

International Journal of Population Data Science

Journal Website: www.ijpds.org



The EEF Data Archive: Unlocking high-quality RCT data and resources for secondary analysis to advance education research.

Belén Parada Zuleta¹ and Rachael Morris¹

¹The Education Endowment Foundation, London, United Kingdom

Objectives

The Education Endowment Foundation (EEF) Data Archive provides researchers with a robust resource to produce research for the public good. The archive enables impactful secondary analyses by linking Randomised Controlled Trial (RCT) data with administrative datasets, such as the National Pupil Database (NPD), advancing evidence-based practices and policies in education.

Conclusions

Whilst trials are the gold standard for robust evidence generation, they are expensive, time-consuming, and logistically challenging. The EEF archive, linked with administrative datasets, enables long-term impact assessments, cross-intervention analyses, and methodological advancements, offering a valuable, time- and cost-effective opportunity to support research for the public good.

Methods

EEF is an independent charity dedicated to reducing the link between children's and young people's academic outcomes and their socio-economic background. Since 2011, the EEF has pioneered the commissioning of rigorous Randomised Controlled Trials (RCTs) in education and archived this high-quality data for future research. Currently, the EEF archive holds data from over 120 projects involving over 1.8 million children in England, a resource that will grow as new evaluations are conducted. With funding from the Evaluation Task Force (Cabinet Office/HM Treasury), EEF has recently created a portfolio of resources to facilitate secure third-party access to this data for secondary analysis.

Results

The EEF archive is hosted in the Office for National Statistics (ONS) Secure Research Service (SRS), which presents both opportunities and challenges for researchers wishing to access the data. This session will introduce the archive's extensive data resources and share insights from initial research projects leveraging its data. Key developments include piloting processes for linking EEF trial data with Department for Education datasets, such as the NPD. To enhance accessibility, the EEF has developed a suite of resources, including low-fidelity synthetic datasets, a feasibility study of creating higher fidelity synthetic data, a comprehensive data catalogue, and a streamlined application process. These efforts aim to facilitate secure, efficient access for researchers while maximising the archive's potential to support impactful education research.



Temporary page!

$\LaTeX$ was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because $\LaTeX$ now knows how many pages to expect for this document.