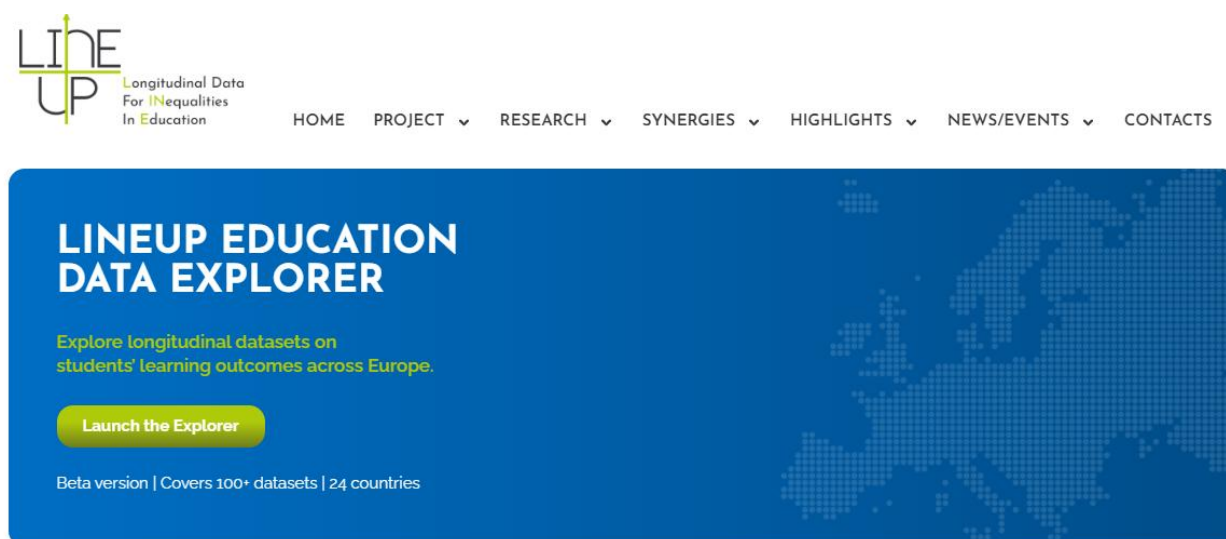
interfaces. It is designed to support both immediate access to high-level information and deeper exploration of individual datasets through layered content and intuitive navigation tools.

The **landing page** (Figure 5) provides the essential background information before users access the full dataset interface. It introduces the LINEup Education Data Explorer as an open-access platform for discovering and comparing longitudinal data on primary and secondary education across Europe.

The landing page is organised into four informative sections:

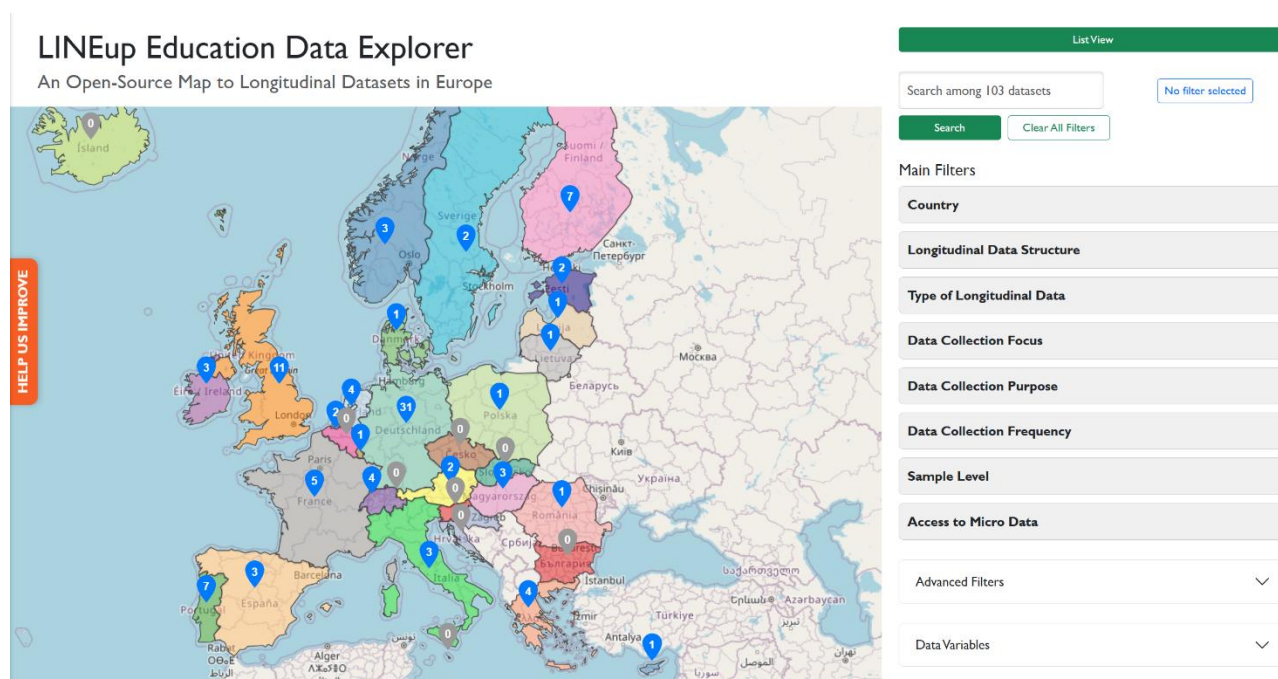
- **What is it?** – Describes the Explorer as a map- and list-based interface giving access to metadata from over 100 datasets, with a focus on academic achievement, learning outcomes, and equity-relevant contextual variables.
- **What kind of data is included?** – Details the types of longitudinal data featured, including survey, administrative, repeated cross-sectional, and hybrid sources, alongside variables such as design, measurement tools, data linkage, and access conditions.
- **Who is it for?** – Outlines the different use cases for researchers, policy analysts, educational stakeholders, and institutions.
- **How can you contribute?** – Explains how users can suggest new datasets, report errors, recommend improvements, or request further information, reinforcing the participatory and evolving nature of the system.

Figure 5: The landing page of the LINEup Education Data Explorer.



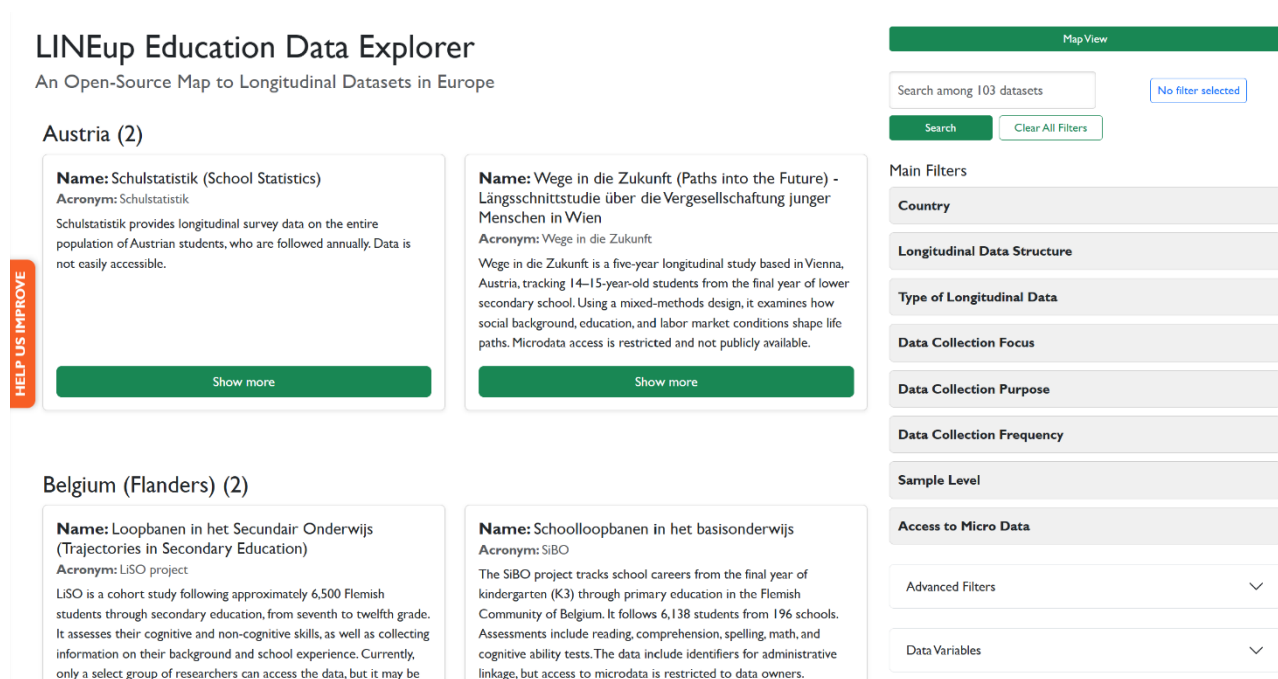
From this page, users can access the full Explorer via the “**Launch the Explorer**” button, which leads to the main interactive environment. The **default view** is a map of Europe, where clickable country markers indicate the number of datasets identified in each country. Hover-over and zoom-in functionalities enable quick access to country-level summaries and more detailed insights (see Figure 6).

Figure 6: LINEup Education Data Explorer - Map View.



Users can also switch to a **list view**, where datasets are sorted alphabetically by country and presented in text-based format (see Figure 7). A toggle button allows switching between views at any time, offering flexibility based on user preference.

Figure 7: LINEup Education Data Explorer - List View.



To ensure usability, each dataset entry is structured into **three progressive levels of information**:

- **Level 1 – Basic overview:** Includes essential metadata such as dataset name, acronym, short description, responsible organisation, longitudinal data structure, type of longitudinal data, data collection purpose, data collection focus, data collection frequency, starting and ending year of data collection, and sample level.
- **Level 2 – Detailed description** (accessible via a “Show more” button): Presents expandable sections covering:
 1. School grades and coverage (including post-school follow-up)
 2. Students’ skills and achievement (domains, measures, and administration)
 3. Sampling strategy (type, weights, unit, size)
 4. Linkability (at individual level)
 5. Accessibility (microdata availability, restrictions, official website)
- **Level 3 – Variables available:** Organised into four expandable categories:
 1. Student-level information
 2. Household-level information
 3. Teacher-level information
 4. School/class-level information

This **multi-level structure** ensures that **users with different levels of expertise** – from general stakeholders to specialised researchers – can access information at the desired depth, supporting both broad overviews and fine-grained analyses.

To optimise dataset discovery, the Explorer includes an integrated **search and filtering system** on the right-hand side of the main page, organised as follows:

- **Main filters** (always visible): Allow users to select datasets by country, longitudinal data structure, type of longitudinal data, data focus, collection frequency, access conditions, and more.
- **Advanced filters** (expandable): Include more detailed attributes such as school grades covered, type of skills measured, administration methods, and sampling criteria.
- **Data variables filters** (expandable): Help users identify datasets containing specific variables across four domains: student, household, teacher, and school/class characteristics.

In addition, a **keyword search bar** allows direct search by dataset name, acronym, or specific topic of interest. On the left-hand side, a permanent “**Help us improve**” button offers users the opportunity to contribute feedback, such as proposing new datasets, flagging outdated or inaccurate information, suggesting improvements, or submitting specific queries.

By offering structured access to hard-to-find metadata, our Explorer helps reduce information asymmetries and lowers the barrier to evidence use in education - enabling more inclusive, comparative, and policy-relevant research.

Next Steps for Development

To complete the development of the LINEup Education Data Explorer, a Beta version is currently available. This initial release includes the core features reported above: an interactive map displaying the available datasets, integrated search and filtering tools, and access to structured dataset metadata. Ensuring the system’s technical robustness and scalability is essential for future enhancements.

The Beta version is therefore undergoing a two-phase testing process. Consortium partners have assessed its usability - including navigation, visualisation, and filtering functions - while selected stakeholders, such as data owners, are currently focusing on verifying the accuracy and completeness of the metadata.

The final release of the Explorer will incorporate iterative improvements based on user feedback, with the aim of enhancing usability, strengthening metadata validation, and introducing additional features such as expanded search capabilities and, eventually, API integration. The platform will be officially launched during the interim LINEup project event in Athens (October 17th)²⁶, accompanied by a dedicated outreach campaign. Community inputs collected during and after the event will inform subsequent rounds of refinement.

In the longer term, the project will establish mechanisms to ensure its sustainability and ongoing relevance, including:

- Periodic metadata verification with dataset providers.
- Community-driven updates based on user contributions.
- Regular validation and integration of user-submitted information into the mapping dataset.

To ensure both timely updates and transparency for research use, the project will adopt a versioning system for the underlying dataset. While individual entries and metadata may be updated continuously, major updates will be released periodically as documented versions (e.g., v1, v2) and archived for reproducibility. This approach allows users to always access the most recent data through the Explorer, while also enabling researchers to reference specific versions of the dataset in publications and analyses.

As such, the Explorer not only consolidates the outputs of LINEup mapping of existing longitudinal data, but also lays the groundwork for future policy dialogue, research collaboration, and enhanced data use across Europe.

²⁶ More information available here: <https://www.lineup-project.eu/lineup-in-athens-2025-understanding-educational-inequalities-through-longitudinal-data/>

6. Discussion: Interpreting the Research Potential and Limitations of Mapped Datasets

Our mapping contributes to current and future social research by providing a replicable framework for assessing the quality and relevance of educational longitudinal datasets. As demonstrated, a structured validation process was developed and applied to ensure transparency and consistency in dataset classification. The result is not only a curated and reliable data foundation, but also a methodology that can be reused and adapted for future updates or extensions.

The adopted classification scheme is based on an inclusive definition of longitudinal data, as outlined in Section 2. This definition goes beyond traditional panel designs and encompasses data sources that allow for time-based analysis without necessarily tracking the same individuals over time by design – such as repeated cross-sectional surveys and pseudo-panels as well as longitudinal administrative data, which follow individuals through linked records across multiple years. Our descriptive analysis provided an overview of the 104 selected datasets, focusing on their analytical potential for studying educational inequalities.

Several key dimensions explored during the mapping led to a set of relevant descriptive results:

- **Methodology:** Most datasets have a pure longitudinal design (68.0%). About two-thirds use student samples rather than population data, with 44.7% using random sampling. Overall, 53.4% of datasets include administrative data - either alone (19.4%) or combined with survey sources (34.0%).
- **Information Collected:** 76.7% of datasets include standardised tests on cognitive skills, and 68.9% include school grades, with over half (53.4%) providing both. Literacy (70.9%) and numeracy (69.9%) are the most frequently assessed domains. Key student characteristics like gender, age, and migration background are often available, while teacher- and school-level variables are more limited. Basic school information (e.g. type, track) is more common than detailed contextual data (e.g. school composition, geo-referencing).
- **Countries Included:** Most datasets (62.1%) have a national scope. The highest number of datasets were identified in Germany (30.1%) and the UK (10.6%), although a number of these have a regional rather than national coverage. Most Central and Eastern European countries display a structural lack of accessible longitudinal data. Southern European countries present more heterogeneous situations: some, like Italy, have population-scale longitudinal data repeated annually, while others lack accessible data or rely on small-scale samples.
- **Duration:** There is considerable variation in the number of observation points: around 20% of datasets include only two time points, while nearly a quarter (22.7%) span over 14 years. Grades 8, 9, and 10 are the most frequently covered. 47.1% of datasets span both primary and secondary education. Just over half of the datasets (51.4%) are still ongoing, helping ensure greater data recency.

Overall, the mapping shows that datasets analytical potential is promising, with many offering large samples, standardized tests, multiple waves, and background information. In sum, there is a relatively large number of datasets across Europe providing valuable information to study students' educational trajectories and social inequalities therein over time, from a longitudinal perspective. However, a more thorough assessment of dataset quality must consider not only the cross-cutting challenges already discussed, but also the specific research questions and analytical purposes the datasets are intended to support. The analysis presented in this report highlights three main challenges that limit what can currently be done with the available datasets:

- **Data access and documentation** remains often complex, due to limited and unclear information, difficulties in timely data acquisition, and significant barriers to cross-border transferability - often forcing parallel national analyses. While 26.2% of the datasets have highly restricted access or no access at all, for 13.6% of them the conditions for data access are not clear or not available. This situation severely limits data reuse and stands in contrast with the principles of open access and research replicability.
- **Timing of measurement** is often difficult to determine, due to the interplay of several factors such as student age, school track, and the type of longitudinal design (e.g. true panel vs. repeated cross-sectional).
- The **presence and measurement of key variables**, although essential to the mapping process, often oversimplify the issue of comparability. The same variable may differ substantially across datasets – and even across waves within the same dataset – in terms of definition, data collection methods, or reporting practices. This affects both student achievement measures (due to variations in test design, stakes, or scaling) and non-cognitive variables. The hierarchical structure of the data and differences in sampling design further complicate meaningful comparison.

To evaluate the suitability of the available datasets for different research questions and analytical goals, we introduce a three-dimensional framework (see Table 29). This framework considers:

- Whether the research has a primarily **descriptive goal** - seeking to map students' educational trajectories and related inequalities - or also an **explanatory goal**, focused on testing theories and assessing the influence of potential independent variables on those trajectories.
- Whether the study adopts a **single-country design**, analysing student trajectories within a national or regional context, or a **cross-country comparative design**, aiming to examine differences across education systems.
- In explanatory designs, whether the focus is limited to **students' ascriptive characteristics** (e.g., socio-economic background, gender, migration background), or also includes **contextual factors** (e.g., teacher characteristics or school composition) that may shape their trajectories.

Adopting such a scheme allows for a concise yet nuanced answer to the question of what can be done with the currently available datasets for longitudinal research on educational inequalities.

Table 29: A framework to assess the availability and quality of longitudinal datasets for the study of education inequalities

Main goal	Explanatory Focus	Single Country Research Design	Comparative Research Design
Descriptive	–	++	+
Explanatory	Ascriptive traits	+	-
	School agency/role	-	- -

Legend: ++ very good, + good with caveats, – limited, – – severely limited

While this framework provides a practical lens to assess dataset usability, its application also reveals several persistent limitations across the mapped datasets, which constrain their use for specific types of research.

First, the dataset collection allows for a solid description of students' trajectories over time - based on cognitive test results or school grades - at the *single-country level*. Most EU countries are covered, along with several non-EU European countries such as Norway, Switzerland, and the UK.

When research shifts to an *explanatory focus at the national level* - using students' ascriptive traits as independent variables - relevant information is generally available, though with an important caveat. Students' social origins can refer to different dimensions, such as social class (based on parental occupation or income) or parental education. As highlighted in the previous section, datasets vary in how many of these dimensions are included and how they are operationalised. The issue is less pronounced, though still present, when it comes to students' migration background: some datasets use citizenship, while others rely on the country of origin of the student or their parents. However, when the research aim is to assess not only the role of ascriptive traits but also *school agency/role* in shaping educational trajectories and inequalities, the situation deteriorates significantly. Very few datasets include even basic information on teachers or school-level characteristics that can be linked to individual students - thus limiting the possibility of analysing their potential effects on student outcomes over time.

Overall, in several countries, it is possible to investigate - *both descriptively and explanatorily* - the role of students' ascriptive traits in shaping their learning trajectories at the national level. However, opportunities to enrich these analyses by examining the contribution of school-level factors - such as resources or teacher characteristics and behaviours - remain rare.

The situation begins to change when the research design shifts to a *descriptive cross-country comparative goal*. In this case, the available datasets are useful but require careful harmonisation. Challenges arise especially when comparing data based on school grades with those including (or substituting) cognitive test results. Additional issues emerge from differences in student evaluation scales (e.g. proficiency levels vs. continuous grading) and in the cognitive domains assessed — for example, the concept of literacy or numeracy may vary significantly across instruments. In some cases, inconsistencies in variable definitions across waves introduce further obstacles to harmonisation within the same dataset over time. The most commonly available variable for cross-country comparison is parental education. However, even this requires careful harmonisation to ensure meaningful comparability across datasets.

Given the findings discussed above, it is not surprising that a *comparative research design* - combining *cross-country comparison with explanatory aims* and a focus on school-level factors - is currently feasible in only a very limited number of countries. This is mainly due to two factors: most datasets lack information on school or teacher characteristics, and where such data are available, there is significant heterogeneity in both content and operationalisation.

In sum, the analysis presented here points to at least **four avenues for future development** in the provision of longitudinal datasets to support research on student learning outcomes and related inequalities.

The first point is that, if longitudinal datasets are to support both policymakers and the scientific community in understanding and addressing educational inequalities, the lack of information on schools and teachers - i.e. on the agency component of the education system - represents a major limitation. On the one hand, neglecting potentially crucial school- or teacher-level factors risks mis-specifying explanatory models of inequality. On the other, it becomes impossible to assess whether

- and to what extent - specific interventions targeting school characteristics or teacher behaviours (e.g. teacher type, classroom composition) are associated with different student outcomes. Overall, the scarcity of information on how schools function limits the potential use of evidence by public decision-makers seeking to design informed policies. Of course, experimental research designs (e.g. randomised control trials) are best suited to identify causal effects, but they often lack external validity due to limited population coverage. In contrast, longitudinal datasets - especially those based on large samples or full population coverage - can offer important insights with stronger generalisability.

The second point is the need for EU institutions - especially Eurostat - to promote expert groups tasked with developing standard definitions for core variables used in the study of educational inequalities, particularly independent variables such as students' social origins. While several established classification schemes already exist, there is still no consistent convergence across datasets. Establishing common standards for dependent variables, such as student performance, is more challenging. However, progress could be made by harmonising conceptual frameworks (e.g. what skills are assessed) and aligning measurement scales to improve comparability.

The third point concerns data access. The fact that 26.2% of the mapped datasets have severe or complete limitations to microdata access, and for almost 14% of them there are no clear information available on data access, is a major barrier. At a minimum, for datasets owned by public institutions in EU member states (e.g. agencies, universities), it is crucial to promote good practices in microdata sharing—supported by a European access framework that includes safeguards to comply with national regulations.

The **final point** concerns the absence of a European-level data infrastructure aimed at tracking students' educational trajectories over. While international assessments such as PISA, TIMSS and PIRLS provide valuable cross-country data, they do not follow the same individuals longitudinally. At present, no EU-wide dataset exists that focuses specifically on education from a longitudinal perspective. As this report has shown, harmonising existing national datasets is possible to some extent, but significant limitations remain - particularly when aiming for cross-country comparisons with explanatory or policy-oriented goals. Rather than calling for the immediate establishment of a new pan-European longitudinal study²⁷, we recommend initiating a coordinated effort to align and strengthen existing data systems. This effort should build on established best practices and tools²⁸, and foster a structured reflection on the long-term feasibility and added value of a European data infrastructure for education. A pragmatic starting point could be the development of a repeated cross-sectional initiative, gradually evolving into a more integrated design capable of supporting robust comparative research on educational inequalities.

In conclusion, **Europe's landscape of longitudinal educational data is rich but fragmented**. Unlocking its full potential will require addressing key challenges - in harmonisation, data access and documentation, and inclusion of school- and teacher-level variables - to support more effective research and policymaking on educational inequalities.

²⁷ The creation of longitudinal data infrastructures at EU level is not without precedent. In other fields - such as labour market dynamics, income inequality, and health - Europe-wide longitudinal datasets or harmonised datasets already exist. Notable examples include the EU Statistics on Income and Living Conditions (EU-SILC), the European Union Labour Force Survey (EU-LFS), and the Survey of Health, Ageing and Retirement in Europe (SHARE). In the field of higher education, initiatives like EUROSTUDENT and EUROGRADUATE have advanced the harmonisation of student data across countries. Similarly, the European Social Survey (ESS) provides high-quality comparative data on social attitudes and behaviours.

²⁸ For example, the COORDINATE network, currently hosted at the University of Essex, contributes to the visibility and harmonisation of child and youth cohort data across Europe, and has developed a harmonisation toolbox to support this effort (see: <https://www.coordinate-network.eu/database>).

Examples of longitudinal dataset's best practices in the EU:

the Italian INVALSI and the German NEPS

To broaden the comparative perspective and illustrate best practices in longitudinal data collection on student learning outcomes, two exemplary datasets are reviewed here: INVALSI (Italy) and NEPS (Germany). These cases were selected for their methodological robustness, relevance to the study of educational inequalities, and complementary design features. While INVALSI adopts a repeated cross-sectional design with partial individual tracking, NEPS implements a true panel approach with multiple cohorts followed over time. Together, they represent two of the most advanced and informative educational data infrastructures currently available in Europe.

INVALSI – the National Institute for the Evaluation of the Education and Training System – is part of the Italian public education infrastructure²⁹. It conducts periodic and systematic assessments of students' knowledge and skills, as well as the overall quality of educational provision in schools. The INVALSI framework is closely aligned with the Italian National Curricular Guidelines and focuses on students' competences in literacy, numeracy, and English as a foreign language (including both listening and reading). These skills are assessed through standardized tests administered in different formats: paper-based in primary school and computer-based – using adaptive software – in secondary school. Tests are conducted in Grade 2 and Grade 5 (primary), Grade 8 (lower secondary), and Grades 10 and 13 (upper secondary).

The dataset has a two-level structure. At the first level, it includes population data (census data collected from the entire student population), covering students' skills, personal characteristics, parental background, and home possessions. At the second level, a national sample-representative at the regional level - adds information on teachers, collected through a dedicated ad hoc questionnaire.

INVALSI adopts a repeated cross-sectional design, in which data are collected from the population of interest at multiple time points, rather than by tracking the same individuals over time. In other words, each data collection wave includes a new, independent cohort of students (e.g., all 5th graders in Italy in a given year). Within the dataset, students are identified via a unique code, which allows researchers to examine trends at the population level (e.g., national average scores over time) and to partially reconstruct individual trajectories across primary and secondary school. This design is particularly suited to monitoring system-wide performance, identifying regional or demographic disparities, and evaluating policy impacts at the cohort level. Its main limitation is that students who leave the school system are no longer tracked, and those who repeat a grade are included in a subsequent wave, potentially compromising the continuity of individual trajectories.

The second example of excellence in educational longitudinal research is **NEPS**, the German National Educational Panel Study, launched in 2010³⁰. Data are collected by the Leibniz Institute for Educational Trajectories and are accessible to researchers upon request.

To provide timely data across different life stages, NEPS follows a multi-cohort design: multiple cohorts were first surveyed around 2010 and are followed up annually. These include: SC1 - Newborns; SC2 - Kindergarten; SC3 - Grade 5; SC4 - Grade 9; SC5 - University students; SC6 - Adults; SC7 - a second cohort of Newborns (starting in 2031); SC8 - a second Grade 5 cohort (since 2022). In total, NEPS has accompanied more than 70,000 participants and about 50,000 contextual persons. Each cohort allows researchers to follow individuals over time, focusing on the development of the following competences: spelling, reading, and listening comprehension in

²⁹ INVALSI: <https://serviziostatistico.invalsi.it/en/>

³⁰ NEPS: <https://www.neps-data.de>

German; English language skills; and, for students with a migration background, competences in their heritage language (currently assessed only for Russian and Turkish speakers). Additional domains include mathematics, science, ICT skills, and cognitive problem-solving.

All competences are assessed through standardized procedures aligned with international benchmarks, such as PISA frameworks for mathematics and science. NEPS also collects school grades and contextual data through questionnaires administered to students, parents, teachers, and school principals. Thanks to its coordinated study design and regular data collection schedule, NEPS makes it possible to distinguish between cohort, age, and period effects - a key feature for robust longitudinal analysis.

INVALSI and NEPS stand out as **best-practice frameworks** due to their **comprehensive design**. Both enable researchers to trace the emergence of social inequalities and identify early predictors of later academic and life outcomes. Their integration of individual-level and contextual data (family, classroom, school) allows for nuanced, multi-level analyses of learning and development. Together, they illustrate the value of combining large-scale assessment with longitudinal tracking, and serve as benchmarks for improving longitudinal education data infrastructures across Europe.

7. Conclusions and Implications for Research, Policy, and Practice

This report presents a comprehensive mapping of longitudinal datasets on student learning outcomes across 32 European countries. Through a rigorous three-step strategy, the LINEup project identified 104 relevant datasets in 24 countries and analysed their structure, coverage, and comparative potential. This effort produced a structured inventory of available data sources and offered deeper insights into their analytical and policy relevance.

Summary of Key Findings

Our mapping exercise confirmed the existence of a substantial body of longitudinal data on student learning outcomes across Europe. The 104 datasets identified differ considerably in design, scope, and quality, but together offer a valuable - albeit uneven - empirical basis for analysing learning trajectories and educational inequalities.

Most datasets provide repeated measures of student performance, particularly in literacy and numeracy. By contrast, non-cognitive outcomes, such as student well-being or motivation, are inconsistently measured in connection with longitudinal data on student performance and rarely comparable across countries. Basic background variables such as gender, parental education, and migration status are included in most datasets, though they are often operationalised differently across sources. However, there is substantial variation in longitudinal depth, grade-level coverage, and the inclusion of contextual variables. The duration of follow-up ranges from 2 to over 14 years, but only a minority of datasets allow for continuous individual-level tracking across multiple levels of schooling - from early education through the end of secondary school.

The mapped datasets adopt a variety of designs - pure panels, repeated cross-sections with unique identifier, and pseudo-panels - and rely on surveys, administrative records, or hybrid approaches. While administrative data offer strong potential for large-scale longitudinal analysis, they remain underused and are often difficult to access or repurpose for research, as they were not originally intended for this purpose.

A notable pattern is the uneven grade coverage: middle grades (particularly grades 8–10) are well represented, while early and late stages of schooling are more fragmented. This distribution reflects national assessment schedules and key educational transition points. Nonetheless, nearly one-third of the mapped datasets (30.1%) include observations starting from kindergarten, providing a valuable, though still limited, entry point for analysing early learning trajectories.

Few datasets include the contextual and process-related variables - such as classroom, teacher, or school-level data - needed to understand the broader dynamics of inequality and the role played by the functioning of education systems. This lack is problematic not only because it limits our understanding of the mechanisms behind the development of inequality, but also because it means that existing evidence offers little guidance to policymakers on key levers that fall within their control.

In terms of comparative potential, only a small subset of datasets displays the methodological features and variable richness necessary for cross-country analyses. In several countries, particularly Germany, Italy, France, and the United Kingdom, the datasets stand out for their national coverage, metadata quality, openness and potential for multilevel analysis, making them strong candidates for future harmonisation efforts. The theoretical and methodological framework (see Section 2), together with the classification and validation system (see Section 4.1), and the dataset analysis and discussion (see Sections 4.2 and 4.3), now represent a key contribution of the project. They lay the foundation for a more structured, transparent, and interoperable use of these datasets in future research.

Implications for Research, Policy, and Practice

The implications of these findings are significant. For the **research community**, our mapping provides a solid foundation for designing comparative analyses of student trajectories and outcomes. At the same time, it highlights persistent structural and technical barriers that researchers must navigate - such as fragmented access regimes, lack of harmonised definitions, insufficient documentation, and the frequent absence of contextual variables needed for multilevel or process-oriented analyses.

For **policymakers and data owners**, our results call for targeted investments in data infrastructures - not only to expand their coverage, but also to improve access, quality, transparency, and long-term reusability. This includes systematically incorporating school- and teacher-level variables (where feasible and legally permitted), strengthening metadata and documentation standards, and enabling secure linkages between different data sources (e.g. test scores, school registries, administrative records). Capacity-building initiatives targeting both data producers and users—such as training on documentation, access protocols, and harmonisation practices - could significantly enhance the usability of existing sources. In addition, legislation enabling the use and linkage of administrative data for research purposes remains a key policy lever. Without addressing these issues, it will remain difficult to produce policy-relevant evidence to monitor disparities, assess the impact of regulations or reforms, or design effective responses.

More broadly, for **educational stakeholders**, our findings point to the need for a shift toward more integrated and research-oriented data governance. Student-level outcomes should be meaningfully linked to their social and institutional contexts to enable robust and reliable analysis of educational inequalities. This also requires more balanced data infrastructures - those that integrate information on achievement with school context, teachers, peer composition, and socio-emotional dimensions. These aspects are often underrepresented, yet they are crucial for understanding the broader dynamics of learning. In addition, more effort is needed to translate research findings from these datasets into formats and messages that are accessible and relevant to practitioners and local stakeholders. Supporting this transformation will require closer collaboration between researchers, national ministries, statistical offices, and EU-level institutions.

In the longer term, the development of a shared European longitudinal data infrastructure should be considered - building on existing best practices and successful models from other policy domains. This would support more systematic and comparable research on educational inequalities across countries. In this regard, the typologies, validation framework, and metadata descriptors developed through this mapping can serve as a reference model for future efforts toward more consistent, transparent, and interoperable education data systems across Europe³¹.

Next Steps

Ultimately, the LINEup mapping activity is a foundational step toward a more strategic and inclusive approach to using longitudinal data in European research and policymaking. However, the mapping alone is not sufficient. The real value of longitudinal data for education lies in their responsible use, careful harmonisation, and meaningful integration into broader systems of knowledge production and decision-making. By making these datasets publicly visible and available, promoting dialogue between data producers and users, and developing tools for their use, the LINEup project aims to

³¹ The Education Data Explorer developed by the project represents a first step in this direction, helping make metadata more accessible and reusable for both researchers and policymakers.

contribute to a more connected, transparent, and equity-oriented educational data ecosystem in Europe. As such, it offers a concrete contribution to advancing inclusion, excellence, and innovation across European educational systems.

Looking ahead, the LINEup project will address these methodological and infrastructural challenges through a feasibility study on data harmonisation. This study will explore how selected longitudinal datasets can be aligned using shared coding schemes and protocols, enabling more robust cross-country comparisons of student background and learning outcomes (Deliverable 3.2). LINEup will also test analytical models across national contexts, assess their replicability, and inform the development of guidelines for future comparative research.

These insights will inform the qualitative components of the project, helping to guide the selection of national case studies and to strengthen the connection between large-scale data and student lived educational experiences. In particular, they will support a deeper investigation into the phenomenon of disengagement - focusing on students who disappear from longitudinal datasets over time. While these dropouts are often treated as a technical issue of panel attrition, they may in fact reflect critical educational processes, such as school exclusion, silent disengagement, or early dropout. WP4 and WP5 will explore these dynamics further, aiming to bring visibility to the social realities behind missing data.

In parallel, the project will continue to update, expand and refine the LINEup Education Data Explorer by promoting cooperation with data owners to improve metadata quality, support secure linkage protocols, and facilitate access through trusted research environments. By improving its filtering options and metadata descriptors, the Explorer will become an increasingly valuable tool for researchers and educational stakeholders seeking to navigate the European education data landscape.

Outlook on Harmonisation Activities - Advancing from Mapping to Comparative Use

While this report focuses on the mapping of existing longitudinal datasets, a further step toward cross-country comparability is currently under way through a dedicated feasibility study on data harmonisation. This work aims at assessing the technical and conceptual conditions under which selected datasets could be aligned and used for comparative research on educational inequalities. The results of this analysis will be presented in a Feasibility Study on Data Harmonisation (due in January 2026).

As a preparatory step, a subset of eight national datasets has been selected to explore the feasibility of harmonising core variables and indicators related to students' background and learning outcomes. The selection includes datasets from countries represented in the LINEup consortium, as well as others chosen for their analytical relevance and open-access availability³². This set of datasets also helps address a significant gap in the literature by strengthening the longitudinal evidence base for countries in Southern and Southwestern Europe, which remain underrepresented in cross-country studies on school achievement. This harmonisation effort will ultimately support cross-country analyses of how educational inequalities emerge and evolve throughout primary and secondary education, with a focus on achievement gaps related to students' socio-economic background, gender, and migration status.

³² The selected datasets are: INVALSI (Italy); DEPP 2007 and DEPP 2011 (France); Education Data at the Catalan Department of Education – RALC (Spain); NEPS SC2 and SC3 (Germany); MySchool (Greece); Netherlands Cohort Study (The Netherlands); ENEB and ENES (Portugal); Millennium Cohort Study (United Kingdom).

Our preliminary work suggests that background variables, such as age, gender, or parental education, are more amenable to harmonisation, despite different levels of effort and final information potential. On the other side, outcome variables are more complex to harmonize due to differences in assessment content, structure, and scoring systems. At the same time, this early phase has already revealed several operational and methodological challenges, including issues of data access, variation in data collection timelines, and the heterogeneity of educational systems and measurement instruments.

These insights will serve as the foundation for the upcoming deliverable, which will assess technical options, identify barriers, and propose shared strategies to support comparative research on educational inequalities in Europe.

8. References

- Blossfeld, H.-P., Buchholz, S., Skopek, J., & Triventi, M. (2016). *Models of Secondary Education and Social Inequality: An International Comparison*. Edward Elgar Publishing.
- Boudon, R. (1974). Basic mechanisms generating inequality of educational opportunity. In T. N. Clark (Ed.), *Education, Opportunity, and Social Inequality. Changing Prospects in Western Society*. Wiley Series in Urban Research.
- Bukodi, E., & Goldthorpe, J. H. (2013). Educational attainment - Relative or absolute mobility in England? *British Journal of Sociology*, 64(3), 393–419.
- Campos, D. G., Ding, Y., Goian, A.-M., Koskinen, A., Lokken, J., Mannsaaker, T., Munkácsy, B., Papadopoulos, D., Patsis, P., Rolfe, V., Tóth, E., Scherer, R., & Olsen, R. (2025). *Longitudinal research on educational inequalities in Europe: A scoping review of trends, gaps, and methodological approaches* (Deliverable D2.1). MapiE Project (Horizon Europe).
- Carnoy, M., & Rothstein, R. (2013). *What do international tests really show about U.S. student performance?* Economic Policy Institute.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of Educational Opportunity (Coleman Report)*. U.S. Department of Health, Education, and Welfare.
- Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum*, 1989(1), Article 8.
- European Commission: European Education and Culture Executive Agency. (2025). *Key data on early childhood education and care in Europe 2025 – Eurydice report*. Luxembourg: Publications Office of the European Union.
- European Commission (2025). *Basic Skills and STEM Action Plan to Support Education and Training*. European Education Area, Directorate-General for Education, Youth, Sport and Culture.
- European Commission (2024). *Education and Training Monitor 2024*. Directorate-General for Education, Youth, Sport and Culture. Luxembourg: Publications Office of the European Union.
- Council of the European Union (2022). *Council Recommendation on Pathways to School Success (2022/C 243/01)*. *Official Journal of the European Union*.
- European Commission. (2022). *Quality investments in education and training: Investing in skills for better socio-economic outcomes*. Publications Office of the European Union.
- Kampylis, P., Fragkiadaki Theodoroulea, M., Kandila, M., Cholezas, I., Mobilio, V., Sampson, D., Lievore, I., Mauro, V., Gunzelmann, S., Lanoë, M., Maurya, P., Moulin, L., Ortiz, L., Passaretta, G., & Moreira, P. (2024). *Tracing educational inequalities in primary and secondary schools: Insights from a systematic review of longitudinal and repeated cross-sectional studies* (Deliverable 2.1). LINEup Project, Horizon Europe.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 Assessment Framework*. TIMSS & PIRLS International Study Center, Boston College.
- NCES (2023). *Early Childhood Longitudinal Study (ECLS)*. U.S. Department of Education, Institute of Education Sciences. Retrieved from <https://nces.ed.gov/ecls>
- OECD (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. OECD Publishing.
- OECD. (2019). *Measuring What Matters for Child Well-being and Policies*. OECD Publishing.

OECD. (2021). *21st-Century Readers: Developing Literacy Skills in a Digital World*. OECD Publishing.

Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press.

Symonds, J. E., Chzhen, Y., Kaye, N., Dominy, J., Campbell, C., Sykes, C., Bastuğ, S. I., Fiasconaro, S., & Heydari Barardehi, I. (2025). A metareview of research on educational inequality and socio-economic disadvantage. *Education Sciences*, 15(6), 740.

Scott, J. (2005). Sociological approaches to educational data. In S. Ball (Ed.), *The RoutledgeFalmer Reader in Sociology of Education* (pp. 112–124). Routledge.

Tourangeau, K., Nord, C., Lê, M., Sorongon, A. G., & Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Combined User's Manual for the ECLS-K Eighth-Grade Data Files and Electronic Codebook (NCES 2009-004)*.

UNESCO (2021). *Global Education Monitoring Report*. UNESCO.

UNESCO-UIS. (2018). *Education Data Quality Guidelines*. UNESCO Institute for Statistics.

UNICEF Innocenti Research Centre. (2012). *Measuring child poverty: New league tables of child poverty in the world's rich countries*. Innocenti Report Card 10. UNICEF Office of Research, Florence.

Walby, S. (2007). Complexity theory, systems theory, and multiple intersecting social inequalities. *Philosophy of the Social Sciences*, 37(4), 449–470.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Santos, L. B. S., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

9. Appendices

9.1 List of mapped and excluded datasets

I Part - Mapped Datasets

Country	Name	Acronym
AT (Austria)	Schulstatistik (School Statistics)	Schulstatistik
	Wege in die Zukunft (Paths into the Future) - Längsschnittstudie über die Vergesellschaftung junger Menschen in Wien	Wege in die Zukunft
BE (Flanders)	Loopbanen in het Secundair Onderwijs (Trajectories in Secondary Education)	LiSO project
BE (Flanders)	Schoolloopbanen in het basisonderwijs	SiBO
CH (Switzerland)	Determinanten der Ausbildungswahl und der Berufsbildungschancen	DAB panel study
	The Swiss Survey on Children and Youth - Cohort 1	COCON (C1)
	Transitions from Education to Employment	TREE II
	The Zurich Project on the Social Development from Childhood to Adulthood	Z-PROSO
CY (Cyprus)	Short- and long-term effects of the home learning environment and teachers on student achievement in mathematics: a longitudinal study	-
DE (Germany)	National Education Panel Study (Nationales Bildungspanel) - Starting Cohort Grade 5 (2010; SC3)	NEPS (SC3)
	National Education Panel Study (Nationales Bildungspanel) - Starting Cohort Grade 9 (2010; SC4)	NEPS (SC4)
	National Education Panel Study (Nationales Bildungspanel) - Starting Cohort Kindergarten (2010; SC2)	NEPS (SC2)
	National Education Panel Study (Nationales Bildungspanel) - Starting Cohort Newborns (2012; SC1)	NEPS (SC1)
	National Education Panel Study (Nationales Bildungspanel) - Starting Cohort Grade 5 (2022; SC8)	NEPS (SC8)
	Social Integration of Migrant Children – Uncovering Family and School Factors Promoting Resilience	SIMCUR
	TwinLife: Eine längsschnittliche Zwillingsfamilienstudie (Cohort 1)	TwinLife (C1)
	TwinLife: Eine längsschnittliche Zwillingsfamilienstudie (Cohort 2)	TwinLife (C2)
	Socio-Economic Panel	SOEP
	Mehrsprachigkeitsentwicklung im Zeitverlauf	MEZ
	Kompetenzerwerb und Lernvoraussetzungen	KuL
	Children of Immigrants Longitudinal Survey in Four European Countries – DE	CILS4EU-DE
	Refugees in the German educational system (Bildungswege von geflüchteten Kindern und Jugendlichen) Refugee Cohort 1	ReGES (RC1)
	Refugees in the German educational system (Bildungswege von geflüchteten Kindern und Jugendlichen) Refugee Cohort 2	ReGES (RC2)
	Panel Analysis of Intimate Relationship and Family Dynamics	Pairfam
	Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter	BiKS 8-18
	Konzeptionelle Weiterentwicklung von Lern- und Übungszeiten durch kooperative Lernformen	StEG-Tandem
	Die Rolle familialer Unterstützung beim Erwerb von Diskurs- und Schreibfähigkeiten in der Sekundarstufe I	FUnDuS

	Bildungssprachliche Kompetenzen - Anforderungen, Sprachverarbeitung und Diagnostik	BiSpra II
	IQB Bildungstrend Mathematik (Trends in Student Achievement)	IQB (Mathematik)
	IQB Bildungstrend Primarstufe (Trends in Student Achievement)	IQB (Primarstufe)
	IQB Bildungstrend Sprachen (Trends in Student Achievement)	IQB (Sprachen)
	Inklusion in und nach der Sekundarstufe in Deutschland	INSIDE
	Aufwachsen in Deutschland Alltagswelten	AID:A
	Bildungsentscheidungen und Bildungsverläufe vor dem Hintergrund struktureller Veränderungen im Berliner Sekundarschulwesen	BERLIN-Studie
	Motivation Jugendlicher in der Schule	MOVE
	Zur Bedeutung der kognitiven und motivationalen Überzeugungen von Lehrkräften und ihres Instruktionsverhalten im Unterricht für die Überzeugungen und den Lernerfolg von Schülerinnen und Schülern	TEACH!
	Perceptions of Inequality and Fairness among Students	PerFair
	Vergleichsarbeiten (Vera)	VERA
	Transformation des Sekundarschulsystems und akademische Karrieren (2002)	TOSCA (2002)
	Transformation des Sekundarschulsystems und akademische Karrieren (2006)	TOSCA (2006)
DK (Denmark)	Administrative Register Data	
EE (Estonia)	Estonian Education Information System	EHIS
	Eksamite Infosüsteemi (Examinations Information System)	EIS
EL/GR (Greece)	Papadopoulou - West Macedonia	
	Athena Studies of Resilient Adaptation	AStRA
	National Diagnostic Tests	Hellenic PISA
	MySchool	MySchool
ES (Spain)	Evaluaciones Diagnósticas (Gobierno Vasco)	ISEI-IVEI
	Education Data at the Catalan Department of Education	RALC
	Evaluaciones de Diagnóstico - The Canarian Agency for Quality Assessment and Accreditation, Santa Cruz de Tenerife	ACCUEE
FI (Finland)	Adolescents' well-being and learning in future society	FRAM
	Bridging the Gaps	GAPS (2017-2021)
	INVEST data	INVEST
	Student engagement, truancy, and cynicism: A longitudinal study from primary school to upper secondary education	
	Dynamics between perceived social support and study engagement among primary school students: A three-year longitudinal survey	
	Mind the GAP (2013-2016)	GAPS (2013-2016)
	Finnish National Agency for Education's national central integration service for study rights and achievements	KOSKI
FR (France)	Direction de l'Évaluation, de la Prospective et de la Performance - Directorate of Evaluation, Forecasting, and Performance - Panel 2007 (entrants en sixième)	DEPP 2007
	Direction de l'Évaluation, de la Prospective et de la Performance - Directorate of Evaluation, Forecasting, and Performance - Panel 2011 (entrants CP)	DEPP 2011

	Étude Longitudinale Française depuis l'Enfance - French Longitudinal Study of Children	ELFE
	Système de Suivi des Candidats (Candidate Tracking System)	SYSCA
	Évaluations repères	Évaluations repères
HU (Hungary)	Országos Kompetenciamérés (National Assessment of Basic Competences)	NABC
	KRTK'S Linked administrative panel dataset - Admin3	KRTK'S Admin3
	KRTK'S Linked administrative panel dataset - Admin4	KRTK'S Admin4
IE (Ireland)	Children's School Lives	CSL
	Growing Up in Ireland, Infant Cohort	GUI (IC)
	Growing up in Ireland, Child cohort	GUI (CC)
IT (Italy)	Anagrafe Nazionale degli Studenti - Open data	ANS OPEN
	National Institute for the Evaluation of the Education System in Italy	INVALSI
	Anagrafe Nazionale Studenti microdata	ANS Microdata
LT (Lithuania)	NACIONALINIAI MOKINIŲ PASIEKIMŲ TYRIMAI (National Survey of Student Achievement)	NMPT (NSSA)
LU (Luxemburg)	Épreuves Standardisées (Standardized tests)	ÉpStan
LV (Latvia)	Centralised Exam Results	
NL (Netherlands)	The Netherlands Cohort Study on Education	NCO
	Children of Immigrants Longitudinal Survey in the Netherlands	CILSNL
	PRIMary to Secondary education (PRIMS), Cohort 1	PRIMS (C1)
	PRIMary to Secondary education (PRIMS), Cohort 2	PRIMS (C2)
NO (Norway)	The Behaviour Outlook Norwegian Developmental Study	BONDS
	Longitudinal Study on Intention to Leave Upper Secondary School	
	National Education Database	SSB/NUDB
PL (Poland)	Central Examination Board	CKE
PT (Portugal)	Observatório de Trajetos dos Estudantes do Ensino Secundário	OTES
	Portuguese Longitudinal Study on School Engagement	
	Provas de Aferição do Ensino Básico, Exames Nacionais do Ensino Básico e Exames Nacionais do Ensino Secundário	PAEB, ENEB, ENES
	Estatísticas Oficiais de Educação e Formação	EEF
	Sistema de Informação do Ministério da Educação e Ciência	MISI
	Escola 360	E360
	Sistema Integrado de Informação e Gestão da Oferta Educativa e Formativa	SIGO
RO (Romania)	Training and Development in Education System	TRAIDES
SE (Sweden)	Evaluation Through Follow-up	UGU
	Gothenburg Educational Longitudinal Database	GOLD
UK (England, Wales, Scotland, Northern Ireland)	Millennium Cohort Study	MCS
	Longitudinal Survey of Young People in England (Cohort 2)	LYSPE2
	Understanding Society	Usoc
	Born in Bradford	BiB
	COVID Social Mobility and Opportunities	COSMO
	National Pupil Database	NPD
	Families and Schools Together	FAST

	Children's University	CU
	Growing Up in Scotland, Birth cohort 1	GUS (C1)
	Growing Up in Scotland, Birth cohort 2	GUS (C2)
	Growing Up in Scotland, Child cohort	GUS (CC)
USA	Early Childhood Longitudinal Studies	ECLS

II Part - Excluded Datasets

Country	Name	Acronym	Reason for exclusion
AT (Austria)	Atlas of Social Protection Indicators of Resilience and Equity	ASPIRE	Data collection not started yet
	Austrian-Socio-Economic Panel	ASEP	Data collection not started yet
BE (Flanders)	Studying Transactions in Adolescence: Testing Genes in Interaction with Environments	STRATEGIES	No student achievement/learning outcomes at individual level
	SONAR	—	Not recent
BE (Wallonia)	Educational Statistics	STAT_ELEVE	No available/clear information
BG (Bulgaria)	Data from the National Statistical Institute	NSI Data	No student achievement/learning outcomes at individual level
	Participation rate in education and training	—	No student achievement/learning outcomes at individual level
	Enrollment rates by educational degree	—	No student achievement/learning outcomes at individual level
	Pupils by education level and modern foreign language studied	—	No student achievement/learning outcomes at individual level
CH (Switzerland)	Swiss Education and Training Survey	SEATS	Comparative assessment
	Learning and Behavior	LABB	No student achievement/learning outcomes at individual level
	The Transition from Education to Employment - Study I	TREE I	Not recent
	Zürcher Longitudinal Study on Educational and Social Integration	ZLSE	Not recent
	Kleine Weltentdecker - Little World Explorers	—	No primary/ secondary education
	Bundesweite Nachfolgestudie zur Ausbildung von Kindern und Jugendlichen	BUNAVIA	No student achievement/learning outcomes at individual level
	Bildungsmonitoring Schweiz - Educational Monitoring Switzerland	VeCoF	Data collection not sufficient (not longitudinal)
CZ (Czechia)	Czech Longitudinal Study in Education	CLoSE (CLoSEread6)	Ethics concerns
DE (Germany)	Deutsch-Englisch-Schülerleistungen-International	DESI	Not recent
	Bildungsprozesse, Kompetenzentwicklung und	BiKS 8-14	Not recent

Selektionsentscheidungen im Vorschul- und Schulalter		
Professionalisierungsmaßnahmen zur bedeutungsfokussierten Sprachförderung im Sachunterricht der Grundschule	ProSach	Not representative
Tradition and Innovation: Developmental processes at non-academic track secondary schools in Baden-Wuerttemberg and Saxony	TRAIN	Not recent
Evaluation eines Vorschultrainings zur Prävention von Schriftspracherwerbsproblemen sowie Verlauf und Entwicklung des Schriftspracherwerbs in der Grundschule	EVES	Not recent
Digitale und datenbezogene Kompetenzen in Deutschland	DataLiteracy	Data collection not started yet
Kompetenzen und Einstellungen von Schülerinnen und Schülern - Skills and Attitudes of School Children	KESS	Not recent
Erhebung zum Lese- und Mathematikverständnis: Entwicklungen in den Jahrgangsstufen 4 bis 6 in Berlin	ELEMENT STUDY	Not recent
Aspects of Learning Background and Learning Development	LAU	Not recent
Duzy, 2013	-	Not representative
Trebits et al., 2022	-	Not representative
Stereotype Threat als Ursache niedriger Leistungen von Schülerinnen und Schülern mit Migrationshintergrund im deutschen Bildungssystem	Stereotype Threat	Not representative
Fischer & Rustemeyer, 2007	-	Not recent
FRéDA - The German Family Demography Panel Study	FRéDA	No student achievement/learning outcomes at individual level
Kompetenzaufbau und Laufbahnen im Schulsystem	KOALA-S	No student achievement/learning outcomes at individual level
Bielefelder Längsschnittstudie zum Lernen in inklusiven und exklusiven Förderarrangements	BiLieF	Not representative
Berufliche Zielorientierung und Kompetenzerwerb bei Lehrkräften	-	Not recent
Nett et al. (2022)	-	Not representative
Study on the development of all-day schools	StEG	Not recent
Selbstorientierung und Selbstständiges Lernen: eine	SEBI1	Not representative

	Analyse von Sozialisations- und Lernumgebungen von Grundschulkindern		
	Selbstprozesse und Schülerkarrieren: Ein interaktionistischer Zugang zur strukturellen Reproduktion	SEBI2	Not longitudinal
DK (Denmark)	Børneforløbsundersøgelsen - Danish Longitudinal Survey of Children	BFU	Same data described
EE (Estonia)	Estonian Open Data	–	No student achievement/learning outcomes at individual level
	Riigi Infosüsteemi Haldussüsteem - Estonian Catalogue of Public Sector Information	RIHA	No student achievement/learning outcomes at individual level
	Statistics Estonia	–	No student achievement/learning outcomes at individual level
ES (Spain)	Pruebas Diagnósticas de la Comunidad de Madrid	-	Not a dataset
	Prueba de Conocimientos y Destrezas Indispensables	CDI	Not a dataset
	Instituto Nacional para la Evaluación de la Educación	INEE	Comparative assessment
	Longitudinal data Cambra de Comer	-	Not longitudinal
	Merino et al 2020	ISCY	Not longitudinal, not representative
	Evaluación General de Diagnóstico	EGD	Not longitudinal, not recent
	Plan de Refuerzo, Orientación y Apoyo	PROA	Not longitudinal, not recent
	Mercader et al., 2017	-	Not representative
	Marchesi et al., 2004	-	Not recent
	Panel de Familia Infancia	PFI	Not recent
FI (Finland)	Higher education achievement register	VIRTA	No primary/secondary education
	Are Schools Alienating Digitally Engaged Students? Longitudinal Relations between Digital Engagement and School Engagement	Hietajärvi et al., 2020	No available/clear information
	Student engagement profiles in vocational education and training: a longitudinal study	Niittylahti et al., 2023	Not representative
	Staying on track of learning	-	Not recent
	Assessment of Undergraduate Students' Learning Outcomes in Finland	KAPPAS2!	Not longitudinal, No primary/secondary education
	METROP	-	No available/clear information
	2-y preschool RCT	-	No primary/secondary education

FR (France)	Évaluation des acquis des élèves - Assessment of Students' Achievements	Insee-Depp-Sies EVA	No primary/secondary education
	ORTHOLEARN project	—	Not representative
	Trajectoires et Origines - Trajectories and Origins	TeO1	Not recent
	Scolarité file of the Bordeaux academy	—	Not recent
HU (Hungary)	Hungarian Life Course Survey	HLCS	Not recent
IC (Iceland)	Reykjavik Risk-Taking Longitudinal Study	—	No student achievement/learning outcomes at individual level
	How self-regulation influences academic achievement and positive youth development: the mediating role of internalizing and externalizing symptoms	Guðmundsdóttir, et al. (2023)	No available/clear information
IE (Ireland)	Living in Ireland	LII	Not recent
	Post Primary Online Dataset	PPOD	No student achievement/learning outcomes at individual level
IT (Italy)	Di Tommaso et al., 2024	—	Experimental design
	Survey on students with BES	BES Survey	Not representative
	INAPP (National Institute for the Analysis of Public Policies) plus	INAPP Plus	No student achievement/learning outcomes at individual level, No primary/ secondary education
	Labor Force Survey	ISTAT ILS	No primary/secondary education
	Health Behavior in School-aged Children	HBSC	No student achievement/learning outcomes at individual level, comparative assessment
	Valutazione della Qualità - Assessment of Quality	AVQ	No student achievement/learning outcomes at individual level
	Administrative data on school buildings	—	No student achievement/learning outcomes at individual level
	Annual survey on schools	—	No student achievement/learning outcomes at individual level
	Indagine Diplomati - Graduates Survey	ISTAT Graduates Survey	No primary/secondary education
	Registro Elettronico - Electronic Register	—	Not a dataset
	Gui, Ercolanoni, Assirelli (2024)	—	Not longitudinal, not representative
LI (Liechtenstein)	Schulstatistik	—	No student achievement/learning outcomes at individual level
LT (Lithuania)	Education Management Information System	EMIS	No student achievement/learning outcomes at individual level
	Open Information, Counselling and Guidance System	AIKOS	No student achievement/learning outcomes at individual level
	Students' Register	—	No primary/secondary education
	Open Data Portal	—	Same data described
	Official Statistics Portal	—	No student achievement/learning outcomes at individual level

LU (Luxemburg)	Young people and COVID19	YAC	No student achievement/learning outcomes at individual level
	Youth Survey Luxembourg	YSL	No student achievement/learning outcomes at individual level
	ChildWeb	–	No student achievement/learning outcomes at individual level
	SCOLARIA	–	Not a dataset
LV (Latvia)	Latvian Qualifications Database	LQD	No student achievement/learning outcomes at individual level
	Academic Information Center	AIC	No student achievement/learning outcomes at individual level
	Latvia's Open Data	–	No student achievement/learning outcomes at individual level
	Graduates from Primary and Secondary Education	–	No student achievement/learning outcomes at individual level
	Further education of primary and secondary school graduates	–	No student achievement/learning outcomes at individual level
NL (Netherlands)	Cohortonderzoek Primair Onderwijs - Primary Education Cohort Study	PRIMA Cohort Study	Not recent
	OnderwijsLoopbanen onder leerlingen van 5 tot 18 jaar - Cohort research on educational careers among students aged 5-18	COOL	Not recent
	Poorthuis et al., 2015	–	Not recent
NO (Norway)	Ung I Norge - Young in Norway	Ung I Norge	Not recent
	Mother and Child Cohort Study (SOL project)	MoBa - SOL project	No student achievement/learning outcomes at individual level
	UngData	UngData	No student achievement/learning outcomes at individual level
	(Young) Hunt	Hunt	No student achievement/learning outcomes at individual level
	Eriksen et al. 2023	–	No student achievement/learning outcomes at individual level
PL (Poland)	System Informacji Oświatowej - Education Information System	SIO	No student achievement/learning outcomes at individual level
PT (Portugal)	Rosário et al., 2017	-	Not representative
	Careers on Doctorate Holders	CDH	No primary/secondary education
	Ribeiro et al., 2023	-	Not representative
	Lemos et al., 2020	-	Not representative
	Census data – Statistics Portugal	-	No student achievement/learning outcomes at individual level
	Labour force survey (Statistics Portugal)	-	No student achievement/learning outcomes at individual level
	Registo dos Alunos e Diplomados do Ensino Superior	RAIDES	No primary/secondary education
	Concurso Nacional de Acesso ao Ensino Superior	CNA	No primary/secondary education
RO (Romania)	PERSEIDA	PERSEIDA	No available/clear information

SE (Sweden)	Swedish Twin Registry - Academic Research	STAR	Not a dataset, ethnic concerns
UK (England, Wales, Scotland, Northern Ireland)	Success For All	EFF	Experimental design
	School Matters	—	Not recent
	Derrington, 2007	—	Not recent, not representative
	The Avion Longitudinal Study of Parents and Children	ALSPAC	Not recent
	The science aspirations and career choice project	ASPIRES	No student achievement/learning outcomes at individual level
	Junior School Project	JSP	Not recent
	British Cohort Study 1970	BCS	Not recent
	Next Steps	LSYPE1	Not recent
USA	School and District Assessment data	SEDA	No student achievement/learning outcomes at individual level
	High School Longitudinal Study of 2009	HSLS:09	Not recent
	Education Longitudinal Study of 2002	ELS:2002	Not recent
	National Assessment of Educational Progress	NAEP	Not longitudinal

9.2 Structure of the Expert Questionnaire

INTRODUCTION (opening page)

Welcome to this questionnaire designed for experts on education!

We are a consortium of universities and research institutions, and we are part of the [LINEup EU Horizon Europe project](#).

We aim at identifying, mapping and analysing **existing longitudinal data on students' learning outcomes at primary and (upper/lower) secondary education**. The objective is to study how educational inequalities emerge and develop over time, by focusing on educational achievement and attainment of basic skills and students' school engagement.

We focus on **school education**, and we are particularly interested in discovering **sources of national/regional/local longitudinal data** that are little known and used, with the final scope of analysing and, when possible, harmonizing them with other longitudinal datasets.

If you are here, you have been identified as an **expert on educational science and research**, and we believe your knowledge and contribution is valuable.

The questionnaire is short and consists of **3 sections**:

- In the first one, we ask you to suggest a national/regional/local longitudinal dataset at primary and/or secondary level;
- In the second section, we ask you to provide the name of another expert we can contact;
- In the final section, we ask some questions about your research.

This will last approximately **15 minutes**. Please notice that all collected information will be protected and analysed in an aggregated way, in compliance with privacy policies.

Thank you very much for your valuable contribution!

The LINEup Team

☐ Button for privacy policy to accept

SECTION 1 – Suggest a longitudinal dataset to be mapped by LINEup

In this first section, we ask you to suggest and partially describe **one or more longitudinal dataset(s) not already known by our research team and available in the country where you mainly work.**

Here the **selection criteria** of the datasets we are looking for:

- Data should be **longitudinal** - e.g. purely longitudinal survey data, pseudo-longitudinal (repeated cross-sectional) survey data, administrative data.
- Data should be about **primary and secondary level of education** (no higher education).
- Data should be on **students' learning outcomes (basic skills and others)** typically measured through standardized test scores; and/or on **academic performance indicators** (e.g. school grades, drop-out, truancy, grade retention, school trajectories); and/or on other measures of **cognitive and non- cognitive development** (e.g. socio-emotional and transversal skills, students' school engagement, sustained attention, etc.)
- Data should come from **EU Member States (27), European Economic Area Associated Countries** (Iceland, Liechtenstein, and Norway) and/or the **UK**.
- Data should be **representative at regional/national level**, or at least based on a large-scale sample.
- The data collection should be **ongoing or recent** (if completed, the latest round should have taken place in the last decade, year 2014 the latest).

Examples of such datasets are: NEPS (DE), INVALSI (IT), ELFE (FR), the Millennium Cohort Study (UK), etc. Do not worry, during the questionnaire, we provide information about the datasets we already identified to save your time and avoid collecting double information.

Please note that international comparative assessments are out of scope for this phase of the research (e.g. PISA, SESS, TIMMS, PIRLS, ICLS, etc.)

As explained, we want to map all existing datasets with these characteristics listed above: feel free to suggest little known or unused datasets, even if they are **not open access**.

Thank you for your collaboration!

SECTION 1	Questions
Q1* In which country is the institution you mainly work with/for? (one option possible)	1. AT (Austria) 2. BE (Wallonia) 3. BE (Flanders) 4. BG (Bulgaria) 5. CY (Cyprus) 6. CZ (Czechia) 7. DE (Germany) 8. DK (Denmark) 9. EE (Estonia) 10. EL/GR (Greece) 11. ES (Spain) 12. FI (Finland) 13. FR (France) 14. HR (Croatia) 15. HU (Hungary) 16. IC/IS (Iceland) 17. IE (Ireland) 18. IT (Italy) 19. LI (Liechtenstein)

	20. LT (Lithuania) 21. LU (Luxemburg) 22. LV (Latvia) 23. MT (Malta) 24. NL (Netherlands) 25. NO (Norway) 26. PL (Poland) 27. PT (Portugal) 28. RO (Romania) 29. SE (Sweden) 30. SI (Slovenia) 31. SK (Slovakia) 32. UK (England, Wales, Scotland, Northern Ireland) 33. Other (specify): _____
Q1.2* For this country, we identified the following dataset(s).	<i>List of dataset(s) appears (if we did not identify any dataset for some countries, Q1.2 and Q1.2.1 do not appear).</i>
Q1.2.1* Can you please tell us what is your opinion about the dataset(s) listed above (for instance strengths and weaknesses you consider important, major problems you saw or found in the dataset(s), personal considerations you want to share with us)? <i>Suggestion: If you do not know, you can type "Not applicable". If you decide to tell us your opinion, please at the beginning of the text state the name of the dataset(s) you refer to. (open answer)</i>	About the dataset: _____ _____
Q2* As far as you know, how many longitudinal datasets are present in the country where your institution is based (the one you selected) BESIDES the ones already listed above? Please, keep in mind the <u>selection criteria</u> mentioned. <i>Suggestion: If the list of datasets we identified for your country does not appear, it means we did not identify any longitudinal dataset for your country. (one option possible)</i>	1. None 2. One dataset 3. More than one dataset 4. I don't know
Filter: If 1. None or 4. I don't know à SECTION 2 If 2. One à Q2.3 and then Q4 and following. If 3. More than one à Q2.1 & Q2.2 and then Q4 and following & Q2.3, Q10	

Q2.1* (filter) Please, write down the name of the datasets you are aware of/refer to in the previous question. (fill in from 2 to 3 answers)	Dataset 1: _____ Dataset 2: _____ Dataset 3: _____
Q2.2* (filter) Indicate the dataset name of the one you want to describe. (open answer)	Dataset: _____
Q2.3* (filter) Indicate the dataset name (open answer)	Dataset: _____
Q3* Indicate the type of dataset (multiple answer possible)	1. Longitudinal survey data 2. Pseudo-longitudinal (repeated cross-sectional) survey data 3. Administrative longitudinal data
Q4* Indicate the official website of the dataset, if available Suggestion: <i>if it is not available, you can type "Not available"</i> (open answer)	Website: _____
Q5* Indicate the level of the sample (one option possible)	1. National 2. Limited only to some country's region (specify): _____ 3. I don't know
Q6* Indicate the school grades covered (Multiple answer possible)	1. Infant school or kindergarten 2. First grade 3. Second grade 4. Third grade 5. Fourth grade 6. Fifth grade 7. Sixth grade 8. Seventh grade 9. Eighth grade 10. Ninth grade 11. Tenth grade 12. Eleventh grade 13. Twelfth grade 14. Thirteenth grade, when present 15. Other (specify): _____ 16. I do not know

Q7* Indicate whether is it possible to access micro-data (one option possible)	1. Open data 2. Full access through requests 3. Limited access through requests 4. I do not know
Filter: If Q7 à 3 (limited access) à Q7.1 & Q8 If Q7 à 1, 2, or 4 à Q8	
Q7.1* (filter) Can you please describe the main limitation in accessing the dataset? (multiple answer possible)	1. Only by paying 2. Only accessible to employees of a specific organization 3. Only accessible in a specific place 4. Other (specify): _____
Q8* Can you please tell us what is your opinion about the dataset you are suggesting (for instance strengths and weaknesses you consider important, major problems you saw or found in the dataset(s), personal considerations you want to share with us)? (open answer)	About the dataset: _____ _____
Q2.4* (filter) Would you like to suggest another dataset for the country you selected as the main one you work in? (one option possible)	1. Yes 2. No
Filter for Q2.4 from Q2: If 1: Yes à Go to Q2.3 (restart) If 2. No à SECTION 2 If Q2.1 fills 3 datasets, then it is possible to describe up to 3 datasets	

SECTION 2 – Your contribution to the LINEup Project

In the second section, we ask you to suggest the name of another expert in education like yourself, and to contribute to our project and to our literature review with your expertise.

SECTION 2	Questions
Q9* Do you want to suggest another expert to be contacted for answering this questionnaire? <i>(one option possible)</i>	1. Yes 2. No
Filter: If 1. Yes → Q9.1 If 2. No → Q11	
Q9.1* (filter) Please, insert below the contact information for the expert you suggested. <i>(open answer)</i>	Expert Name: _____ Expert Surname: _____ Expert Organisation: _____ Expert E-mail: _____
Q10* We are performing a literature review on empirical research through longitudinal datasets on the topic of the project. Would you be willing to contribute to such literature review by indicating publications in this respect? <i>(one option possible)</i>	1. Yes 2. No
Filter: If 1. Yes → Q10.1 If 2. No → SECTION 3	
Q10.1* (filter) Please indicate up to three relevant publications written by you or others based on analysis of longitudinal data in primary and secondary education. Feel free to consider studies in your own language, hence not easily retrievable in international databases. Suggestion: <i>Please state at least: Surname of the author(s), title, and year of the publication.</i> <i>(fill in from 1 to 3 answers)</i>	Publication 1: _____ Publication 2: _____ Publication 3: _____

SECTION 3 – About you

In this final section, we ask some information about you. Please remember that all collected information will be protected and analysed in an aggregated way, in compliance with privacy policies. If you decide to provide your contact details, you will be able to receive updates about the LINEup Project.

SECTION 3	Questions
Q11* Would you like to receive updates about the LINEup project? <i>(one option possible)</i>	1. Yes 2. No
Filter: If 1: yes → Q11.1 & Q11.2 If 2. No → Q12	
Q11.1* (filter) What type of updates would you like to receive? <i>(multiple answers)</i>	1. General newsletter (did you know? – every 6 months) 2. Scientific newsletter (lessons learned – every 6 months) 3. Policy briefs (LINEup results – mid and end of project)
Q11.2* (filter) Please add your email so you can receive updates about the LINEup project. <i>(open answer)</i>	E-mail: _____
Q12 To allow us not to contact you again with this request, please insert your name, surname, and institution. <i>(open optional answer)</i>	Name: _____ Surname: _____ Main institution: _____
Q13 Use this space if there is anything you want to add. <i>(open optional answer)</i>	_____ _____ _____

END OF QUESTIONNAIRE (ending page)

Thank you for participating and contributing to the LINEup project with your expertise. Please notice that the dataset(s) you suggested will be further analysed and there is no responsibility for possible mistakes or inaccurate information. It is possible for you to add or modify information, contacting us directly here: ricerca@fondazioneeduca.it

9.3 Structure of Dataset Description Form

Introduction

This is the form designed for LINEup internal members for describing the dataset(s) identified in each of the countries we need to cover according to the following distribution:

Partners	Countries to be covered
PHKA	AT, BG, DE, LI + UK in cooperation with UPF
KEPE	CY, DK, EL, NL
PIREUS	HU
UPF	ES, IE, MT, U.S. + UK in cooperation with PHKA
INED	BE (WALLONIA), BE (FLANDERS), FR, LU
UTAD	EE, LT, LV, PT
FpS	IT, PL, NO, HR
UNIMI/BICOCCA	SE, CZ, SK, CH
UNIMC	FI, IC, RO, SI

In Yellow = additional countries, on top of the ones we promised to cover in the project application.

In this form, you are expected to describe a dataset that complies with the following criteria:

- Data should be **longitudinal** - e.g. purely longitudinal survey data, pseudo-longitudinal (repeated cross-sectional) survey data, administrative data.
- Data should be about **primary and secondary level of education** (no higher education).
- Data should be on **students' learning outcomes (basic skills and others)** typically measured through standardized test scores and/or on **academic performance indicators** (e.g. school grades, drop-out, truancy, grade retention, school trajectories).
- Data should be **representative at regional/national level**, or at least based on a large-scale sample.
- The data collection should be **ongoing or recent** (if completed, the latest round should have taken place in the last decade, year 2014 the latest).

Please note that international comparative assessments are out of scope (e.g. PISA).

The form is made of the following six parts:

1. Basic information on the identified datasets.
2. Basic information on the skills investigated.
3. Additional information available in the dataset: students' characteristics.
4. Additional information available in the dataset: students' household characteristics.
5. Additional information available in the dataset: teachers' characteristics.
6. Additional information available in the dataset: school characteristics.

Please:

- Try to reply to all questions and limit the use of the answer option 'I do not know/not clear'.
- Use the free text part you find at the end of each section of the form to add details about the peculiarity of the dataset.
- Be aware that you cannot change the form once it is submitted – hence we strongly advise to have a first draft on paper before filling the online form in.

Section 1: Basic information on the identified datasets

SECTION 1	Text
Dataset country <i>(One choice)</i>	1. AT (Austria) 2. BE (Wallonia) 3. BE (Flanders) 4. BG (Bulgaria) 5. CY (Cyprus) 6. CZ (Czechia) 7. DE (Germany) 8. DK (Denmark) 9. EE (Estonia) 10. EL/GR (Greece) 11. ES (Spain) 12. FI (Finland) 13. FR (France) 14. HR (Croatia) 15. HU (Hungary) 16. IC/IS (Iceland) 17. IE (Ireland) 18. IT (Italy) 19. LI (Liechtenstein) 20. LT (Lithuania) 21. LU (Luxemburg) 22. LV (Latvia) 23. MT (Malta) 24. NL (Netherlands) 25. NO (Norway) 26. PL (Poland) 27. PT (Portugal) 28. RO (Romania) 29. SE (Sweden) 30. SI (Slovenia) 31. SK (Slovakia) 32. UK (England, Wales, Scotland, Northern Ireland) 33. Other (specify): _____
Dataset name	Name: _____
Data owner/collector/producer <i>(multiple answers)</i>	1. Public authority (ministry or agencies) 2. University or public research center 3. Private organisation 4. Other (specify: _____)
Purpose of data collection <i>(multiple answers)</i>	1. Academic research 2. School system monitoring/evaluation 3. Educational institutions monitoring/evaluation 4. Individuals' assessment (without implications on their school career or grading – low stake) 5. Individuals' assessment (with implications on their school career or grading – high stake) 6. Other (Specify: _____)
Official website (if available)	Website: _____
Was the data collection initially designed as longitudinal at individual level?	1. No, but it is possible to develop a pseudo-panel thanks to repeated cross-sectional waves 2. No, but it is possible to follow students over time thanks to an individual student's identifier/identification code

	3. Yes, the test design is longitudinal, following students for n° years: _____
Level of the sample	4. National 5. Limited only to some country's region (specify): _____
Type of the sample & Sample size (multiple answers)	1. Population (all students) 2. Random sample of students (please comment in the box the average number of students in the sample across waves: sample size: n° _____) 3. Non random sample of students but trying to achieve national/regional representativeness 4. Other - but relevant because: _____
School tracks	1. All school tracks available in the education system are included in the sample 2. Only some school tracks available in the education system are included (specify): _____
Sampling levels (multiple answers)	1. Counties/Cities 2. Schools 3. Classes 4. Pupils 5. Other (specify: _____)
Frequency of data collection – also in case of cross-sectional assessments	1. Yearly (or more than once per year) 2. Every other year 3. Every three years 4. Every four years or more
Starting year of administration of the test	Year: _____
Ending year of administration of the test (if ongoing, type “ongoing”)	Year: _____
School grades included in the design (multiple answers)	1. First grade 2. Second grade 3. Third grade 4. Fourth grade 5. Fifth grade 6. Sixth grade 7. Seventh grade 8. Eighth grade 9. Ninth grade 10. Tenth grade 11. Eleventh grade 12. Twelfth grade 13. Thirteenth grade, when present
Early childhood education or pre-primary education included in the longitudinal design for primary and secondary education	1. Yes 2. No

Linkability of the data at individual level	1. Data include an identification code that allows linkability/matching to other (administrative) data to all researchers (at least through ad hoc requests) 2. Data include an identification code that allows linkability to other (administrative) data only for data-owners 3. Data do not include an identification code that allows linkability to other (administrative) data
Data collection – only if linkable at student's level <i>(multiple answers)</i>	1. Students' questionnaire 2. Students' test 3. School questionnaire (i.e. structural information or available resources) 4. Principal/Headmaster questionnaire 5. Teachers' questionnaire 6. Parents' questionnaire 7. Others (Specify):
Data accessibility	1. On a personal computer without restriction 2. On a personal computer with a specific interface 3. On a computer belonging to the producer of the data, which is not connected to internet or to another computer (in this case the producer examines all the results output) 4. Other (specify: _____)
Access to micro-data <i>(multiple answers)</i>	1. Open data 2. Full access through requests 3. Limited access through requests – main limitations: _____
Additional notes on the dataset (strengths, weaknesses, personal considerations, etc)	

Section 2: Basic information on the skills investigated

SECTION 2	Text
Type of skills analysed in the test (specify sub content areas if available) <i>(multiple answers)</i>	1. Literacy (reading/writing, etc.) (specify): _____ 2. Numeracy (mathematics) (specify): _____ 3. Science (scientific literacy) (specify): _____ 4. Foreign language (specify): _____ 5. Other skills (specify): _____
Type of scale/measurement of skills/achievement used <i>(multiple answers)</i>	1. Continuous score 2. Proficiency levels 3. Others (specify): _____
Type of assessment	1. Computer-based 2. Paper-based 3. Both (specify criteria, e.g. year of change, different for specific grades): _____
Additional notes	

Section 3: Additional information available: students' characteristics

SECTION 3	Text
Gender	1. Yes 2. No 3. Not clear
Age	1. Yes 2. No 3. Not clear
Citizenship	1. Yes 2. No 3. Not clear
Foreign birth country	1. Yes 2. No 3. Not clear
Specific birth country	1. Yes 2. No 3. Not clear
Town of residence	1. Yes 2. No 3. Not clear
Province of residence	1. Yes 2. No 3. Not clear
Region of residence	1. Yes 2. No 3. Not clear
Belonging to a recognised ethnic minority	1. Yes 2. No 3. Not clear
Attendance to early childhood education and care	1. Yes 2. No 3. Not clear
Previous grade retention	1. Yes 2. No 3. Not clear
Learning impairments	1. Yes 2. No 3. Not clear
Physical impairments	1. Yes 2. No 3. Not clear
Attitude or motivation toward education/school	1. Yes 2. No 3. Not clear
School grades	1. Yes 2. No 3. Not clear
Presence of student allowance/scholarship	1. Yes 2. No 3. Not clear
The information on students flagged above are:	1. Self-reported (provided by the student)

	2. Administrative (provided by the institution or others) 3. Both self-reported and administrative 4. Not clear
Other information on students (specify):	
Additional notes:	

Section 4: Additional information available: student's household characteristics

SECTION 4	Text
Number of parents in the household	1. Yes 2. No 3. Not clear
Presence of stepparents	1. Yes 2. No 3. Not clear
Siblings	1. Yes, number 2. Yes, but only presence/absence 3. No 4. Not clear
Parental working status (i.e. employed, unemployed...)	1. Yes, for both parents 2. Yes, but for only one parent 3. No 4. Not clear
Parental occupation	1. Yes, detailed information (e.g. ISCO code) for both parents 2. Yes, detailed information for only one parent 3. Yes, simplified information for both parents 4. Yes, simplified information for only one parent 5. No information 6. Not clear
Parental education	1. Yes, for both parents 2. Yes, but only for one parent 3. No 4. Not clear
Parental education level	1. Yes, full ISCED for both parents 2. Yes, full ISCED for only one parent 3. Yes, simplified ISCED for both parents 4. Yes, simplified ISCED for only one parent 5. No 6. Not clear
Parental migration background	1. Yes, for both parents 2. Yes, but only for one parent 3. No

	4. Not clear
Parents age	1. Yes 2. No 3. Not clear
Parents place of birth (if migrants)	1. Yes, specific country 2. Yes, geographical area 3. No 4. Not clear
Parental income or wealth	1. Yes 2. No 3. Not clear
Parental level of knowledge of host country's language	1. Yes 2. No 3. Not clear
N of books in the household	1. Yes 2. No 3. Not clear
N of digital devices in the household	1. Yes 2. No 3. Not clear
Ownership of the apartment/house they live in	1. Yes 2. No 3. Not clear
The information on student household flagged above are:	1. Self-reported (provided by the student) 2. Administrative (provided by the institution or others) 3. Both self-reported and administrative 4. Not clear
Other information on parents or household (specify):	
Additional notes	

Section 5: Additional information available: teachers' characteristics

SECTION 5	Text
Teacher age	1. Yes 2. No 3. Not clear
Teacher gender	1. Yes 2. No 3. Not clear
Teacher seniority	1. Yes 2. No 3. Not clear
Teacher educational degree	1. Yes 2. No 3. Not clear
Teacher type of contract (i.e. permanent, fixed term)	1. Yes

	2. No 3. Not clear
The information on teachers flagged above are:	1. Self-reported (provided by teachers) 2. Administrative (provided by the institution or others) 3. Not clear
Linkability: the information on teachers flagged above are:	1. Linkable to the ones of students 2. Not linkable to the ones of students 3. Not clear
Other information on teachers (specify):	
Additional notes:	

Section 6: Additional information available: school characteristics

SECTION 6	Text
Geo-referencing of the school	1. Yes 2. No 3. Not clear
Type of school (i.e. public, private)	1. Yes 2. No 3. Not clear
Track followed (when relevant)	1. Yes 2. No 3. Not clear
School size	1. Yes 2. No 3. Not clear
Classroom size	1. Yes 2. No 3. Not clear
School composition (i.e. migration/socio-economic background)	1. Yes 2. No 3. Not clear
Class composition (i.e. migration/socio-economic background)	1. Yes 2. No 3. Not clear
Other information on the school or class (specify):	
Additional notes:	

9.4 Glossary of Technical Terms

This glossary provides a quick and concise reference for key technical terms used throughout the Mapping Research Report. For full conceptual elaboration, see section 2.

Term	Definition
Academic performance	A subset of learning outcomes that includes grades, grade retention, truancy, or school progression.
Access conditions	The terms under which microdata can be consulted or used for research purposes, including restrictions such as registration, data agreements, or on-site use.
Administrative data	Data collected by public authorities or institutions as part of routine operations, often linked to official registers.
Cohort	A group of individuals who share a common characteristic (e.g. year of birth or school entry) and are followed over time in longitudinal research.
Data harmonisation	The process of aligning data from different sources to ensure comparability across datasets and countries.
Database	A broader system or infrastructure that may include multiple datasets collected under the same programme or initiative, potentially covering different cohorts, regions, or time periods.
Dataset	A single, structured collection of data that shares the same design and unit of analysis. In the context of this report, each dataset is considered a distinct unit for mapping and analysis.
Data accessibility	The degree to which data can be accessed, used, and reused. This may include conditions such as open access, restricted access, or institutional permission.
Data collection round or wave	A single point in time at which data are collected from the sample in a longitudinal dataset. Longitudinal datasets include multiple waves.
Data source	The origin or institutional provider of a dataset, such as a ministry, statistical agency, or research institute. Distinct from the dataset itself, which refers to the specific structured data collected.
Data linkage	The process of connecting information from different datasets or administrative sources about the same individuals, often through anonymised identifiers.
Educational inequality	Differences in educational opportunities or outcomes that are systematically associated with social or demographic characteristics.
Hybrid data	A dataset combining features of both survey-based and administrative sources or combining different types of longitudinal formats (e.g. panel and pseudo-panel).
Inclusion criteria	The conditions defined to determine which datasets could be included in the mapping activity, such as relevance to primary/secondary education, longitudinal structure, and data availability on learning outcomes.
Learning outcomes	Measures of what students know and can do, typically assessed through tests, grades, or various performance indicators.
Longitudinal data	Data collected from the same subjects repeatedly over time to study changes and developments.

Metadata	Structured information describing the content, structure, and access conditions of datasets. In this report, metadata enables the development of the online Data Explorer.
Metadata file	A structured document or database that compiles standardised descriptions of multiple datasets, including variables, methods, and access conditions. In LINEup, the metadata file underpins the LINEup Education Data Explorer.
Microdata	Individual-level data that provide detailed information for each observation (e.g., student), as opposed to aggregated statistics. Microdata is often subject to access restrictions due to privacy concerns.
Panel studies (or pure longitudinal data)	A type of longitudinal dataset that tracks the same individuals across multiple waves.
Pseudo-panel	A dataset constructed by grouping individuals into cohorts and tracking these groups over time using repeated cross-sections.
Repeated cross-section	A dataset that collects data from different individuals in the same population at multiple time points.
Sampling weights	Values used to adjust for unequal probabilities of selection or non-response in a sample.
Standardised test	An assessment administered and scored in a consistent way, often used to compare performance across time (anchoring) or populations.
Unit of analysis	The primary subject or entity that data are collected about in a dataset (e.g., student, classroom, school). In this report, each dataset is treated as a separate unit of analysis for the mapping when it follows a coherent design and a consistent unit of analysis across waves (cohort).