

Creating a unique world class resource of multi-country longitudinal data

July 2018



- Young Lives has created a dataset of exceptional quality, strength, depth and uniqueness which is of enormous value to all those with an interest in child poverty.
- Young Lives data and other resources are openly shared through public archives, and data visualisations are making the data even more accessible.
- The quality of the Young Lives study is reflected in the world-class researchers who seek to collaborate with us, and in the number of high-impact publications.
- Young Lives data have been used extensively by organisations including the World Bank and UNICEF/UN in analyses that have had, or will have, influence on global policy.

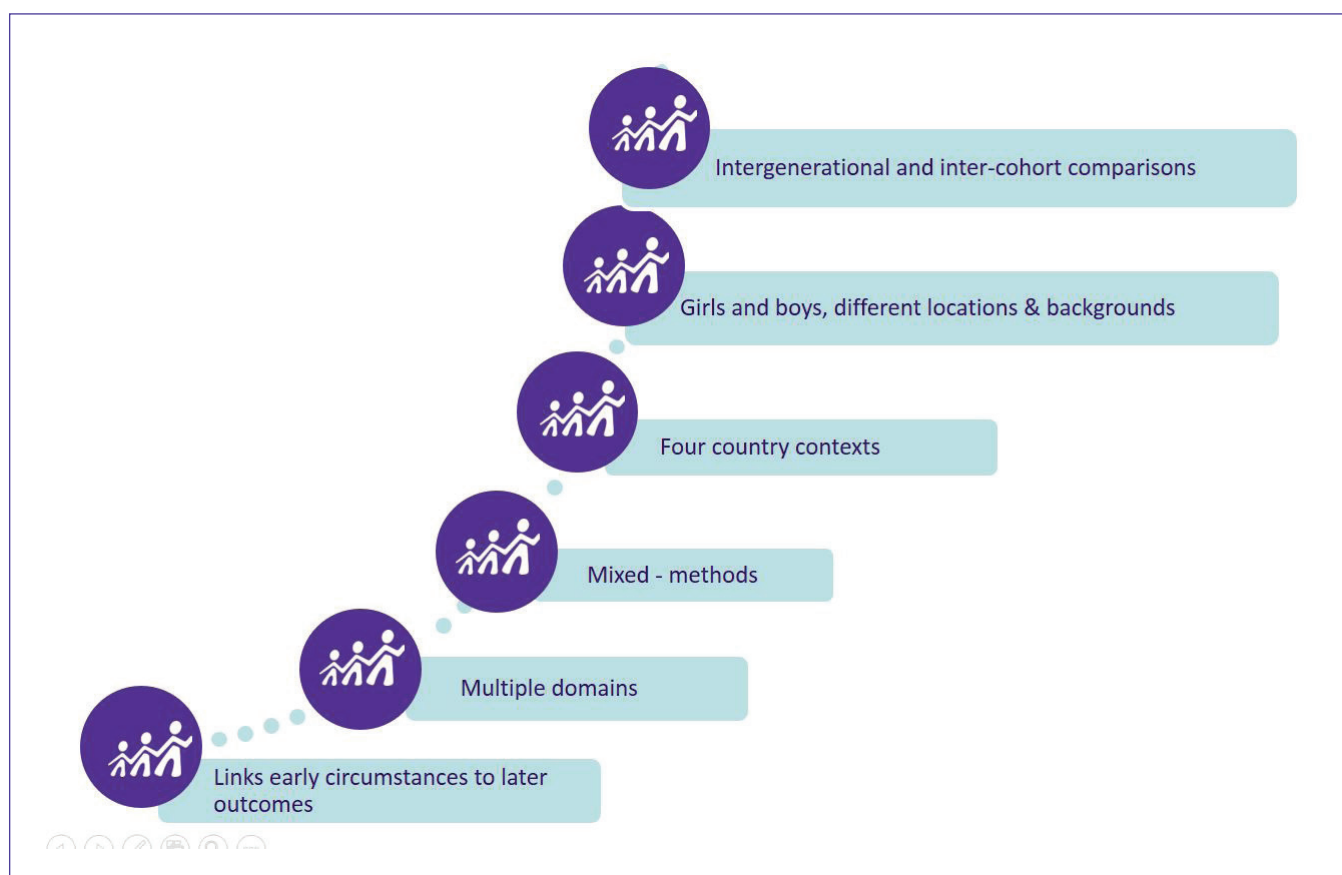
Setting the scene: the strengths and characteristics of Young Lives

One of the most significant impacts that Young Lives seeks to achieve is 'the legacy of a mature and unique longitudinal dataset, which can continue to be mined, and provides important social science infrastructure for research, teaching and learning in low- and middle-income countries, and indeed globally' ([Theory of Change, 2015](#)). Young Lives is one of the few studies worldwide that combines children's own perspectives on their lives with quantitative measures of wellbeing, time use and cognition, together with information on households and communities and the effectiveness of children's schools; this gives it a strength and depth which makes it of great value to many other individuals and organisations with an interest in child poverty. The public archiving of Young Lives data makes it accessible to the research community, and new online data visualisations tools widen access.

Longitudinal cohort data yield a wealth of insights that cannot be generated through shorter-term or single-focus research. In high-income countries cohort data have become established

as an essential part of informing change and improving policy and practice (for example the UK's [CLOSER](#) initiative). In low- and middle-income countries, however, far fewer studies are carried out, especially on children, making Young Lives a particularly important resource. The four country nature of the study gives it a relevance to a wider set of countries which may not have access to such data themselves.

The exceptional quality, strength and uniqueness of the Young Lives study can be evidenced in many ways (see also figure following). It is the only longitudinal cohort study comparable across countries, and the only one that uses diverse qualitative methods as well as quantitative measures. It gathers multi-dimensional information about communities, households, caregivers, schools, siblings and children's peers, as well as about the individual children themselves. This means that it is capable of identifying the determinants of children's outcomes at many different levels, and providing a holistic view of child development in contexts of risk and adversity. The data collected is multipurpose and is designed to allow insights and impacts in areas that were not necessarily the specific focus of the study (for example, the discovery that some recovery may be possible from the effects of early poor nutrition).



Young Lives' unique data features

Being a two-cohort study, it enables comparisons and measurements of improvement or deterioration to be made across the cohorts (such as educational performance at a given age). For more details on the Young Lives study design and methodology, please visit: <http://younglives.org.uk/content/our-research-methods>. In addition, Young Lives is beginning to collect data on a new generation of babies (the children of the Older Cohort); this is creating new intergenerational evidence to add to that already obtained through study of caregivers and Young Lives children. Another unique feature of Young Lives research is the surveys gathering data on school effectiveness, which are entering their second round; no other study in low- and middle-income countries brings together child, household and school data in this way.

World-class collaborators and high-impact publications

Researchers from across the world have actively sought to collaborate on the Young Lives study, reflecting the high esteem in which the study is held. The calibre of collaborators is extremely high, with respected academics from many different disciplines wanting to do joint research or use the data. Examples include Professors Paul Glewwe, Jere Behrman, Pauline Rose, Aryeh Stein, Orazio Attanasio and Robert Blum. The ability to attract and retain high-level collaborators significantly increases the reach and influence of Young Lives research.

“The Young Lives data, with rich comparable longitudinal information across four developing countries on individual children, their families and the communities in which they live, is unique ... The many strengths of the Young Lives data include a prospective design with longitudinal follow ups, extensive multi-dimensional data, rich community information that permit investigating causal relations, low attrition rates and the use of common instruments and research methods in Asia, Africa and Latin America.”

Professor Jere Behrman, University of Pennsylvania

Young Lives data has generated high-impact journal publications in, for example, *Social Science and Medicine*, *Journal of Nutrition*, *Public Health Nutrition*, *Economics and Human Biology*, *World Development*, *BMC Public Health*, *The Routledge Handbook of International Education and Development* and *International Journal of Behavioural Development*. In May 2016 the Lancet published a major new commission on Adolescent Health and Wellbeing, which contains analysis from three of the Young Lives countries. Public archiving of Young Lives data facilitates the publication of high-quality papers by external researchers and so this, coupled with the unique design features and efforts put into sharing the data, suggest this legacy will grow.

Open sharing of data

Young Lives' quantitative data is publicly archived via the UK Data Service (UKDS), and is internationally accessible. Public archiving is not common, especially in low- and middle-income countries. The dataset is vast, with a total of around 15 million data points collected in each survey round. There is little missing data, and few children drop out of the study, especially compared to studies in high-income countries, making this a high quality source. Young Lives also produces other resources including fieldworker manuals and technical notes, and panel data of key variables across all survey rounds. This additional information widens access and use of the data.

Users of the data has increased steadily. Of the users recorded by the principle repository (ESDS), one in five users came from outside the Organisation for Economic Co-operation and Development (OECD). Cumulative downloads of data have also risen sharply in recent years and by June 2018 stood at 6,080 (downloads are higher than users as Young Lives is made up of multiple datasets and users could access several datasets). The increase in users and downloads reflects our expectation that significantly more downloads would be made once additional rounds of data had increased the power of the longitudinal data. The fifth round is now available but downloads of this are not yet captured in the figures. We expect a substantial further increase in users. In addition to the public archive, Young Lives makes considerable efforts to help people actually use the data, especially in low- and middle-income countries. Data is released on CD-ROM to facilitate its use by those who do not have internet access. Young Lives country teams distribute the data directly to local researchers and also run training sessions on how to use the data and to increase data usage in the study countries. The Peru team is particularly active in this respect, having trained hundreds of students, academics, governmental and NGO staff in the use of Young Lives data in Peru.

In addition, Young Lives has developed [data visualisations](#) on the themes of Nutrition, Gender, Education, Skills and Work, and household wealth. The aim is to make fifteen years' worth of the complex and comprehensive data from five survey rounds even more accessible, interactive and powerful. The visualisations are designed to create narratives around what it is that impacts children in poverty, and prompt discussion, reflections and implications for what must be done to break intergenerational cycles of poverty. Generated charts can be downloaded. Multi-layered variables and determinants can be simultaneously visualised, creating a very useful tool for engaging with non-technical audiences as well as with policymakers. The team were awarded an OxTALENT award in 2018 in recognition of their innovative data visualisation platform aiming to make complex data accessible.

The World Bank's online Microdata Library has links to the Young Lives data archive, which facilitates access and also demonstrates that the World Bank regards the data as being of high quality and is keen to recommend to others.

Use of Young Lives data

Young Lives data is routinely used by major global institutions to provide evidence that is of national and international significance.

- In 2016, Young Lives Peru data was used in the World Bank report '[Left Behind: Chronic Poverty in Latin America and the Caribbean](#)' to provide insights into intergenerational poverty that would have been difficult to obtain through other non-longitudinal research (see quotation below). Peru data was also used in the CAF (Development Bank of Latin America) [RED Report 2016](#) on skills for work and life.
- Young Lives data from Vietnam was used extensively in the preparation of the OECD's Development Centre Social Cohesion Policy Review of Viet Nam (2014) and the World Bank's report [Skilling up Vietnam: preparing the workforce for a modern market economy](#) (2014).
- Other Young Lives data has been used by the World Bank in preparation of two major reports: the [World Development Report 2018: Learning to Realize Education's Promise](#), and in the preparation of a forthcoming report on early child development and skills.

“ One of the most difficult tasks we have as researchers in development is to understand welfare dynamics. Our simple cross-sectional data simply is not up for the task. Young Lives is one of those projects that has really facilitated breaking the data gaps. The unique ability to follow up children across different settings and over their childhoods has provided us with a rare glimpse of life trajectories and new insights about both the economic, social and psychological factors that go into play in the process of development and early childhood. One example that has struck me as extremely useful is from our work for our Chronic Poverty book in Latin America. The Young Lives data from Peru allowed us to better document the low intra-generational mobility not only in the education space but also in aspirations. That opens up a whole new agenda for us to seriously think how to break the intra-generational transmission of poverty. Looking forward to seeing more insights coming out of this project. ”

Renos Vakis, Lead Economist, Head - Behavioral Initiatives Poverty and Equity Global Practice & LSMS Team, World Bank

“ While doing research on inequality of opportunity at the World Bank, I had the chance of working with the Young Lives data. The possibility of having an understanding of the role of individual, family and community circumstances, and the panel features, provide opportunities to test hypotheses that are not possible in other designs. ... There are only few projects that have made the initial investment needed to build a real panel with the wealth of information that Young Lives provides. ”

Jaime Saavedra-Chanduvi, then Director of Poverty Reduction and Equity Department at the World Bank (and later Minister of Education in Peru)