



Image © Young Lives/ Niños del Milenio

A Guide to Young Lives Research

May 2017

Section 1: Young Lives – A Multi-disciplinary Longitudinal Study of Childhood Poverty

May 2017

Young Lives is an international study which is tracing the changing lives of 12,000 children over 15 years. It is following two groups of children, with approximately equal numbers of girls and boys, from poor households in Ethiopia, Peru, Vietnam, and the states of Andhra Pradesh and Telangana in India. The first group were born in 2001-02 and the second in 1994–95. The research uses a range of methods to provide evidence and insights into the changing nature of child poverty at the beginning of the twenty-first century. The core research questions are about the causes and consequences of childhood poverty and about the means by which poverty is transmitted across generations. The study's goal is to offer credible evidence for use in policy engagement to ensure that children are considered in the shaping and implementation of pro-poor and poverty reduction strategies.

Drawn from each of the four major regions of the developing world, the study countries have diverse socio-economic and political characteristics. Twenty sites were selected in each country to illustrate diversity in terms of rural and urban location, ethnicity and religion. The sample comprises approximately 2,000 children from the younger age-group, and 1,000 children from the older age-group in each country. The sampled children were selected randomly from children of the right age in each of the study sites.

Young Lives takes a multi-dimensional view of poverty, seeing it as complex, dynamic and manifested primarily by diverse material disadvantages, susceptibility to risk and constraint on choice. Poverty is very often linked to social exclusion and exacerbated by restricted access to basic services, notably health, education and social protection. Childhood is the major analytic focus of the study and the children and their households are key units of observation and sources of evidence. The study measures children's experiences of poverty and its outcomes across many domains of well-being and development, including the physical, psycho-social, and cognitive. The study design recognises that children's lives are inseparable from the settings, institutions, systems of relationship and cultural processes in which their health, well-being, learning and development are embedded.

One of the major strengths of Young Lives is its innovative and comprehensive methodology, which harnesses the power of longitudinal research to illuminate the patterns of change and causal processes affecting children as they move from infancy to adulthood. It collects qualitative and quantitative data from children, caregivers, siblings, teachers and other community representatives. Young Lives has several distinct but linked components:

- **Child, household, and community surveys** – child, household and community questionnaires are designed to regularly gather information on the changing context of each community, the demographic and economic composition of households, children's access to and experience of health and education services, their daily activities and experiences, attitudes to work and school, and hopes and aspirations for the future. They also collect time-use data for family members, and information about children's weight, height and cognitive skills, including maths, language and literacy.
- **Qualitative longitudinal research** – 50 case study children from each country were selected from both age cohorts of the main sample, and four rounds of qualitative data collection documented the changing trajectories of their individual lives. Each round had a key theme. Fieldworkers used multiple methods based on observation, semi-structured interviewing, visual methods and group activities to record the understandings and perspectives of the children and the key adults in their lives.
- **School surveys** – a study of a sample of schools in Young Lives sites, including schools attended by the cohort children in all four countries, was introduced in 2010. The school surveys examine the student, class, teacher and school factors associated with the quality of learning outcomes at primary and secondary level. By linking information about schools and education quality to individual children and their households, the surveys provide evidence about the effectiveness of education in combatting inequality and preventing the intergenerational transfer of poverty. The survey design in each country is based on consultation with education stakeholders to ensure that they offer evidence relevant to current national policy. Data are collected through survey methods, including cognitive tests and psychosocial measures.
- **Sub-studies** – these are usually shorter enquiries based on sub-samples, using qualitative methods or surveys. They are based on context-specific issues, such as orphanhood in Ethiopia or the impact of the National Rural Employment Guarantee Scheme in India.
- **Collaborative research initiatives** – Young Lives has been involved in a number of collaborative research initiatives which either match Young Lives data with other datasets, or use them to help frame new research projects. These initiatives include studies of health and children's access to school, environment and family life in India and the UK, and the prevalence, timing, and benefits of recovery from early childhood growth failure in all four study countries.

The cohort surveys, school surveys and qualitative longitudinal research components are part of the same iterative research cycle, with each building on and elucidating the others. Analysis of survey data indicates research subjects and topics for more intensive focus, and provides patterns into which more detailed information is likely to fit. Qualitative case study data inform the design of the child survey instrument, and are used to identify research issues for inclusion in later surveys and possible new models for analysing survey data. These interactions between different types of data and analysis are a continuous process in Young Lives.

This approach has produced a unique, multilingual dataset which situates children's experiences of poverty in relation to the people around them, and the socio-cultural context, institutions, services and policies that shape their lives and opportunities. This dataset is being used to:

- understand how children are affected by poverty at every stage of their lives, and how boys and girls are affected differently.
- monitor the impact of macro-economic changes on children and their families.
- learn why some children do better than others by capturing relationships between diverse social forces including gender, religion, caste and ethnicity.
- trace the ways in which poverty is passed on across generations and examine why some households either move out of or fall into poverty.
- explore how poverty exacerbates or is exacerbated by other forms of adversity such as environmental and family shocks.
- examine perspectives on services and programmes, and understand whether and how households and children benefit from provision.
- analyse the impact of policies and services on children's life course and outcomes.
- trace the trends, processes of change and policy challenges that each study country faces, and consider the implications for other countries with similar trajectories of growth and development.

A Guide to Young Lives Research

This Guide provides an overview of how the Young Lives study has been carried out, giving insight into the diverse methods and processes involved in a complex longitudinal study made up of many different components. It offers lessons on building and managing research partnerships, designing and conducting multidisciplinary research, managing and analysing data, and using research to influence policy.

The Guide has been updated in 2017 by Virginia Morrow with Gina Crivello, Rosaleen Cunningham, Patricia Espinoza Revollo, Rhiannon Moore, and Anne Solon, based on the 2011 Young Lives Methods Guide written by Karen Brock and Caroline Knowles. It reflects on 15 years of experience, with each section summarising a different part of the research process, and considering:

- what has been done in each area
- what challenges have arisen
- how ways of working have changed and developed
- what lessons have been learned.

The sections of the Guide are:

1. [Young Lives: Study Overview](#)
2. [What Can Comparative Country Research Tell Us About Child Poverty?](#)
3. [What Can Longitudinal Research Tell Us About Children's Life-chances?](#)
4. [Research Ethics](#)
5. [Sampling](#)
6. [Child, Household and Community Surveys](#)
7. [Qualitative Longitudinal Research](#)
8. [School Surveys](#)
9. [Sub-studies](#)
10. [Piloting: Testing Instruments and Training Field Teams](#)
11. [Planning and Managing Fieldwork](#)
12. [Cohort Maintenance: Tracking and Attrition](#)
13. [Computer-assisted Personal Interviewing](#)
14. [Methods for Analysis](#)

If you are interested in more details about any particular aspect of the research process or section of the Guide, please email younglives@younglives.org.uk.

As well as producing academic and policy-related publications, Young Lives communicates research findings through numerous platforms, including a series of illustrated mixed methods books profiling children's biographies over time, social media, videos, podcasts, photography, and data visualisations.

We have also documented the impact of our research. See [Capturing a Picture of Change](#) and our [Theory of Change](#).

Section 2: What Can Comparative Country Research Tell Us About Child Poverty?

May 2017

‘Cross-national research is valuable, even indispensable, for establishing the validity of interpretations derived from single-nation studies. In no other way can we be certain that what we believe to be social-structural regularities are not merely particularities, the product of some limited set of historical or cultural or political circumstances.’ (Kohn 1987: 713)

We are often asked why Young Lives is conducting research in four countries, and why specifically Ethiopia, India (the states of Andhra Pradesh and Telangana), Peru and Vietnam. This section attempts to answer these questions.

Why Young Lives?

There was huge optimism at the turn of the century about the Millennium Declaration and the international community made a commitment to the Millennium Development Goals, many of which related to children and childhood – ending poverty, expanding enrolment in primary education, improving access to clear water, and reducing child mortality. DFID (the UK Department for International Development) wanted to understand the drivers and impacts of child poverty in low- and middle-income countries (LMICs) and to generate evidence to help design better programmes and policies. One way to achieve this was to initiate a long-term study to track children’s lives over the 15-year lifespan of the MDGs, in a range of countries. Young Lives was commissioned by DFID to do this.

The world has changed rapidly since 2000 and following the agreement of the Sustainable Development Goals (SDGs) in 2015, it is timely to set out a rationale for longitudinal research that helps to understand children’s development in a range of poverty contexts, by exploring what we have learned that will be useful for the proposed data revolution within the SDGs when: ‘the world must acquire a new ‘data literacy’ in order to be equipped with the tools, methodologies, capacities, and information necessary to shine a light on the challenges of responding to the new agenda’ (UN Secretary General 2014: 38).

Why these four countries?

Young Lives was designed to explore the correlates and outcomes of child poverty and well-being, and to contribute to international efforts to understand the consequences of poverty during childhood. Because it was intended from inception to be a policy-relevant study, the extent to which governments and civil society organisations in each country were committed to poverty reduction was a factor in country selection. The aim was to ensure a comparative perspective that also reflected a diversity of political and economic circumstances, and geographical,

social and cultural contexts and circumstances, including economic liberalisation, indebtedness and debt relief, conflict and natural disasters, and inequality. The existence of institutions in each country with capacity to undertake long-term research was also crucial to the choice of country (Attawell 2003).

The study countries were selected from 24 possible options. The countries chosen are in the four continents of the global South – sub-Saharan Africa, the Indian sub-continent, South-East Asia and Latin America. By undertaking research in a range of circumstances, Young Lives can explore children’s experiences of and responses to poverty, highlighting the diverse ways in which poverty affects children in specific communities, regions and countries.

What comparative research can tell us

While direct comparisons between countries are not possible, simultaneous research in four countries enables Young Lives to explore how patterns of relationships are similar or different across the countries, in relation to the long-term effects of poverty, the intergenerational transmission of poverty, and the unequal outcomes of economic and social development for children and young people. This is essential for understanding the causes and consequences of poverty in relation to the life-course, and for the possibilities of generalising (or not) from particular countries to other LMIC countries. Over time, it has become clear that the four-country design enables better understanding of the apparent failure of economic growth to alleviate childhood poverty in three countries now categorised as ‘middle-income’ (India, Peru and Vietnam). So while general living standards are improving, Young Lives households continue to be affected by negative events such as droughts, flooding or illness. Some children and households remain in poverty and, in some cases, families fall into poverty in spite of the increasing resources around them.

Undertaking research with children and families in four countries also enables us to track the effects of specific social policies over time: for example, relating to the consequences of increased enrolment in school and the variations in schooling effectiveness; policies relating to child labour; the impacts of policies relating specific issues (such as early marriage in Ethiopia and India); and the implications for children and their families of a multiplicity of social protection schemes. These include the Productive Safety Net Programme (PSNP) in Ethiopia, the Mahatma Gandhi National Rural Employment Scheme (NREGS) in India, the conditional cash transfer programme *Juntos* in Peru, and Programme 135, an area-based poverty reduction programme in Vietnam. Understanding why and how specific policies or programmes are effective in one country may help to pose important questions.

- Comparative analysis can give greater confidence that evidence in one country is applicable to others. For example, Young Lives evidence shows some children appear to recover in terms of height growth after malnutrition in infancy. This suggests the benefit of and need for continued nutritional support throughout childhood. That this pattern recurs across the four countries strengthens the argument that these findings are not country-specific, have wider applicability, and are not mere 'peculiarities'. This preliminary Young Lives finding is now being tested by other researchers using other datasets, with similar results.
- Comparative analysis also shows how norms vary between countries, and what this might mean for children. For example, global debates tend to emphasise the significant commonalities in discrimination experienced by girls and women. In relation to school expenditure, girls are systematically disadvantaged in India, but in Ethiopia and Vietnam we find that more Older Cohort girls are enrolled in school than boys (at age 18 to 19). Collecting data from children in a number of countries helps identify similarities and differences in these patterns, and enables us to explore underlying determinants of such disadvantage.
- Social policies are designed differently across countries, and comparisons highlight both national implications and messages for global learning. For example, analysis shows that children in primary schools in Vietnam learn more quickly than in the other countries, posing questions about why the Vietnamese school system seems to be more effective.
- Undertaking research simultaneously in multiple locations means that Young Lives contributes to learning in relation to methods, including the possibilities and limitations of trying to develop measures that can be used across cultures (for example testing children's learning levels in schools), which enables us to analyse policy effectiveness more closely.

Possibilities for long-term comparative research

Priorities in international development have shifted considerably over the lifetime of Young Lives. The world changed on September 11, 2001, and the focus of the international community shifted away from childhood poverty and towards the intersections of poverty and conflict and the potential for civil unrest. However, the renewed focus on poverty and inequality within the SDGs means that there is potential for Young Lives to contribute learning from its comparative longitudinal research with children. Research within and between middle- and low-income countries is essential in order to understand deepening inequalities. Findings from India, Peru and Vietnam, now middle-income countries, enable us to report on trends that are likely to affect low-income countries in the future.

REFERENCES

This section can be read alongside Section 3: [What Can Longitudinal Research Tell Us About Children's life-chances?](#)

Attawell, K. (2003) *International Longitudinal Research on Childhood Poverty: Practical Guidelines and Lessons Learned from Young Lives*, Working Paper 11, Oxford: Young Lives.

Brock, K. and C. Knowles (2012) '[Doing Longitudinal Research: Opportunities and Challenges in a Study of Childhood](#)', in J. Boyden and M. Bourdillon (eds) *Childhood Poverty: Multidisciplinary Approaches*, Basingstoke: Palgrave Macmillan.

Dercon, S. (2012) '[Understanding Child Poverty in Developing Countries: Measurement and Analysis](#)', in J. Boyden and M. Bourdillon (eds) *Childhood Poverty: Multidisciplinary Approaches*, Basingstoke: Palgrave Macmillan.

Kohn, M. (1987) 'Cross-national Research as an Analytic Strategy', *American Sociological Review* 52: 713-731.

UN Secretary General (2014) *The Road to Dignity by 2030: Ending Poverty, Transforming All Lives and Protecting the Planet*, New York: United Nations.

Wilson, I., S. Huttly and B. Fenn (2006) 'A Case Study of Sample Design for Longitudinal Research', *International Journal of Social Research Methodology* 9.5: 351-365.

Section 3: What Can Longitudinal Research Tell Us About Children's Life-chances?

May 2017

Most research used to gather household-level information in low-income countries is cross-sectional, collecting information on individuals at one point in time. Cross-sectional research monitors the magnitude of phenomena or problems, and who is affected, but is limited in its capacity to evaluate how and why problems persist. Studies like Young Lives go beyond this by collecting information about the same children over time as they grow up, and on the different elements that affect their lives: moving from a snapshot of children's lives to a filmstrip. This section describes how longitudinal research like Young Lives adds value for policy debates.

Cohort surveys: using numbers to establish patterns and relationships

There are two key benefits to longitudinal cohort studies:

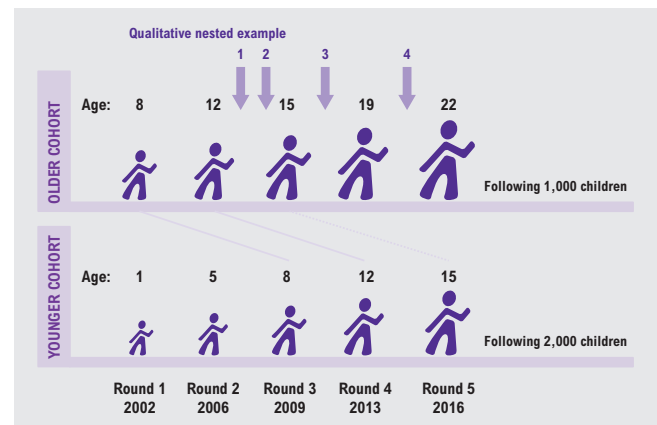
- They allow researchers to identify links between earlier circumstances and later outcomes.
- They show how persistent particular circumstances are, and thus enable evaluation of the differing impacts of continuing circumstances (or one-off changes) on later well-being.

Taken together, these advantages allow for policy-relevant insights into *which* children face particular disadvantages, how children develop, what matters, when it matters, and how policy can support children more effectively.

In a cohort study, a group of individuals sharing a common characteristic, often age, are followed over time. Cohort samples are sometimes representative of the group studied (for example, a random sample of children of a particular age) but that is not necessarily the key focus, since the primary aim is to study links between early experiences or characteristics and later outcomes, rather than to measure the scale of something within a given population. Observational cohort and panel studies form an important part of the data that social scientists analyse in order to understand social problems and to inform public policy, especially in high-income countries, and increasingly, in middle- and low-income countries.

Young Lives is following the lives of around 12,000 boys and girls in four low- and middle-income countries from early life and into adulthood. Young Lives uses an observational cohort design, set out in Figure 1. The study has two cohorts of children, born seven years apart, sequenced to collect information at the same age points (age 1, age 5, age 8, age 12, age 15 and age 19). This makes it possible to explore the relative contributions of age and historical time. Where cohort-sequential analysis is available, we can show the effects of events or policy changes that have affected one cohort rather than the other.

Figure 1: Young Lives study design



Identifying when differences emerge

The Young Lives cohort design enables analysis of children's physical, cognitive, or psychosocial developmental trajectories, and so by extension the timing of when inequalities emerge between children distinguished, for example, by gender, ethnicity or economic status. A range of approaches can be used to identify how early factors shape later outcomes – for example, identifying how long particular groups of children typically stay in school, how different groups of children perform in cognitive tests by particular ages, and so on. Such analysis can make it possible to identify which children most need support as well as the timing of potential interventions. Where information is available through childhood and into adulthood, this can be used to inform policy debates, for example, about the extent of social mobility, equality of opportunity and the intergenerational transmission of poverty.

Analysing what shapes later well-being

Background contextual information – such as parental education, socio-economic status, risks experienced, or services received – can be linked to children's development trajectories. For example, regression analysis enables researchers to 'control' for multiple possible relationships, and so identify underlying associations. Such techniques

contribute to identifying how poverty shapes children's opportunities and development. The use of data from different points in time reduces the problem experienced in cross-sectional studies of 'reverse causation' where, for example, low cognitive test scores seem to result in early school leaving, not the other way around. So-called 'natural experiments' (such as a new road, or a new public policy intervention) often arise during the course of longitudinal research and their effects on children's well-being/ outcomes can be explored. 'Quasi-experimental' techniques can be used to compare similar households, where only some are affected by a particular change, and with evaluation of the resulting differences between groups. Quantitative approaches identify statistical links and who is typically affected. Such knowledge can then inform analysis of qualitative research which seeks to understand the processes and mechanisms which shape well-being or outcomes.

Testing the 'dynamics' of well-being

Cross-sectional research (for example, the Demographic and Health Surveys or other studies such as UNICEF's Multiple Indicator Cluster Surveys) can show how many or which households are poor, and which children are stunted, but cannot show whether households remain poor or move in and out of poverty over time, or whether children remain stunted. Such dynamics – whether they reveal persistence or change – are substantively important, both because prolonged chronic poverty may matter more than short-term dips in and out of poverty, and because a study of dynamics also shows how earlier factors shape which households become poor. Identifying which groups become poor also enables analysis of the risk and resilience factors associated with disadvantage. For example, analysis has shown that some children can recover physically from early malnutrition, while others falter in their growth. This analysis has also linked relative height gain among children who were previously stunted with better-than-expected performance in cognitive tests.

Triangulating research approaches to inform policy

A weakness of observational studies is that not everything is measured, and so analysis risks so-called 'omitted variable bias'. The statistical models are only as good as the data collected and if key information is missing, then results might be misleading. This is an important concern for analysis of observational data. This risk is reduced by collecting a wide range of relevant background indicators and analysing them with statistical techniques such as regression analysis which control for multiple factors.

Comparisons are sometimes made between observational longitudinal studies, and randomised control trials (RCTs), where an intervention (for example, a new health promotion programme) is applied to one group randomly, and parallel information is collected from similar groups who do not experience the programme (a control group).

Observational longitudinal studies that collect data on many aspects of children's lives can be used to inform a wide range of policy questions, while RCTs can be used to give precise answers to specific questions – evaluating the specific changes in well-being attributed to a particular programme. Because RCTs rely on a random allocation of participants to an intervention and control groups, such an approach overcomes the problem of omitted variable bias (since it is expected that the impact of any unknown factors apply equally to intervention and control participants). Experimental approaches therefore contribute further to evidence-based policy, but suffer the weakness that while they can give precise answers to specific questions, they can only answer the question posed by the trial.

The key for evidence-based policy, therefore, is not to see observational or intervention approaches as competing methodologies, but rather to employ each to triangulate between methods, and to use one to inform the other; using multi-purpose observational cohort studies, for example, to identify areas worth examining in greater detail with experimental techniques or qualitative research.

Qualitative longitudinal research – deepening understandings

Young Lives is unusual in including qualitative longitudinal research with a nested sample of children – enabling qualitative analysis to be combined with analysis of the household survey data. Repeat visits to the same children show how experiences, circumstances, motivations and perceptions change with age and experience.

Qualitative research enables us to:

- **explore children's experiences**, their agency, priorities, and their interpretations and understandings of their situations, and how these change over time. This helps to explain the dynamics of childhood poverty. Findings from qualitative longitudinal research show how children and families are vulnerable to economic difficulties that accumulate over time, and how changing circumstances (at home, work or in policy) affect everyday lives over time.
- **capture the links** between differing aspects of children's lives. This enhances theory-building related to the life-course, showing the intersections between social determinants/structural factors – such as availability of resources, including economic, educational, health – and individual lives over time, from children's point of view (Morrow and Crivello 2015).
- **explain diverging experiences and trajectories**. Using mixed-methods approaches has been vital for policy and communications purposes, adding richness and depth. It also enables us to question and challenge dominant assumptions about children as passive recipients of social change, by exploring how children actively navigate their way through childhood in resource-poor settings.

Research example: understanding migration in childhood

Qualitative longitudinal research deepens insights from the survey about children's movements across place and time. The complexity of these movements is difficult to study through large-scale cross sectional surveys. For example, Elmer, in Peru, had migrated from his place of birth to Lima at the age of 12 to help his sister look after her children while she and her husband worked. In exchange, she paid for Elmer's upkeep and schooling. The following year, Elmer returned to the village. His parents had moved to a village where they had purchased a plot of land, and the children, including Elmer, moved to a different village to attend school. Each weekend the children walked three hours to help their parents in the fields. In 2013 we found Elmer still living in the village, but by 2014 he had returned to his sister in Lima. Comparing cases across multiple rounds of data means we can explore children's mobility and migration histories in greater depth by tracing their biographies (Crivello 2015).

Research example: why do girls marry early? Understanding accumulated disadvantage

In all the Young Lives study countries, young people say they want to delay marriage until they are in their mid-20s, yet cohort data allow us to compare earlier aspirations with later realities – many of girls in the Young Lives sample in Ethiopia are still marrying below the legal age of 18. Longitudinal analysis demonstrates the complex reasons why some girls marry early, and while survey research can show factors that increase the chances of early marriage, it cannot demonstrate how multiple difficulties accumulate to affect girls' lives. For example, Haymanot's mother's ill health meant that Haymanot worked from an early age to support her family and had to miss school. She married at age 15 which meant her family situation improved, she could support her mother with access to better food, and she no longer needed to work so hard. However, her husband divorced her, and she was last reported living with her mother and baby. Disadvantage accumulated over time for Haymanot, but there were key intervention points – at school, or through access to health care for her mother – that could have improved her life and reduced the chances that she would marry young (Morrow and Crivello 2015).

Making longitudinal research useful for policy

Longitudinal research allows exploration of the cumulative experience of particular policies on young people's trajectories – the 'long view' rather than the 'short view'. A key theme in recent years has been identifying the critical period of early childhood in improving long-term outcomes: longitudinal analysis is needed to form such conclusions. Longitudinal studies help separate out groups affected by 'episodic' deprivation from those who experience 'persistent' deprivation, and so both examine which groups are facing chronic disadvantage. By collecting information before change happens, cohort studies can go beyond counting who is disadvantaged to understanding why disadvantage occurs, by identifying earlier factors associated with later disadvantage and by taking a holistic view of how the different domains of children's lives – their health, learning and social development – are shaped. Longitudinal research reveals key points when policy interventions are most timely and how investments in one area of children's lives, such as nutrition, may support development in another, such as learning, showing the importance of working across social policy silos.

Debates on the Sustainable Development Goals have emphasised the need for a data revolution, with better and timelier statistics to improve monitoring and measurement (UN 2014). Clearly this is crucial, but better policy requires tools to evaluate, not only to measure, social problems – and longitudinal analysis can play this role.

Cohort studies give a powerful sense of what matters in people's lives. There is ongoing interest in funding more longitudinal studies, and plans to start a community of practice (a [Global Longitudinal Research Initiative](#)). Such studies are investments for the future, as their value and the power of the data increases with each round of research. There is also value in longitudinal research maintaining a 'generalist' and general purpose, broad design, so that data collected today can be used and analysed flexibly to inform future, as yet unknown, policy questions.

Policy knowledge from the UK cohort studies

There are a number of national birth cohort studies in the UK: the National Child Development Study (of children born in 1958), the British Cohort Study (1970), and Millennium Cohort Study (2000), all of which provide vital evidence across a range of policy domains.

The UK Academy of Social Sciences has identified contributions from key cohort studies, with examples including to:

- help identify those groups with the highest needs, and thereby focus the attention of organisations aiming to reduce disadvantage towards those groups.
- bring together a clear evidence base on what mattered for pre-school interventions, motivating a more effective joined-up approach.
- identify those young people at risk of offending, and working with them to help address the underlying reasons and to prevent offending.
- inform the UK's policy approach to child poverty by identifying multiple disadvantages poor children experience.

For example, findings from the 1970 cohort of the British Cohort Study about children's cognitive development and socio-economic background informed the introduction of free part-time childcare for under-4 year olds, and there are numerous other examples of longitudinal research evidence informing social policy in the UK. The latest UK birth cohort study is the Life Study, involving more than 80,000 babies born between 2014 and 2018 and their families.

Source: Academy of Social Sciences 2013

Cohort studies in low- and middle-income countries

Longitudinal birth-cohort studies are unusual in low- and middle-income countries, although there are some important ones:

- [Cebu Longitudinal Health and Nutrition Survey](#), The Philippines, established 1983
- [New Delhi Birth Cohort Study](#), India, established 1969-1972
- [Pelotas Birth Cohort Study](#), Brazil, established 1982
- [Birth to Twenty \(BT20\)](#), Johannesburg-Soweto, South Africa, established 1990
- [Mauritius Child Health Project](#), established 1972
- [Gansu Survey of Children and Families](#), China, established in 2000
- [Kagera Health and Development Survey 2](#), Tanzania, established 1991
- [Kwazulu-Natal Income Dynamics Study \(KIDS\)](#), South Africa, established 1998
- [Chilean Longitudinal Survey of Early Childhood \(Encuesta Longitudinal de la Primera Infancia\)](#), established 2009
- [The Jamaican 1986 Birth Cohort Study](#)

REFERENCES

This section can be read alongside Section 2: [What Can Comparative Country Research Tell Us About Child Poverty?](#)

Academy of Social Sciences (2013) *Making the Case for the Social Sciences: Longitudinal Studies*, London: Academy of Social Sciences/Sage.

Boyden, J., S. Dercon and A. Singh (2014) 'Child Development in a Changing World: Risks and Opportunities', *World Bank Research Observer*, early online publication Sept 2014.

Brock, K., and C. Knowles (2012) '[Doing Longitudinal Research: Opportunities and Challenges in a Study of Childhood](#)', in J. Boyden and M. Bourdillon (eds.) *Childhood Poverty: Multidisciplinary Approaches*, Basingstoke: Palgrave Macmillan.

Centre for Longitudinal Studies (2010) 'Case Study on the Impact of IOE Research: The British Birth Cohort Studies', London: Institute of Education.

Crivello, G. (2015) '["There's no future here": The Time and Place of Young People's Migration Aspirations in Peru](#)', *Geoforum*, early online publication April 2015.

Dercon, S. (2012) '[Understanding Child Poverty in Developing Countries: Measurement and Analysis](#)', in J. Boyden and M. Bourdillon (eds) *Childhood Poverty: Multidisciplinary Approaches*, Basingstoke: Palgrave Macmillan.

Dorman, P. and M. Woodhead (2015) *How Inequalities Develop through Childhood*, UNICEF Office of Research–Innocenti Discussion Paper Series: Perspectives on Equity, Florence: UNICEF Office of Research.

Morrow, V., and G. Crivello (2015) 'What Is the Value of Qualitative Longitudinal Research with Children and Families for International Development?', *International Journal of Social Research Methodology: Theory and Practice* 18.3: 267-280.

Morrow, V., and R. Singh (2015) 'Children's and Parents' Perceptions of Corporal Punishment in Schools in Andhra Pradesh, India', in J. Parkes (ed.) *Gender Violence in Poverty Contexts: The Educational Challenge*, Abingdon: Routledge.

Singh, A. (2014) *Emergence and Evolution of Learning Gaps across Countries: Panel Evidence from Ethiopia, India, Peru and Vietnam*, Working Paper 124, Oxford: Young Lives.

Society for Longitudinal and Life-course Studies www.slls.org.uk

UNICEF Office of Research (2014) 'Cohort and Longitudinal Research', www.unicef-irc.org/research/266

United Nations Secretary General (2014) 'The Road to Dignity by 2030: Ending Poverty, Transforming All Lives and Protecting the Planet', New York: United Nations.

Section 4: Research Ethics

May 2017

Research ethics exist to ensure that the principles of justice, respect and avoiding harm are upheld in research processes through the use of agreed standards. While these basic principles are universal, they are open to differing interpretations and understandings which revolve around the central question of balancing the goals of a piece of research with the interests and rights of its subjects. Young Lives takes a positive view of research ethics as enabling high-quality research while respecting these key principles.

There are particular ethics questions which arise when doing research with children and families, and with poor communities in developing countries. There are also ethical challenges involved in research that aims to influence policy. Young Lives has had to develop awareness of the ethical dimensions of the study through all its stages, particularly in respect of the power relations between research teams and the children and families who participate in the study.

Background debates in research ethics

In developing an approach to ethical social research, Young Lives has drawn on a growing literature on the governance of social research which identifies the key qualities of integrity and transparency, and the basic principles of free and informed consent, confidentiality and anonymity (ESRC 2010). It has also drawn on existing protocols designed to protect children from abuse through awareness, prevention, reporting and responding (Save the Children 2003).

The approach to research ethics taken in the design stages of Young Lives was strongly grounded in the paradigm of medical and epidemiological research, where there are often direct, visible links between research and its risks for and effects on participants. In this field, the development of committees, standards and ethical protocols are well established in high-income countries. Nonetheless, there may be considerable disjunction between protocol and practice, which depends on how the procedures intended to implement protocols are actually used by fieldworkers, and how research participants interpret and experience them (Fairhead et al. 2005). This highlights the importance both of developing clear, transparent structures for research governance, and of monitoring and understanding what happens when they are used with research participants. It also points to the potential for misunderstanding between researchers and researched, which is mediated by the power relations between them and frequently shaped by wealth, social class, gender, ethnicity, caste, or age.

In contrast with the field of medicine, where the effects of research on participants are often physical, it is harder to trace the impacts of social research. Possible negative outcomes include damage to people's futures, reputations and relationships through public reports and influence on policies or practices. There is also the potential to exploit research participants from poor communities by failing to consider how they might benefit from the research.

Social science research uses a range of methods. Responding to a questionnaire survey, for example, involves a regular encounter between enumerators and respondent within the clearly defined boundary of the questionnaire. Qualitative research may use mixed and multiple methods to work iteratively and reflexively, and often builds the trust of the research participant in order to learn about their concerns in depth.

Cutting across these ethics questions is the need to develop clear communication about the study and why it is being done. Translation and understanding has particular importance when research covers multiple countries, cultures and disciplines.

Building an ethical social research programme

The foundations of the approach Young Lives takes to ethics were established before its research activities began. The study proposal was checked against the ethics standards of each of its six original partner institutions, and the study was reviewed by the London School of Hygiene and Tropical Medicine ethics committee. A pilot phase which developed and tested the questionnaire in South Africa in 2001–02 was given ethical approval by the Rand Afrikaans University. Save the Children's 2003 child protection policy was influential in shaping the ethics approach in the first survey round. The study subsequently received approval from research ethics committees from the Social Science Division of Oxford University, IIN in Peru, and in 2016, from research ethics committees in Ethiopia, India and Vietnam (see Nuffield Council on Bioethics 2016).

As the programme moved towards the pilot phase of fieldwork, country teams in Ethiopia, India, Peru and Vietnam began to translate ethics from paper into practice during the fieldwork stages. As far as possible, techniques for achieving this were developed collaboratively. Training for qualitative and survey fieldworkers was designed and delivered by the country teams with support and input from the coordination

team in the UK, and included sessions discussing ethics. Survey manuals contained detailed ethics guidance, while a set of ethics protocols for the qualitative research was prepared and adapted to be locally relevant in each country.

Following piloting of the qualitative research methods in 2007, a Memorandum of Understanding for qualitative fieldworkers was developed in collaboration with the qualitative research teams. This sets out guidance about research procedures and respectful communication with research participants. It was adopted across the study, and is also used with the survey teams. As far as possible, Young Lives has tried to work with the same field teams in successive rounds, and training for fieldworkers is ongoing, with sessions taking place before each new round of survey or qualitative research.

Once data collection was under way, it became important to ensure a consistent focus on ethics throughout the rest of the research process, from data storage and analysis to the use of findings to influence policy. The longitudinal character of Young Lives means that there are many rounds of visits to survey sites. Each visit generates reports and information on ethics. Ethics questions are recorded as they arise, and qualitative data are coded and analysed for participants' views of their involvement with Young Lives. This has generated the collaborative, iterative development of a shared understanding of, and collective approach to, ethics by researchers in different places whose work focuses on diverse themes and activities. The approach to ethics that has emerged has several key cornerstones:

- *Informed consent.* Young Lives works on the principle that researchers must obtain informed consent from parents or caregivers and from children themselves, from as early an age as possible. The purpose of the research is clearly explained every time fieldworkers visit a community, emphasising that Young Lives is a study, not a development project. Consent is understood as an ongoing process, and is frequently re-checked.
- *Anonymity.* The Young Lives children and their families share a great deal of personal information and we have a responsibility to ensure that their confidentiality and identities remain protected. Names of people and places are removed from Young Lives data before archiving, and a set of pseudonyms is used in publications.
- *Respect and protection for children.* The Memorandum of Understanding covers how to behave respectfully towards children, to have an awareness of potential signs of child abuse, and establishes a structure for reporting and responding when concerns arise. We are also conscious of the need to maintain a gender balance within fieldwork teams, particularly as the children enter adolescence.
- *Working with local researchers.* This helps minimise the risk of inadvertently causing damage to participants through misunderstanding local contexts.
- *Flexibility about rewards and compensation.* The research teams in study countries each take a culturally appropriate approach towards compensating research participants, ranging from paying them for their time to giving small gifts to thank them.

- *Reporting back to communities.* With each study round we have developed new ways to provide information about Young Lives research findings to respondents, enabling the respectful implementation of the study. Findings are presented at meetings in a range of locally relevant ways that are intended to be accessible to all members of the communities, and that highlight the usefulness of the data they are providing.

Practical challenges and lessons learned

Many of the challenges encountered in implementing ethical research practice relate to misunderstandings between researchers and participants, in particular about the nature and purpose of the research. In Ethiopia and Peru, for example, this has meant researchers having to assuage the fears of parents that Young Lives will take their children away.

In all four countries, the presence of researchers has raised the expectations of people who live in poor communities that they will benefit directly from their participation. Despite consistent efforts at clear communication, researchers have encountered widely differing understandings of the research among participants. In particular, they have found that the use of the word 'project' is loaded with expectations of financial and material benefits. In Ethiopia and Vietnam, where government departments are involved in data collection, and sites where Save the Children had been a research partner, association between Young Lives, government and NGOs sometimes further contributed to misunderstandings about the research's purpose and outcomes. One consequence in some places has been difficulty in renewing informed consent as participants have come to fully understand that the research is not going to directly improve their lives.

High expectations about possible benefits of the research are closely related to the question of compensating participants for their time. Here the challenge has been to balance different understandings of the value of people's time, their willingness to undertake research activities for the common good, and the reality of their having to take time away from work to talk to researchers. Over the course of the study, the question of payment has become more important as local economies have become increasingly monetised, and people have become more aware of the financial value of their time.

While explaining that the research will not bring direct benefits in the form of 'projects' or 'programmes', teams do explain to participants that the information they provide will be used to try to improve the situation of children more broadly. This raises the questions of reciprocity in research and of how best to learn about what people think would be useful to them without this being seen as a promise or an aid intervention. It also demands that research teams explain to participants and their communities how messages from the research are being used to advocate for change.

The key strategy for meeting the challenges outlined above has been a reflexive approach to ethics which continues to

develop. In many cases, this has meant making relatively small changes in response to particular challenges – for example, describing Young Lives as a ‘study’ rather than a ‘project’, or ensuring that research teams do not travel to sites in vehicles with NGO logos. It has also meant investing a great deal of researchers’ time in patient explanation of the research and reassurance about its motives. An example from India, which records a researcher explaining random sampling to parents, illustrates how this has been undertaken in locally relevant terms:

“Now, let me explain why we have selected [your child] for the research. While cooking rice, you will take some grains and test whether it’s cooked or not. You will not check the whole rice. In the same way, we select some children to know how they are and to know about their lives, and to know how the lives of children are in [this community]. That’s why [your child] has been chosen.”

Equally important is working in a way that allows people to express their concerns and worries, as illustrated in this extract from group discussions with local authority workers in Peru, one of whom has said that there are rumours in the community that Young Lives is going to take children away:

Fieldworker: “It is good that you’re mentioning this because, as the authorities, it is good that you’re informed [...] No one is going to take any of the children, no way are we taking them away from their homes. In fact, what we want is to see how they grow up in their homes, how some improve and others do not, and the reasons why some make progress and others do not [...] You can tell us any fears or worries that you have. In each visit we hand out a leaflet, a letter for the families with telephone numbers and address, and you can call

and ask any time. We’ve also handed out a letter in the municipality, where you’ll find our phone numbers and addresses [...]”

Local authority worker: “Are you all Peruvians?”

Fieldworker: “Everyone; we are all as Peruvian as *yucca* [cassava] and potatoes!”

While broad shared ethics practices are crucial, these need to be applied with some flexibility according to each situation that arises. The importance of understanding dynamic local contexts cannot be overemphasised. When research teams visit sites, they are not going into neutral situations. Circumstances can change very rapidly, and these changes themselves need careful documentation.

Understanding local contexts is equally crucial to explanations of how research participants respond to being involved in a longitudinal data-gathering exercise. An integral part of an ethical approach to this kind of study is to follow the effects of participation on children and their families over time, partly to try to ensure that they are not negative.

Maintaining an iterative approach means continuing to adapt the existing Memorandum of Understanding with research teams and fieldworkers, through learning from participants and adapting methods and standards to fit their views more closely. Similarly, questions on informed consent and managing raised expectations need constant reflection. For Young Lives perhaps the most challenging part of maintaining a responsive approach to ethics is following through on policy work in ways that can demonstrate potential change in children’s lives, thereby meeting its fundamental ethical responsibilities to children, families and communities.

REFERENCES

Crivello, G. with V. Rojas, Y. Tafere, U. Vennam, and Vu Thi Tranh Huong (in preparation) ‘We’re like family now’: negotiating relationships, reciprocity and closure at the end of a qualitative longitudinal research collaboration.’

Alderson, P. and V. Morrow (2011) *The Ethics of Research with Children and Young People: A Practical Handbook*, London: Sage.

Economic and Social Research Council (2015) *Framework for Research Ethics*, Swindon: ESRC.

UNICEF (2013) *Ethical research involving children* www.childethics.com

Morrow, V. (2009) *The Ethics of Social Research with Children and Families in Young Lives: Practical Experiences*, Working Paper 53, Oxford: Young Lives.

Nuffield Council on Bioethics (2016) *Research with children: ethical challenges around the world*, London: Nuffield Council on Bioethics.

Save the Children (2003) *Child Protection Policy*, London: Save the Children Alliance.

Section 5: Sampling

May 2017

The overall objective of Young Lives research is to produce detailed, long-term panel data about the causes and consequences of childhood poverty, the impact of pro-poor policies, and the means by which poverty is transmitted across generations in Ethiopia, India, Peru and Vietnam.

In a cohort study collecting data about the same group of people over a specified period of time, initial decisions about sample selection are a crucial determinant of the outcome of the research and the ways the data can be used.

Key considerations and challenges in sample design

Designing a sampling strategy for Young Lives involved striking a balance between many competing needs (Wilson, Huttly and Fenn 2006). Perhaps the most basic of these was the tension between selecting a statistically viable sample which was not only within the study's budget but also feasible to manage given the geographic and infrastructural characteristics of the four countries and the degree of detail demanded by the research objectives. As a result, the Young Lives sampling method was never designed to be nationally representative of children of a specific age, as achieving this within the available budget would have meant limiting the number of countries in the study. Instead, the sampling method was intended to generate a large enough sample for general statistical analysis, and to be systematic and clearly justified. This has shaped the character of Young Lives as 'an in-depth study of relationships between pieces of information, rather than an instrument to collect national statistical results' (Wilson, Huttly and Fenn 2006: 358).

The objectives of Young Lives, established prior to the sample design stage, were particularly important in shaping the approach taken to sampling. Studying the causes and consequences of childhood poverty meant designing a sample that included a high proportion of poor children, but which also included other children with whom their experiences could be compared. This was achieved by over-sampling poor areas, and then randomly selecting children of the right age within the selected communities. Avoiding a sample comprised exclusively of poor children not only provided opportunities to compare poor and better-off children, but also minimised the chance that the study results would be rejected on the grounds of not being representative.

The sample also had to be suitable for use to obtain data about children's experiences of poverty at different levels, including

the community and the household. This need for detailed site-level data, together with the logistical considerations that arise from widely dispersed rural populations poorly served by transport infrastructure, determined that children would be selected in geographically compact sites rather than randomly across countries. As well as being predominantly located in poor areas, these sites were selected to reflect heterogeneity of ethnicity and religion in country populations.

These two needs – to over-sample the poor and to produce in-depth data about sites as well as children – were reconciled through the development of a multi-stage sampling procedure, adapted from sentinel site monitoring methods. The concept of sentinel site monitoring comes from public health studies, and involves the purposive sampling of a small number of settings, deemed to represent a certain type of population or area, which are then studied in a consistent way at relatively long intervals. Under the sentinel site monitoring system adopted by Young Lives:

- sentinel sites in each study country were selected non-randomly, with rich areas excluded from the sample and poor areas purposively over-sampled.
- children in the right age group in the selected sites were sampled randomly.

Implemented in 2002, this procedure resulted in the random selection of 2,000 infants (aged between 6 and 18 months) living in 20 sites mostly located in poor areas of each country. At the same time, 1,000 older children (aged 7 to 8 years) were also randomly selected in the same sites.¹ Initially, work with these older children was intended to be limited to the testing of instruments and methods for later use with the younger children. Subsequently, however, the decision was taken to retain the older cohort because of the value of inter-cohort analysis which provides unique information about changes over time. As such, the two age cohorts of children form the panel for the Young Lives longitudinal survey rounds, as well as the foundation from which sub-samples for other elements of Young Lives – such as the [qualitative research](#) and [school surveys](#) – were later drawn.

Sentinel site selection

For each country, site selection protocols were written to transparently describe the sequence of decisions that were made in selecting and defining sites and to systematise procedures for over-sampling poor areas. Proposed criteria and procedures for site selection were extensively discussed

¹ Because of resource constraints, only 700 older children were selected in Peru.

with the national Young Lives Advisory Panels and amended according to these discussions. Each of the country study teams used slightly different processes to arrive at a non-random selection of sites. Each process involved several stages.

In Ethiopia (see Outes-Leon and Sanchez 2008):

- Five regions were selected out of a total of nine, accounting for 96 per cent of the national population.
- Three to five districts were selected in each region, with a balanced representation of food-deficient rural and urban districts. Where official statistics were not available, this classification was made through consultation with local officials.
- Since districts were too large, in terms of both area and population, to be considered as sentinel sites, at least one peasant association or *kebele* (the lowest level of administration in rural and urban areas respectively) per district was selected as a sentinel site, with the key criterion being the possibility of finding at least 100 households with a 1-year-old child and 50 households with an 8-year-old child.
- A village was randomly selected within each sentinel site.

In Andhra Pradesh in India (see Kumra 2008):

- Site selection aimed to ensure a uniform distribution of sample districts across the state's three agro-climatic regions, and the inclusion of at least one poor and one non-poor district from each region.
- In order to make this selection, districts were classified and ranked according to a relative development index which aggregated economic, human development and infrastructure indicators. A representative group of 12 poor and non-poor districts was chosen from a total of 23, covering 28 per cent of the population of the state.
- *Mandals*, administrative areas containing between 20 and 40 villages, were deemed to be the appropriate size to be sentinel sites. The second step of sampling was choosing mandals within the selected districts. All the mandals in each district were ranked and selected based on a second set of economic, human development and infrastructural indicators constructed using available mandal-level data.
- Each mandal was divided into four contiguous geographical areas and one village was randomly selected from each.

In Vietnam (see Nguyen 2008):

- Five out of a total of nine provinces were selected to over-emphasise poor regions and to ensure even coverage of urban, rural and mountainous areas, and of the north, central and southern regions. The selection was made through a process of iterative consultation with a range of different actors including government, donors and NGOs.
- Working groups of provincial government staff were established to select sentinel sites in each province. All communes in each province were ranked by poverty level according to the degree of infrastructural development, the

percentage of poor households, and child malnutrition status. As well as level of poverty, other criteria included commitment to the research from local government officials, logistical feasibility, and adequate population to constitute a sample of children of the right age.

- Four communes were selected as sentinel sites in each selected province, 48 per cent from those ranked as poor, 29 per cent from those ranked as average and 23 per cent from those ranked as above average.

In Peru, while the research team followed the general principles of sampling agreed for the whole study, there were significant differences in sample design. Here, the sentinel sites were chosen using a multi-stage, cluster-stratified, random sampling approach (see Escobal and Flores 2008).

- Sentinel sites in Peru are districts, of which there were 1,818 at the time of sampling. A national poverty map developed in 2000 by the *Fondo Nacional de Compensación y Desarrollo Social* (National Fund for Compensation and Social Development) was used as the basis for site selection. This map ranked all districts according to a poverty index calculated from variables which included infant mortality rates, housing, schooling, roads and access to services.
- To achieve over-sampling of poor areas, the 5 per cent of highest-ranking districts were excluded from the sampling process. The remaining districts were listed in rank order with their population sizes and divided into equal population groups. A random starting point was selected and a systematic sample of districts was chosen using the population list. Selection runs were made by computer and the resulting samples of districts were examined for their coverage of rural, urban, peri-urban and Amazonian areas, and for logistical feasibility. The sample of districts that best satisfied the requirements of the study was selected.
- Maps of census tracts (small geographical areas that can be covered by one census worker in a short time) were obtained for each of the selected districts, and one tract per district was selected using random number tables. In each selected tract, all *manzanas* (blocks of housing) and *centros poblados* (clusters of housing) were counted, and one was randomly selected for each district.

Child selection

Having selected 20 sentinel sites in poor areas, households containing children in the right age groups were randomly selected. While the exact procedures used by each study team were adapted to local circumstances, there was careful and transparent documentation of protocols to ensure:

- cost-effective field procedures for traversing each site.
- reasonable control of biases, for example due to the unavailability of any respondent from a household during the listing sweep through the site.
- a sample equivalent to one drawn at random from all possible qualifying households in the area (Wilson, Huttly and Fenn 2006).

In some cases, the local procedure required an exhaustive screening sweep through an administrative area like a sub-district to create a numbered list of all qualifying households, and then drawing a random sample from this list. In other cases, where a defined area was to be sampled rather than fully covered, the process included a stage adapted to the geography of households. In some densely populated urban areas, for example, this entailed selecting particular streets or alleyways as sub-units for seeking qualifying households. In some sparsely populated areas, by contrast, it entailed the use of line transects, which involved walking in a straight line between identifiable landmarks and selecting all households within 50 metres of the line (Wilson, Huttly and Fenn 2006).

The approach in each of the four countries was as follows:

- In Ethiopia, a village within each sentinel site was randomly selected and all the households on the periphery were interviewed until 150 eligible households were located.
- In Andhra Pradesh, a door-to-door listing schedule was completed in order to identify eligible children.
- In Vietnam, a door-to-door screening survey for children the right age was carried out in each commune, and simple random sampling applied to the list.
- In Peru, all households in each selected manzana or centro poblado were visited by fieldworkers to identify children of the right age. If not enough children were found using this method, then neighbouring manzanas and centros poblados were visited until the total was achieved
- The Ethiopian sample was compared with the 2000 Demographic and Health Survey (DHS) and the 2000 Welfare Monitoring Survey. The analyses showed that households in the Young Lives sample were slightly better-off and had better access to basic services than the average household in Ethiopia, but that they held less land, owned less livestock, and were less likely to own a house (Outes-Leon and Sanchez 2008).
- The Andhra Pradesh sample was compared with the 1998/9 DHS. The analysis showed that households in the Young Lives sample were slightly wealthier than households in the DHS sample. They had better access to public services and owned more assets, but they were less likely to own their own house, and the mothers of Young Lives children were less likely to breastfeed or to have received an antenatal visit (Kumra 2008).
- The Vietnam sample was compared with the 2002 DHS and the 2002 Vietnam Household Living Standard Survey. The analysis showed that households in the Young Lives sample were slightly poorer than the households in the other samples. They owned fewer assets, were less likely to own their own house, and were more likely to be registered as poor by their local authorities (Nguyen 2008).
- The Peru sample was compared with the 2000 DHS, the 2001 Peru Living Standard Measurement Survey (LSMS) and the 2005 National Census. The analysis showed that the poverty rates of the Young Lives sample were similar to the urban and rural averages derived from the LSMS, and slightly wealthier than households in the DHS. Young Lives households owned more assets and had better access to public services such as electricity and drinking water than households in the other surveys (Escobal and Flores 2008).

The Young Lives sample and national datasets

Although the Young Lives sample is not and was never intended to be nationally representative, it is important to understand how it compares with larger samples from other studies and surveys which are. In 2008, each Young Lives country sample was compared with one or two other samples to examine and discuss differences and highlight both expected and unexpected biases. This was an important step in situating the Young Lives samples in broader national contexts, and understanding what inferences could be drawn from the findings of the study.

In all four cases, analysis showed that despite biases, the Young Lives sample covered the diversity of children in each country. Therefore, while not suited for simple monitoring of child outcome indicators, the Young Lives study is an appropriate and valuable instrument for analysing causal relations, and modelling child welfare and its longitudinal dynamics.

REFERENCES

Galab, S., P.P. Reddy, and R. Singh (2014) *Young Lives Survey. Design and Sampling in Andhra Pradesh, India*, Oxford: Young Lives.

Pankhurst, A., and T. Woldehanna (2014) *Young Lives Survey. Design and Sampling in Ethiopia*, Oxford: Young Lives.

Sanchez, A., M. Penny, and M. Lizama (2015) *Young Lives Survey. Design and Sampling in Peru*, Oxford: Young Lives.

Thang, N. and L.T. Duc (2014) *Young Lives Survey Design and Sampling in Vietnam*, Oxford: Young Lives.

Wilson, I., S. Huttly and B. Fenn (2006) 'A Case Study of Sample Design for Longitudinal Research: Young Lives', *International Journal of Social Research Methodology* 9.3: 351-365.

Section 6: Child, Household and Community Surveys

May 2017

The longitudinal survey at the centre of Young Lives consists of a set of questionnaires administered by interviewers every three or four years with all 12,000 children, their primary caregivers, and key informants in their communities. Together with Young Lives qualitative longitudinal research, which involves successive rounds of in-depth research with sub-samples of the children, and the Young Lives school survey, it forms the foundation of the longitudinal study.

Young Lives views [childhood poverty as a complex, multi-dimensional phenomenon](#) (Boyden and Dornan 2011). To understand more about its causes, consequences and transmission across generations, the study must therefore gather a broad range of data about the Young Lives children, structured to allow both [multi-level and longitudinal analysis of a range of determinants and outcomes of poverty](#). To achieve this, each survey round consists of three closely linked components – child, household and community context surveys – which make use of several tools and are applied to different respondents.

- **The child/individual survey** has been designed to be administered to the sample children after they reached the age of 8, and therefore provides data at the level of the individual. In these surveys children were asked about their perceptions of well-being, their daily activities, their attitudes to school and work, how they feel they are treated by others, and their future aspirations. In subsequent rounds, the child survey also asked children about their time use, mobility, and complete school histories. In early rounds, when the Younger Cohort children were under 8 years old, similar questions about the health, well-being and care of the child from birth onwards were asked to the caregiver as part of the household survey.
- **The household survey** covers basic information about all household members, as well as covering a range of subjects including parental background and education, livelihood activities, assets, time use, food and non-food consumption and expenditure, recent economic change, social capital, household members' health, and access to basic services. The children's primary caregivers were asked about child care dynamics, their perceptions of, and attitudes towards, a range of subjects, and their aspirations for their child and family.
- **The community context survey** provides background information about the social, economic and environmental context of each community where the Young Lives children live, covering topics including population, ethnicity, religion and language, economic activity and employment, and infrastructure. It also provides a detailed information map of the health, education and child protection services that are available to community members.

As well as content, key considerations in designing the protocols for each component have been respondent burden, question clarity, potential for recall error, cultural sensitivity and developing clear definitions of basic terms like 'household'. For the community context survey, central considerations have been devising questions suitable for both rural and urban settings, and deciding what kind of community profile is necessary to inform the analysis of the household and child data.

In each round, the [research protocols are piloted and revised](#) before they are finalised, and detailed justification documents drawn up for each section that explain why particular approaches and emphases were favoured. While some basic household data were collected in each round, each component was reviewed and new modules have been introduced to reflect the age of children and the issues that they and their families face at each phase of childhood, spanning infancy through to early adulthood. Adaptations have also taken account of learning from each experience of applying the survey, and of conceptual and theoretical developments over time. Country-specific questions about policies and programmes affecting children are also included. Throughout the five rounds, the survey design has been adapted and altered in several different areas.

- **Respondents:** The principal survey respondents are the Young Lives children, their adult carers, and key informants in the community. In each round, however, there are changes in the distribution of questions between these informants; as the children get older, more questions are directly addressed to them. Some rounds have also introduced new informants. Since Round 3, for example, Younger Cohort siblings have been interviewed on a number of topics, and the community survey became more focused on services, requiring different key informants to be sought.
- **Content of sections:** Retaining core content unchanged across rounds is an important principle of longitudinal survey design, as this allows direct comparison between rounds. Nonetheless, while most section content does not change, some alterations are essential to take account of life course and contextual changes. Some sections have been developed to address questions emerging from baseline data collected in Round 1, while others have been strengthened to gather more detailed information in particular thematic areas such as health and education. Other sections have become shorter as questions have been dropped where information is unlikely to have changed – for example, about a child's first language. In some places, questions – for example, those about political capital – have been dropped in particular countries because of anticipated contextual bias.

- **Number of sections:** In some cases, whole questionnaire sections are revised or removed to reflect life-course changes, such as the pregnancy, delivery and breast-feeding section of the Round 1 Younger Cohort child survey. In others cases, new sections are added. This can either reflect a shift of emphasis in the conceptual framework of the whole study – such as the addition of a consumption and expenditure section in the Round 2 household questionnaire or the addition of a employment and earnings module since Round 4 for the Older Cohort – or to the need to add country-specific sections, which usually gather data about specific policy initiatives of particular relevance to childhood poverty.
- **Style of questions and answers:** The way that some questions are asked has altered according to what has been learned in previous rounds. For instance, some children were upset by negative questions asked in Round 2, so these were reframed positively in Round 3. Faces illustrating different moods were chosen to supplement words on some Likert-type answer scales in Round 3; these were discontinued in Round 4 on the premise that 12 and 19 year olds were old enough to understand the scales without the help of illustrations, and that the faces were not strictly representative of the answer scales.

There are several overarching challenges involved in designing research protocols for each successive round. These include:

- maintaining a balance between preserving the continuity of core questions for longitudinal purposes and responding to shifts and changes in contextual debates on poverty and development policy.
- ensuring that questions are age-appropriate and adequately reflect variations in outlook, capacities and communication skills of the children in different countries.
- ensuring that each of the three principal components complements the others, and that overlaps between them contribute to triangulation.

- keeping the surveys at a reasonable length, without overburdening respondents.

Developing the Round 1 survey

The research protocols for the Round 1 survey, carried out in 2002, were designed to provide baseline information both for subsequent rounds, and for the detailed thematic components that were included in the original plan for Young Lives. They aimed to produce data that favoured breadth over depth (Attawell 2003).

The process of developing the surveys was informed by a [pilot study in South Africa](#) and by the varied disciplinary perspectives of the study team. A literature review which drew together information about poverty and children from different sources was used to identify key topics for analysis. This led to the prioritisation of six child welfare outcomes: physical health, nutrition, mental health, developmental stage, life skills, and perceptions of well-being. Having identified these key outcomes, flow charts were constructed to elaborate causal pathways and determinants for each outcome at the micro and macro levels. Three key ‘storylines’ – livelihoods, social relations, and access to services – cut across all six flow charts, as well as reflected contemporary development narratives (Attawell 2003). The outcomes and storylines formed the conceptual foundation of the Round 1 survey protocols. Table 1 summarises the content of the child and household questionnaires that emerged from this process.

Devising specific interview questions, well-being measures and child development assessments involved a lengthy process of negotiation and compromise as the enormous number of potential questions was whittled down to those considered essential to provide both adequate breadth and a balance of variables useful for both cross-sectional and longitudinal analysis.

The design of some questions drew heavily from existing instruments. The caregiver mental health questions, for

Table 1. Content of child, household and community questionnaires at Round 1

	Household questionnaire	Child questionnaire	Community questionnaire
Both cohorts	Household composition Caregiver background Child health Household livelihoods Economic changes and events Socio-economic status Social status Child height and weight		Physical environment Social environment Infrastructure and amenities Economy Health and education Prices
Younger Cohort only (age 6 to 18 months)	Pregnancy, delivery and breastfeeding Child care Caregiver mental health	Children too young to answer direct questions	
Older Cohort only (age 7 to 8 years)	Child mental health Child education and daily activities	Perceptions of well-being Social capital School and work Health Literacy, numeracy and child development	

example, were derived from a World Health Organization questionnaire; the questions on work in the daily activities section of the Older Cohort household survey were based on a standardised, tested International Labour Organization survey methodology; and the child development test comprised Raven's Colour Progressive Matrices, a psychometric tool built around a series of visual problems, requiring the child to identify the missing elements in a series of patterns. It was selected as a non-verbal tool that has been widely used in cross-cultural research. Other areas used relatively new conceptual frameworks – such as livelihoods, social capital, vulnerability and coping strategies – where there was less experience of measurement to draw on. Here, researchers relied more heavily on their own expertise and innovation to design simple methods that could be administered as part of a large survey instrument.

Round 2: learning from Round 1 and looking to the future

Research protocol design for Round 2, carried out in 2006, took into account many of the same key considerations that informed the design of Round 1. In addition, it also had to:

- respond to challenges which emerged from using particular questions and methods in Round 1
- respond to findings which emerged from Round 1
- respond to contextual changes in the research and policy arenas
- reflect the life-course stage influencing the two cohorts, now aged around 5 and 12 years old respectively
- reflect differences in policy, culture and research team priorities between the four countries by including more country-specific questions and sections
- take into consideration how the survey would be linked with the first round of the qualitative longitudinal research component, which was planned for the following year.

The design of protocols for Round 2 reflected an enhanced commitment to a strong child focus. The protocol design of the child component was influenced by qualitative researchers with expertise in child development, and the child focus was reflected by interviewing 12-year-old children directly about their own perspectives and aspirations (Johnson 2008). This raised some challenges. As a Young Lives researcher observed: 'potentially the most important issue about conducting research with children as opposed to adults is that there exists an even greater power differential between adult researchers and child participants than between two adults, due to the lesser power and freedoms of children relative to adults in all cultures' (Johnson 2008: 3). This power differential was particularly stark in some areas of the study countries, particularly where children were not familiar with being asked their opinion.

Many of the Younger Cohort children were approaching the age of primary school enrolment, and the challenge was to find tools to understand children's readiness for school, as well as to measure educational achievement for both cohorts. The selection of tools had to find a balance between using validated, standardised psychometric tests necessary to contribute to debates on cognitive development, and

finding measures that could be applied in developing country contexts. After extensive piloting of several cognitive development and achievement tests, the Peabody Picture Vocabulary Test (PPVT) and a Cognitive Developmental Assessment (CDA) developed by the International Association for the Evaluation of Educational Achievement were selected to assess Younger Cohort children's verbal and quantitative abilities. In addition, the PPVT, two reading and writing items from Round 1, and a Mathematics Achievement Test were selected to assess Older Cohort children's verbal and quantitative abilities. All the tests were translated and back-translated for use in different countries.

As well as the development and selection of new tools, changes in content and structure were made in the Round 2 survey.

- A preliminary interview was introduced, partly to arrange the household interview but also to collect key pieces of information that were previously in the main household survey, thus reducing respondent burden.
- More detailed background information was sought on household members, including the highest level of education reached, and details of non-resident biological fathers.
- The livelihood section of the household survey was subdivided into five areas to better reflect the connection between livelihood strategies and asset structure. More indicators were included about who in the household makes decisions about key assets.
- A new section on food and non-food consumption and expenditure was introduced to the household survey to facilitate more extensive analysis of economic relationships, including measuring poverty using consumption-based welfare measures.
- The social capital section of the household survey was revised to include questions on how social relationships are formed, perceived and used, access and lack of access to services and information, and participation in collective action.
- The child health and development section of the household survey added considerable detail by asking for more information about long-term health problems and disabilities, immunisation, use of health services and dietary diversity.
- In the Older Cohort child component, sections were added to find out about parents and household issues, perceptions of wealth, the community and the future, and children's aspirations, feelings and attitudes, including investigation of discrimination, self-esteem and self-efficacy. A single section on school and activities included more detailed questions about time use.
- Caregiver mental health questions were replaced with questions on psychosocial well-being which closely reflected those posed to the Older Cohort children, partly to see whether the feelings and perspectives of caregivers influence the children they care for.
- The community context survey was restructured into three modules. The first collected updated information on community profiles from Round 1. The second collected detailed information on child-specific services, focusing on health, education and child protection. The third was an optional country-specific module.

- The socio-economic status section was maintained without changes. This section provides the data needed to construct wealth indices, which are the main instrument Young Lives uses to measure and compare the relative socio-economic status of the households [in the sample](#). Maintaining this section unchanged was essential to ensure longitudinal consistency.

Round 3: reinforcing the child focus

Carried out in 2009, when the children were around 8 and 15 years old, Round 3 was the first to include child survey items for the Younger Cohort children. The new Younger Cohort child protocol was heavily based on some sections of the Round 2 Older Cohort protocol, with core sections on school and work activities, feelings and attitudes and social networks, skills and support adapted to the interests and capacities of the age group. It also included a simple game designed to assess risk preference, in order to understand whether children who are willing to take risks have better outcomes than other children.

Several other completely new elements were also introduced.

- A self-administered questionnaire for the Older Cohort asked questions in areas that young people may have felt uncomfortable discussing in a face-to-face interview, such as psychological well-being, experiences of violence, intra-household issues, tobacco and alcohol consumption, and sexual and reproductive health.
- A section was added to the household survey to collect data about the health and nutrition of the closest-in-age siblings of Younger Cohort children, in order to better understand intra-household differences and dynamics. In addition, a receptive vocabulary test (PPVT) was administered to these siblings in Ethiopia, Peru, and Vietnam.
- Several new country-specific sections were added to the household component. In Ethiopia, these concerned access to, and perceptions of, credit support and social protection programmes. In India, they focused on a range of programmes aimed at girls, marginalised groups, rural employment and the abolition of child labour. In Peru, respondents were asked extra questions about access to key services and programmes, and children were given an eye test. In Vietnam, these elements concentrated on an education aid programme, health insurance and experiences of extra schooling.
- Round 3 coincided with the [first round of the school component](#), which provides detailed data about the schools attended by a sub-sample of the Young Lives children, expanding the scope for analysis of the impacts of education.

As in Round 2, alterations and amendments were also made to the core sections of the surveys.

- In response to hypothesised links between climate change and migration, and the considerable rates of migration by Young Lives children in Peru, India and Ethiopia, questions were added to the household and child questionnaires in order to better document children's mobility and the temporal character of their migrations.

- In response to feedback from fieldwork teams, the livelihoods section was made more concise. In place of questions on assets and earnings, a simple seed game developed and piloted by the Peruvian team was introduced in Round 3. Respondents listed all their income sources and then distributed 20 seeds across the list to provide an estimate of the relative importance of each. They were then asked in detail about the amount of income they obtain from the largest source, and this was used to estimate the value of other sources according to the distribution of the seeds.
- The social capital section of the household survey was shortened as, while the household was the main source of social capital for the children when they were young, its importance declines as they grow older and establish social networks independent of the household.
- Questions about fast food, physical activity and tobacco use were added to the health section of the household component, which also included a more comprehensive food security model in order to allow the calculation of a food security status for the whole household.
- Various adaptations and translations were made to increase the cultural relevance of PPVT tests.
- The Older Cohort child survey became more detailed to the extent that it was a challenge for researchers to keep the instrument short enough to apply. While more complex questions were asked in some core areas, other questions were dropped if Round 2 analysis showed high levels of non-response or non-applicable answers.
- The community context questionnaire was shortened to fit better with the rest of the Round 3 survey, gathering information on prices and service delivery, and completing and updating the inventory of schools, social protection and education programmes begun in Round 2.

Round 4: focusing on life-course changes

The fourth survey round, carried out in 2013/14 when the Young Lives children were about 12 and 19 years old, looked into maintaining a critical balance between preserving the longitudinal core elements of the questionnaire, keeping inter-cohort comparability for the 12 year olds (in relation to the Older Cohort in Round 2), and capturing new life situations of young adults, some of whom had formed new households, had children, or had completed full-time education. In addition, there was high analytical demand for re-interviewing the siblings of the younger children (initially interviewed in Round 3) which led to the creation of another panel element. Thus, in Round 4, each study country tracked and interviewed these children, which in many instances involved revisiting the households or travelling to other villages.

Changes made to the core sections of the questionnaire were:

- Several questions and sections previously asked in the household questionnaire (and asked to the main caregiver) were moved to the Older Cohort child questionnaire as 19 year olds were best placed to answer questions about their physical and socio-economic well-being and different activities.

- A tracking questionnaire, designed to document the process of finding households and children, was developed further in order to allow us to identify the 'relevant household'. This is the household that best reflects the socio-economic situation of the child and, thus, where the household questionnaire should be administered. This was a challenge in Round 4 given the different living arrangements in which Young Lives children, especially the 19 year olds, could be found.
- A 'mini-community questionnaire' was designed to get basic information about children who migrated to communities outside the existing ones, where we had collected information since Round 2. Each country adopted a different set of criteria for opening a mini-community questionnaire, for example, a minimum number of five children living in the same locality and a travelling distance of more than 8 km from an existing Young Lives community. The questionnaire was a shortened version of the context instruments.
- Mathematics, reading and comprehension tests were adapted to portray different levels of skill for different ages within each country. Different forms of the tests were piloted in urban and rural/private and public schools, and the best performing items were selected for final booklets administered with the child questionnaire.
- PPVT was adapted in India, Ethiopia, and Vietnam by selecting specific items that reflected increasing levels of vocabulary skills in each of the local languages.
- Siblings of the Younger Cohort children – first interviewed in Round 3 – were tracked and interviewed again in Round 4. Younger siblings were administered a sub-set of psychosocial questions, a cognitive test (PPVT in Ethiopia, Peru, and Vietnam, and mathematics in India), and simple anthropometric measurements (weight and height).
- The self-administered questionnaire was reduced drastically and administered only to 19 year olds in Ethiopia and India. Peru administered a more extensive version, but also only to the Older Cohort.

In response to contextual changes in research and policy and life-course changes faced by 12 and 19 year olds, the new elements added to the questionnaires in Round 4 were:

- A new test of cognitive skills was used in Ethiopia and Peru for Younger Cohort children and their siblings. This was a short, computer-based test of the children's executive functioning skills.
- A comprehensive employment module, which gives an overview of all the paid and unpaid activities that the young adult was involved in, was developed to be administered in the Older Cohort individual questionnaire. The module included labour force participation, detailed information about the main work activity and the acquisition of formal and informal training.
- Relationships, marriage and fertility information was also collected in a new section included in the Older Cohort individual questionnaire. This section included details on spouse/partner characteristics, as well as pre-marriage assets (in Ethiopia), and gifts at marriage (in Ethiopia and India). Older Cohort girls and boys were also asked

about fertility expectations and for the whole history of births if they had had children of their own.

- A new section on decision-making, consisting of a list of hypothetical decisions, and a self-evaluation of level of involvement of the Older Cohort child and other household members in the decision-making process of each one of them.
- Movement histories, which included detailed information of all movements outside the locality that lasted more than two months since the last round, were recorded for children of both cohorts (in the household questionnaire for the Younger Cohort and the child questionnaire for the Older Cohort). In addition, migration aspirations and preferences, and personal and financial links with the main caregiver of the child in Round 3 (in cases where the child was no longer living with this person), were recorded for the 19 year olds.

Round 5: focusing on life-course outcomes

Round 5 was carried out in 2016-17, when the two cohorts were 15 and 22 years old. While many of the fundamental design considerations remained the same, the round introduced new areas of questioning that reflected the new life situations of these young people. The survey design for the 15 year olds focused on keeping a balance with previous core sections to ensure comparability, covering key areas that were asked to 15 year olds in 2009, and adding new contextual and policy relevant sections for this critical age. The design for the 22 year olds, in turn, focused on selecting outcome-oriented instruments and others that would reflect their transition into the labour market.

Building on the structure of Round 4, the development of Round 5 involved the following changes to the core sections of the questionnaire:

- Sections related to the child's education, health, and movement history, previously asked in the household questionnaire for the Younger Cohort (usually to the main caregiver) were moved to the child questionnaire as 15 year olds were best placed to answer these questions.
- Cognitive achievement tests were adapted following extensive piloting in the four countries. Mathematics, reading comprehension, and receptive vocabulary (PPVT) tests were administered only to 15 year olds; discontinuing the administration of achievement tests to 22 year olds. Also, following the administration procedure in Round 4, siblings of Younger Cohort young people were tested in receptive vocabulary (in Ethiopia, Peru, and Vietnam), and mathematics (in India).
- Short self-administered questionnaires (SAQs) were introduced for Younger Cohort young people in Ethiopia, India, and Peru; the latter with a slightly more extensive version. The questionnaires consisted of similar questions asked to the Older Cohort at age 15, in 2009. With some small variations across countries, the SAQs included questions on access to contraception and knowledge on sexual and reproductive health. Ethiopia, India, and Peru also continued administering the SAQ to Older Cohort young people.

In addition, the Round 5 survey included new elements for both 15 and 22 year olds:

- To assess gender attitudes among adolescents and young adults, the survey included the Attitudes toward Women Scale for Adolescents (AWSA); a 12-item scale to which individuals are asked to respond on a 4-point Likert scale ranging from strongly agree to strongly disagree. The statements refer to the rights, freedoms, and roles of girls and boys in education, sports, dating, and families, and to adult roles in parenting and housework.
- A new section aimed at identifying different levels of difficulties in functioning in six core domains was included in both the Younger Cohort and Older Cohort individual questionnaires. The set of questions correspond to the Washington Group (WG) Short Set of questions which ask whether people have difficulties performing basic activities such as walking, seeing, hearing, self-caring, communicating, and learning/concentrating.
- A module on access, frequency of use, and level of skills in using digital devices and internet was developed for both 15 and 22 year olds. After careful piloting, it was determined that in Ethiopia and India, this module would include questions about access and frequency of use for the 15 year olds and that only 22 year olds would be further asked about their skills. In Peru and Vietnam, both 15 and 22 year olds were asked about all three domains.
- Building on an existing section on education and job expectations and aspirations, Round 5 also asked about expectations of future earnings (at the age of 25), in relation to different education scenarios, and in relation to the job that the individual expected to have in the future. As the age of Older Cohort individuals at Round 5 (around 22 years old) was too close to the benchmark age of the expectation questions, the administration of the existing module was discontinued for them in this survey round.

- Fifteen year olds were also asked about their perceptions of marriage and parenthood through an extensive section that covered questions on the ideal age at marriage, ideal number of children, ideal birth spacing, etc.

For the 22 year olds, the focus for Round 5 was on their transition to the labour market. The following sections were therefore developed for the individual questionnaire:

- Soft skills for the labour market in terms of leadership ability and cooperative teamwork were included as self-reported instruments. The selected measures are sub-scales of the Review of Personal Effectiveness with Locus of Control (ROPELOC) instrument created to measure individuals' abilities and beliefs. Specifically in the case of the selected sub-scales, their aim is to measure individuals' social abilities.
- Additional personality traits that predict achievement and well-being such as perseverance and self-control – measured by a short Grit Scale – and two sub-scales of the Big-Five Personality Test (i.e. conscientiousness and neuroticism) were included together with the socio-emotional scales (self-efficacy, self-esteem, etc.) included since Round 2.
- A short section on other relevant skills for the labour market, such as knowledge and fluency in different languages, as well as the ability to operate different types of vehicles and machinery.

Finally, given that many Older Cohort individuals have already had children, and that basic information was collected in Round 4, Round 5 included a module to continue collecting information on the development of these children, creating thus a new panel element. For them, and for children born to the 22 year olds in the period between Round 4 and Round 5, information on health (antenatal care, breast feeding, vaccinations, etc.) and education (attendance at nursery, crèche or pre-school) has been extensively collected in Round 5.

REFERENCES

Attawell, K. (2003) *International Longitudinal Research on Childhood Poverty: Practical Guidelines and Lessons Learned from Young Lives*, Working Paper 11, Oxford: Young Lives.

Boyden, J. and P. Dornan (2011) *Putting Children at the Centre of Poverty Debates*, Policy Brief 12, Oxford: Young Lives.

Johnston, J. (2008) *Methods, Tools and Instruments for Use with Children*, Technical Note 11, Oxford: Young Lives.

Justification documents for the Young Lives survey can be found at:

[Young Lives: an International Study of Childhood Poverty: Round 1, 2002](#)

[Young Lives: an International Study of Childhood Poverty: Round 2, 2006](#)

[Young Lives: an International Study of Childhood Poverty: Round 3, 2009](#)

[Young Lives: an International Study of Childhood Poverty: Round 4, 2013-2014](#)

Section 7: Qualitative Longitudinal Research

May 2017

Young Lives conducts in-depth qualitative research focused on a nested sample of the Young Lives children, but also involving their parents/caregivers, peers and other members of their communities. This work consists of two main strands:

- a longitudinal component which tracks 50 children in each study country, documenting their changing life trajectories over time.
- shorter, thematically-focused enquiries on particular topics, such as children's experiences of parental death in Ethiopia and the impact of the National Rural Employment Guarantee Scheme in India.

In contrast with other components of the study, the qualitative research focuses on the practices and experiences which explain the diverging trajectories of Young Lives children, complementing the survey measurement of various aspects of their lives at particular points in time. Its major strength is that it offers an opportunity to include the detailed, narrative perspectives of children and families about their experiences in a way that is embedded in the longitudinal design of Young Lives. This means that the qualitative research can simultaneously illuminate and draw on the principal quantitative elements of the study, particularly the child, household and community surveys and the school surveys. It also strengthens the capacity of the study to situate the children's experiences of poverty in relation to the people around them, and the social, cultural and institutional contexts that shape their lives and opportunities.

The first plans for Young Lives did not include qualitative longitudinal research; the original research consortium conceived the study with a survey-based panel design. Planned thematic studies to explore particular areas in more depth may have included the use of qualitative methods, but this was the full extent of anticipated qualitative enquiry. However, following the Round 1 survey, an external review highlighted the limitations of a mainly quantitative research design and recommended the inclusion of a qualitative longitudinal component. The main funder of Young Lives, the UK Department for International Development, requested that two rounds of qualitative research be carried out within a relatively short time-frame, so that by the time of the Round 4 child and household survey (planned for 2013), the quantitative and qualitative components would be on a comparable footing.

Planning for the qualitative component began in 2006 with a review of qualitative methods for research with children which mirrored the structure of the Round 2 child and

household surveys (Johnson 2008). A pilot of child-centred, qualitative methods was carried out in Peru six months later, working with seven small groups of children, mostly aged 11 to 12, to test group methods for investigating each of the study's key themes. Qualitative research teams – including anthropologists, education specialists, psychologists, social workers and sociologists – were appointed in the study countries early in 2007.

Preparing for the qualitative longitudinal research: design process and sample structure

The process of designing the qualitative longitudinal research component was iterative, reflexive and fully inclusive of country-based qualitative researchers' views and expertise. Using the review of methods and Peru pilot as a starting point, the pilot phase relied on successive stages of design, testing and refinement of methods and questions, as well as incorporating [various aspects of training](#). Pilot studies were carried out in a rural and an urban sentinel site in each country in mid-2007. The pilot phase produced a set of refined research questions for Round 1 of the qualitative research, and a toolkit from which country teams could select methods appropriate to the contexts in which they were working.

While a key aim of the pilot phase was to develop the methods that would be used in the first full round of qualitative research, it was equally important to ensure that the design would anchor the qualitative research to the panel survey sample. This was achieved through using the pilot studies, to validate measures and indicators used in the surveys and to address themes that were emerging from analysis of the survey data, but also through the selection of a sub-sample of children who would make up the 200 case studies at the centre of the longitudinal qualitative research component.

In common with the process of selecting the full survey [sample](#), the first stage of selecting the qualitative sub-sample was the choice of field sites. This sought to enable the exploration of variations in location, ethnicity and socio-economic status, so in each country sites were selected:

- from different regions, reflecting the main ethnic or caste groups in the country
- to reflect an equal balance between rural and urban sites
- to reflect an equal balance between sites that had been classified as poor and those classified as less poor.

In India, Peru and Vietnam this resulted in the selection of four sites, but in Ethiopia five were needed to reflect the ethnic diversity of the country and the full sample. In each site, equal numbers of boys and girls were selected from the Younger and Older Cohorts. As well as gender and age, the sub-sample children were chosen according to the socio-economic status and structure of their households, and their attendance at pre-school. Potential replacement children were also identified, who could be included in the sub-sample if selected children were not available either because their families had moved or because the children or their caregivers did not grant permission or time for interviews.

Qualitative methods for longitudinal research

The overarching question which frames the qualitative longitudinal research component is ‘how does poverty interact with other factors at individual, household, community and intergenerational levels to shape children’s life trajectories over time?’ The qualitative research is explicitly based on the premise that children’s experiences and perceptions are a major resource for providing answers to this question, and the view that children are social actors who offer valid and useful insights and understandings about their own lives.

The methods used are designed to allow and encourage children of different ages to communicate their points of view about the key themes of the study in a way that also allows systematic recording and analysis. The methodology is flexible enough to adapt to different settings and to thematic emphases that vary from country to country, and adequately reflexive to incorporate learning from successive research rounds.

The review of age-specific tools and instruments for use with children that was carried out at the start of the pilot phase gathered information about a wide variety of possible methods. Criteria were developed to choose methods that best met the needs of the study. Selected methods had to be:

- semi-structured, to ensure that core themes could be studied consistently
- applicable in diverse cultural settings
- implementable by fieldworkers with very variable research training, orientation and experience
- flexible enough to allow children to identify themes and issues that are important to them
- able to generate data that can be relatively simply and efficiently recorded, in contrast to recording full focus group transcripts, which can be very costly in time and money
- adaptable, given the variation in educational levels and preferred methods of communicating among the Young Lives children.

The methods selected were tested in the country pilots, adapted and refined, and a final selection was included in the methods toolkit for the first round of the qualitative research (Camfield, Crivello and Woodhead 2013a). Combined with research questions on each of the key study themes – children’s time-use, well-being, poverty and social worlds – this gave each country team a menu of methods that could be applied to fieldwork, rather than a fixed list of tools that had to be used to research each theme. On the menu were individual methods, group methods and observation.

- The basic method for engaging with **individuals** was the semi-structured interview. For each interview, a checklist was drawn up of questions relevant to the respondent, whether child, caregiver or community member. With children, interviewing was often combined with an optional range of more dynamic and visual tools and exercises including games, life-course drawings and social network maps. Interviews with children were also often carried out over more than one session to prevent them becoming tired or bored.
- Tools for working with **groups** were clustered around each key theme. These included drawings, a time-use bucket activity, activity worksheets, community mapping and guided tours for understanding children’s time use; social network maps and a story completion exercise for examining children’s social worlds; and a body map, an indicators of well-being exercise, and a poverty tree diagram for looking at poverty and well-being.
- **Observation** techniques add to the picture of what people say they do with their time. Researchers in the first round spent time sharing children’s daily activities as well as time in the community making observations of the wider environment.

Different tools were selected to gather a range of information which was used to build up a ‘mosaic’ of children’s lived experiences (Clark and Moss 2001). This composite image could then be supplemented and supported by existing survey data.

While each country team had a common point of departure – the collectively agreed set of methods and research questions – the combination and type of methods they used varied according to different considerations. These included country focus on particular research themes, the findings of the pilot studies and the need to use the methods in sequences that would progressively build up relationships of trust and empathy between children and researchers (Ames 2011).

A similar balance between collective frameworks and protocols and country priorities was taken in approaching data analysis. Systematic protocols for data transcription and management were developed centrally and followed by all teams. A meta-framework for coding the data was developed based on the key research themes, to allow for consistency and comparability [across countries and between rounds](#). Country teams, however, elaborated on the coding framework by constructing more detailed codes and conducting data analysis according to their specific research focus.

New tools have been introduced in each successive round of qualitative research, and existing ones revised and refined. Updating the toolkit to ensure that it matches evolving research questions is a key aspect of [piloting and training in every round](#).

Our approach to ethics has been developed collaboratively with our research teams, following fieldworker training, piloting and reports from fieldworkers after each round of visits to our study sites.

Evolving research questions

The questions which frame the qualitative longitudinal component are closely connected to the key themes and preoccupations of the survey components. In the first round, the enquiry was structured around the three thematic areas which were derived from the overall emphasis of Young Lives. The framing questions derived from these themes for the first round were:

- What are the key **transitions** in children's lives, how are they experienced, and what influences these experiences?
- How is children's **well-being** understood and evaluated by children, caregivers and other stakeholders?
- How do policies, programmes and **services** shape children's transition and well-being?

The first round of qualitative research in 2007 provided baseline information on these three areas and complemented existing household and community-level data on shocks, coping and services (Camfield, Crivello and Woodhead 2013a). The second round, in 2008, followed the same lines of enquiry to begin to document changes in each area. New methods were created and adapted to approach the questions with children who were now a year older (Camfield, Crivello and Woodhead 2013b).

The third round, in 2011, provided another opportunity to document changes, but this time a stronger emphasis was placed on collecting information to understand factors at the household and community level that contribute to diverging life trajectories, and the extent to which children are involved in making the key decisions that affect their lives (Crivello, Morrow and Streuli 2013).

Key questions included:

- What shape have children's life **trajectories** taken, and what are the processes explaining these patterns, including factors related to poverty, intergenerational change or difference?
- What have been the major **transitions** influencing changes in children's life trajectories? How have these been experienced by children and families, including children's own roles in decision-making?

- What have been the main sources of **support** and **risk** to children's transitions?
- How have different transitions shaped children's life trajectories? Have these opened up or constrained **opportunities** for children's present and future lives?

In addition to these questions, three areas of children's life trajectories were prioritised in order to respond to the wider policy and research priorities of Young Lives. At the level of individual case study children, these were changes and continuities in:

- **school trajectories:** transitions through grades, classrooms, institutions, schooling types and locations, including the transition out of school.
- **work trajectories:** capturing the variety of paid and unpaid contributions young people make to their families and to themselves through work, how this supports or interferes with their schooling and with their evolving social identities.
- **social trajectories:** young people's integration into households and communities through their changing roles, responsibilities and identities.

A fourth round of qualitative longitudinal research was undertaken in 2014, building on the findings from the first four survey and three qualitative rounds (Crivello and Wilson 2016).

Key challenges

Key challenges in the ongoing process of designing and adapting the qualitative longitudinal research include:

- developing methods suited to different cultural contexts, ages and experiences within the same study
- developing methods that are cognisant of power asymmetries between adults and children, in particular children who are shy or reserved due to social exclusion
- encouraging creativity and flexibility in adapting questions and methods to local contexts, while also ensuring comparability of datasets
- a relative lack of tradition in conducting research with children that prioritises their views and accounts of their everyday lives, in cultures where children may not be expected to speak up
- language – not simply translation between languages, but the need to find and agree on words that reflect the concepts in the research questions, but also allow comparison between countries
- maintaining and managing a large, multilingual, multimedia, qualitative longitudinal dataset

- [migration of older children away](#) from research sites for work or education
- developing robust methods for analysing qualitative longitudinal data, where few published resources exist
- making best use of quantitative data in integrated analyses and developing qualitative researchers' confidence and skills in this area
- ensuring that policy concerns and priorities are represented in research planning and execution, and data analysis
- coordinating a conceptually and methodologically complex study using country-based research teams with different disciplinary backgrounds and levels of experience
- balancing the needs and timelines of other project components.

REFERENCES

Camfield, L., G. Crivello and M. Woodhead (2013a) *Young Lives. Qualitative Fieldwork Guide: Round One (2007)*, Technical Note 27, Oxford: Young Lives.

Camfield, L., G. Crivello and M. Woodhead (2013b) *Young Lives. Qualitative Fieldwork Guide: Round Two (2008)*, Technical Note 28, Oxford: Young Lives.

Crivello, G. and E. Wilson (2016) *Young Lives. Qualitative Fieldwork Guide: Round Four (2013/14)*, Technical Note 34, Oxford: Young Lives.

Crivello, G. (2015) 'Qualitative Longitudinal Research with Children and Young People,' in *Handbook of Geographies of Children and Young People*, edited by Tracy Skelton, London: Springer.

Crivello, G., V. Morrow and N. Streuli (2013) *Young Lives. Qualitative Fieldwork Guide: Round Three (2010/11)*, Technical Note 29, Oxford: Young Lives.

Crivello, G., V. Morrow and E. Wilson (2013) *Young Lives. Longitudinal Qualitative Research: A Guide for Researchers*, Technical Note 26, Oxford: Young Lives.

Johnston, J. (2008) *Children's Perspectives on their Young Lives: Report on Methods for Sub-Studies Peru Pilot*, April 2006, Technical Note 10, Oxford: Young Lives.

Morrow, V., and G. Crivello. 2015. 'What is the value of qualitative longitudinal research with children and young people for international development', *International Journal of Social Research Methodology* 18.3: 267-280.

Section 8: School Surveys

May 2017

Young Lives has been conducting school surveys across the four study countries since 2010. Recognising the increasingly central role that schooling plays in children's lives, these surveys have sought to improve our understanding of the impact of diverse school environments on children's lives and outcomes. With the addition of school-level data, Young Lives has become one of only a handful of surveys which gathers detailed longitudinal information not only about children and their households and communities, but also about their schools, providing a unique resource for research and policy.

The first round of school surveys took place in all four countries between 2010-13, and focused on school effectiveness at the primary level. Additional secondary school surveys are taking place in 2016-17 in all four countries. Survey design has evolved over time and varies between countries, to take account of important contextual differences, but the research is framed by two broad questions:

- How do the relationships between poverty and child development manifest themselves in and impact upon children's educational experiences and outcomes?
- To what extent does educational experience reinforce or compensate for disadvantage in terms of child development and multidimensional poverty?

A common framework and contextual diversity

In order to take into account the range of contexts and key issues in education policy, the exact content and design of the school surveys differs slightly in each study country. However, all school survey instruments share a common framework, intended to capture the children's experiences of education in a way that allows statistical analysis but is also flexible enough to adjust to contextual diversity. In all countries and at both the primary and secondary levels, the school surveys examine the same core dimensions of education and educational experience:

- educational access and progression
- quality of education
- equity in education
- the role of education in the policy cycle.

Each school survey relies on this common framework and comprises a set of country-specific research questions which are guided by the broad themes of the research. Variations between countries reflect not only an understanding of

which areas the school survey is best placed to investigate given the other data being collected by Young Lives, but also a careful analysis of the policy context in each country. Consultation and discussion with policy actors from the survey design stage through to the dissemination and discussion of results of analysis has been an important element of the research in each country.

Sampling and design is a second key area of variation between countries. The first primary surveys in India and Ethiopia went to the schools of a selection of Young Lives children, adding school-level data to the household data. Later primary school surveys and the forthcoming secondary surveys have moved away from this approach to focus on answering questions of school effectiveness. This has involved the selection of schools in the Young Lives sites, and the sampling of both Young Lives children and their peers, often in a specific grade or class in those schools.

These surveys have also involved a 'test and retest' design in which survey fieldwork has taken place at both the beginning and end of the school year, to enable analysis of the determinants of pupil progress in a single academic year. In each instance, the aim has been to find a balance between capturing as many Young Lives children as possible and producing a balanced school and class level sample for the purposes of conducting research on school effectiveness. Achieving this balance depends partly on logistical and resource limitations, and is also shaped by the same contextual policy analysis that underpins the country-specific research questions.

In Vietnam, for example, where initial access to primary education is now nearly universal, there is an emerging emphasis on the question of primary school completion. The attention of policymakers has also begun to turn towards quality and equity, especially because in respect of children's achievement, regional variation and variation based on household characteristics is very wide, with considerable advantage being afforded by urban location and to the ethnic majority Kinh. The Vietnam school survey in 2011-12 therefore focused on the Younger Cohort children, whose age was ideal to give an understanding of what happens during the latter stages of primary education and the implications of this for later transitions and outcomes. The survey focused on those Young Lives children in Grade 5 at the time of fieldwork, and each school attended by one or more of these children was selected. A sample was also taken of class peers of Young Lives children studying in these schools, allowing for a more thorough examination of the variation between schools and classes.

Data collection methods

Each country survey uses a combination of up to eight separate instruments to collect school data, with the exact mix and design varying by country and survey. These include:

- **Principal questionnaire:** personal details and attitudes of the school principal; general information about the school, including governance and links to higher authorities and the community; school-level information about pupils and teachers.
- **Teacher questionnaire:** personal details and attitudes of a teacher of the surveyed children; general information about the teaching methods and classroom procedures for the classes in which the surveyed children are studying.
- **Child questionnaire:** personal details; attitudes and opinions about school.
- **Child tests:** in maths and reading comprehension (at primary level) and maths, functional English and transferable skills (at secondary level)
- **Teacher methods assessments:** giving examples of mistakes made by children when doing maths problems, and asking teachers to explain how they would correct the errors.
- **School observation:** looking at the infrastructure, facilities and management of the school and the availability of resources such as textbooks and toilet facilities.
- **Child observation:** observing behaviour and environment in the classroom, and assessing homework books.
- **Teacher observation:** observing teaching methods and language.

In addition to the eight basic instruments of the school component, [qualitative sub-studies](#) have been carried out in some locations focused on specific research questions of direct policy interest, which arise out of the main dataset.

Challenges

- In common with the broader Young Lives study, there is a tension in the school component between its dual purpose as a public good providing a repository of knowledge on a broad theme and as an analytic study. This tension gives rise to challenges of how to define questions and in what sequence to use different tools.
- Also in common with the broader study, the school component is producing very large datasets, which can present challenges for systematic analysis and prioritising and sequencing analytic questions.
- The different approach taken to sampling in each country means that particular care is needed in discussing exactly what the data represent. Matching data between components of the Young Lives study is made more challenging by the use of slightly [different approaches to sampling](#) at different stages of the main survey, and this demands constant vigilance.
- Consulting key actors in education policy at the design stage of the school component may have created expectations that the data it produces will answer very specific policy questions, which may not be the case. The pre-existing Young Lives sample is the foundation of the school component sample and data, and this precludes direct nationally representative measurement of the impact of particular policies, teaching methodologies or approaches.
- Some of the pupil tests used in the school component have required considerable adaptation to make them culturally appropriate to the range of contexts.

REFERENCE

Rolleston, C., Z. James, L. Pasquier-Dourmer and Tran Ngo Thi Min Tam (2013) *Making Progress: Report of the Young Lives School Survey in Vietnam*, Working Paper 100. Oxford: Young Lives.

Section 9: Sub-studies

May 2017

Alongside the major survey and qualitative longitudinal components of Young Lives, several additional studies have been undertaken that focus on particular themes in selected countries. While sub-studies carry out new fieldwork designed to address their thematic focus, they also have the advantage of being able to rely on existing survey and qualitative data to shape their research design and contextualise their findings. This section briefly describes several Young Lives sub-studies, and lists the publications associated with them. In most cases, the same fieldwork teams undertook the research who had undertaken previous qualitative longitudinal research, to build on existing relationships.

National Rural Employment Guarantee Scheme (NREGS) in Andhra Pradesh, India

This research in 2009 involved 30 Young Lives households. It aimed to explore young people's experiences of NREGS, and whether the scheme is sustainable in the light of problems it may be causing within communities.

The fieldwork took place in three rural Young Lives sites and included both families who were participating in NREGS and some who were not. Fieldwork was conducted with Young Lives qualitative children and other Young Lives children who were purposively sampled to (i) ensure that the areas of interest were covered and (ii) include children who could be followed over the next qualitative rounds. Interviews and group discussions took place with key informants (for example, Panchayat officials, employers), male and female caregivers, and around 12 children (Younger Cohort children and older siblings) in each site over 10 days. Researchers used semi-structured interviews designed in conjunction with the survey team to cover similar areas to those addressed in the Round 3 survey.

- Camfield, L., and U. Vennam (2012) *From Policy to Implementation: An In-depth Exploration of Mahatma Gandhi National Rural Employment Guarantee Scheme in Andhra Pradesh*, Working Paper 82, Oxford: Young Lives.

Productive Safety Net Project (PSNP) in Ethiopia

This research in 2009 aimed to explore the effects of PSNP on children's well-being, the impacts of local and household economic shocks, implementation of PSNP, and the possible

impacts of PSNP on schooling, work, health and food consumption.

The fieldwork took place in four rural Young Lives sites and included some families who are participating in PSNP and some who are not. Fieldwork was conducted with Young Lives qualitative children and other Young Lives children who were purposively sampled to (i) ensure that the areas of interest were covered and (ii) include children who could be followed over the next qualitative rounds. Semi-structured interviews and group discussions were held with key informants (for example, local officials), male and female caregivers, and around 12 children (Younger Cohort children and older siblings) in each site over 10 days.

- Tafere, Y., and T. Woldehanna (2012) *Beyond Food Security: Transforming the Productive Safety Net Programme in Ethiopia for the Well-being of Children*, Working Paper 83, Oxford: Young Lives.

Risk, vulnerability and resilience in Ethiopia and Andhra Pradesh, India

This sub-study, conducted between 2009 and 2011, aimed to:

- explore the challenges of translating research into policy and practice on child protection and child poverty.
- carry out research on locally relevant aspects of how risk, vulnerability and resilience affect children's lives.

To meet these goals, researchers designed reflective, multi-stakeholder processes to identify barriers and explore challenges to translating research into practice. The study team began by using Young Lives findings and their contextual knowledge of each country to select a subject within the broad theme of risk, vulnerability and resilience which had particular relevance to policies for protecting children and alleviating child poverty. In Ethiopia, where one in five of the Young Lives children has lost one or both parents by age 12, orphanhood and vulnerability were the focus of research, while in Andhra Pradesh, child labour was selected, as it is a common phenomenon and a highly politicised matter.

Researchers, policymakers and practitioners were then brought together to reflect on the challenges and opportunities for improving the use of research in policy and practice in these two areas. This reflection served both to challenge prevailing assumptions about what puts children at risk, and to identify priority research questions for the fieldwork. The design of the sub-study differed in each

country according to variations in stakeholder priorities and prevailing policy processes.

In Ethiopia, fieldwork was carried out in 2010 in three Young Lives sites and focused on a sub-sample of children from both Younger and Older Cohorts, around half of whom had experienced the death of a parent. One hundred individual interviews and six group discussions were carried out with the children, their caregivers and community members, and researchers also drew on existing Young Lives data about the children. Data were also collected through observation and life histories.

In India, the sub-study focused on children involved in agricultural work in Andhra Pradesh (in 2011). The study aimed to explore understandings of risks at work (and in daily life) with a view to contributing to debates about what constitutes hazardous and non-hazardous work from children's perspectives. The fieldwork took place in two Young Lives sites, interviewing 44 children who were around 14-15 years old – 26 Young Lives children and 18 non-Young Lives children. The non-Young Lives children were included to ensure that the sample was large enough and that it included examples of children who worked full-time as well as those who combined work with school. Interviews and group discussions took place with around 20 children per site over one week and key informants (for example, *sarpanch*, health workers). Data were also collected through body mapping.

Findings from both elements challenged policymakers, especially within the field of child protection, to move beyond a focus on categories of children at risk, such as 'orphans' and 'child labourers', and to encourage policies aimed at addressing the underlying causes of poverty and inequality which put children at risk of poorer outcomes and hazardous work situations.

Lessons from the consultative processes illuminated the reasons why translating such findings into changes in policy and practice remains difficult. Principal challenges, summarised by the stakeholder group in Ethiopia, included:

- A lack of dialogue and consultation between researchers, practitioners and policymakers.
- Research is often not oriented towards action.
- Research results are frequently not adequately disseminated or easily found, and findings seldom translated into concrete recommendations.
- Donors rarely require projects to be based on research, while funding for research rarely specifies the need for consultation with other stakeholders.

One outcome has been the establishment of the [Child Research and Practice Forum](#) (CRPF) in Addis Ababa. The CRPF brings together researchers, practitioners and policymakers working on child poverty regularly to share current research and learning at the Ministry of Women and Children's Affairs, and build local capacity for using and engaging with research. Activities include monthly meetings, seminars, published summaries of presentations and a regular newsletter. The CRPF hosts events and consultations to facilitate evidence-based debate on topics such as harmful

traditional practices, child migration and child work. Its success indicates how valuable such a forum can be.

- Crivello, G. and N. Chuta (2012) 'Rethinking Orphanhood and Vulnerability in Ethiopia', *Development in Practice* 22.4: 536-48.
- Crivello, G., and H. Murry (2012) [Why strengthening the linkages between policy and research is important](#), Policy Brief 19, Oxford: Young Lives.
- Morrow, V., and U. Vennam (2012) 'How Children Manage Risk in Daily Life: Identifying Hazards in Agricultural Work in Andhra Pradesh, India', *Development in Practice* 22.4: 549-61.

Parents' views of quality education, India

This was a sub-study into factors affecting school choice in Andhra Pradesh in 2011. It involved interviewing a sub-sample of Young Lives children (aged 9-10) and the key household decision-maker in three Young Lives sites. The sites and children were purposively selected from the Young Lives school component sample because they had demonstrated high rates of children moving between government and low-fee private schools between Round 3 and the school survey. The in-depth semi-structured interviews explored perspectives on education, perceptions of differences between specific private, government and other schools in each community, and detailed information on schooling decisions. The aim was to build a multi-layered picture of the forces driving, and the impacts of, differentiated and changing school choices.

- James, Z., and M. Woodhead (2014) 'Choosing and changing schools in India's private and government sectors: Young Lives evidence from Andhra Pradesh', *Oxford Review of Education* 40.1: 73-90.
- Vennam, U., A. Komanduri and U. Duggani (2014) [Changing Schools in Andhra Pradesh: The Experiences of Children and their Caregivers](#), Working Paper 131, Oxford: Young Lives.
- Morrow V., and E. Wilson (2013) *Parents' perspectives on quality of schooling in Andhra Pradesh India*. London: Save the Children.

Beyond urban relocation: expectations and concerns of children and caregivers in Addis Ababa and Hawasa, Ethiopia

Carried out in January 2012, this relocation study comprised qualitative and quantitative components. A survey was conducted in four urban sites where slums were targeted for clearance and redevelopment. The survey included 466 caregivers and 451 children. Of the 451 children, two-thirds were from the Younger Cohort (aged 11 and 12 years old), and a third were from the Older Cohort (aged 17 and 18). In the qualitative component, interviews were conducted with 79 children and their caregivers, 10 boys and 10 girls in each of the four sites.

Forty-five key informant interviews were conducted with people from formal and customary institutions in each community. Group discussions were also conducted with children, with their caregivers, and with influential community members. Three Young Lives families who had moved to live in condominiums were contacted for in-depth interviews to obtain their views about changes in their lives resulting from the move. Three briefs translated into Amharic were produced (one per working paper).

- Pankhurst, A., and A. Tiemelissan (2013) *Living in Urban Areas due for Redevelopment: Views of Children and their Families in Addis Ababa and Hawassa*, Working Paper 105, Oxford: Young Lives.
- Pankhurst, A., and A. Tiemelissan (2013) *Moving to Condominium Housing? Views about the Prospect among Children and their Families in Addis Ababa and Hawassa*, Working Paper 106, Oxford: Young Lives.
- Pankhurst, A., and A. Tiemelissan (2014) *Beyond Urban Relocation? Expectations and Concerns of Children and Caregivers in Addis Ababa and Hawassa*, Working Paper 126, Oxford: Young Lives.

Stimulating evidence-based approaches to child work/labour in Ethiopia

This was a sub-study of children's work in 2013 in three Young Lives sites (two urban, one rural). The study aimed to deepen understanding of the socio-biographical factors motivating children's pathways into and through work, and to record children's perspectives on the risks and benefits associated with their work. A team of experienced Ethiopian fieldworkers conducted group discussions and interviews with 88 children and young people aged between 9 and 19 years (from within and outside the Young Lives sample), evenly distributed on gender lines, but slightly weighted towards younger children who were below the legal minimum age of 14 years for waged work. We also interviewed a total of 23 caregivers and 45 key informants in the three sites.

The fieldwork was preceded by three consultations (two at regional level and one at national level) with stakeholders working on child poverty and well-being in Ethiopia in order to investigate a range of children's work activities and histories. The results were presented at regional and national consultations in Ethiopia.

- Pankhurst, A., M. Bourdillon and G. Crivello (2015) (eds.) *Children's Work and Labour in East Africa: Social Context and Implications for Policy*. Addis Ababa: Organisation for Social Science Research in Eastern and Southern Africa.

- Pankhurst, A., G. Crivello and A. Tiemelissan (2016) *Children's Work in Family and Community Contexts: Examples from Young Lives Ethiopia*, Working Paper 147, Oxford: Young Lives.
- Pankhurst, A., G. Crivello and A. Tiemelissan (2015) *Work in children's lives in Ethiopia: examples from Young Lives communities*, Addis Ababa: Young Lives.
- Pells, K. (2014) *Children's Well-Being and Work in Sub-Saharan Africa*, Young Lives, Save the Children and the African Child Policy Forum Policy Brief, Addis Ababa: Young Lives
- Tafere, Y., and A. Pankhurst (2016) 'Children combining school and work in Ethiopian communities,' in A. Pankhurst, M. Bourdillon and G. Crivello (2015).

Child and early marriage in Ethiopia

This was qualitative research conducted in December 2015, that aimed to explore how young people and families negotiate the move from childhood to social adulthood in contexts of poverty; the changing role of marriage and parenthood; the benefits, risks, and harms associated with pathways to marriage and parenthood among children and young people; what shapes the quality of married life and young women's status within it; and whether the type and timing/age of marriage matter for young women's life chances.

The research involved 99 interviews, with 40 young people, 37 caregivers, 15 husbands, two wives, and five key informants.

- Crivello, G. (2016) 'Girls' diverging pathways to marriage', Oxford: University of Oxford.
- Chuta, N. (2017) *Young Women's Household Bargaining Power in Marriage and Parenthood in Ethiopia*, Working Paper 166, Oxford: Young Lives.

Child and early marriage in India

This is ongoing qualitative research on decision-making and early childbearing involving the participation of married and unmarried girls (aged 15-22), couples, families and community members in four Young Lives sites in Andhra Pradesh and Telangana, India. The fieldwork was undertaken in late 2016. The study focus is on those girls and young women who married below the age of 18 since this category is of special interest to policymakers in the country and globally. The study aims to generate new knowledge about adolescent sexual and reproductive health, household dynamics of young couples, their fertility decisions, and their experiences of parenthood. The findings will be communicated to relevant stakeholders who influence policies and programmes in India.

Section 10: Piloting – Testing Instruments and Training Field Teams

May 2017

The complex, cyclic and evolving research design of the Young Lives study provides the tools and structures needed to collect qualitative and quantitative data at various levels and from various respondents, and to store and maintain it in a format suitable for longitudinal analysis. This is achieved using a diverse set of instruments, which has changed and expanded with each round as the children grow up and research priorities develop and shift. Piloting tests these instruments and begins training staff to use them.

All instruments in Young Lives are piloted, from panel questionnaires to qualitative toolkits. Piloting aims to:

- ensure that research questions work in the field and are consistent with local situations
- ensure that methods are appropriate to the changing capacities of the children as they develop
- train field teams and learn from their practical experience of fieldwork to improve instrument design
- produce accurate instrument manuals and protocols
- identify and begin to strengthen the skills field teams will need to apply the instrument
- initiate, build and maintain positive team dynamics and mutual respect
- ensure that data collection systems are in place.

Piloting forms part of a wider set of training activities for fieldwork teams. Each country has a committed, long-term study team which includes researchers from different disciplines, survey and data managers, and field supervisors. Fieldworkers are temporarily employed for particular instruments or sub-studies. Although in many cases fieldworkers have returned to carry out successive rounds of the study, the shifting composition of field teams increases the importance of training and teambuilding as ongoing processes. In this way the skills of team members are frequently upgraded through training on each successive instrument or sub-study.

Piloting exercises in Young Lives have varied in scale, format and frequency.

- A year-long pilot stage (2001–02) tested the study's original research design in South Africa, a country outside the study sample.
- The child and household surveys are piloted in non-Young Lives sites in all four countries before each full round of data collection.

- Electronic data collection was piloted initially with the Round 3 survey to assess the possibilities for rolling it out in the following rounds.
- Qualitative research methods were developed collaboratively through piloting and training, and there have been pilots before each subsequent full round of data collection.

Each of these exercises is an essential part of ensuring that all necessary information will be captured and processed effectively during fieldwork. Each has produced lessons to improve subsequent research rounds. The examination below highlights some of the objectives and challenges of piloting different instruments at various stages in a longitudinal study.

Piloting a longitudinal study of child poverty

Young Lives was originally conceived as a longitudinal, survey-based panel study. In 2001–02, over a year, the entire study design was piloted in South Africa (Seager and de Wet 2003). Research questions were selected and survey instruments developed. A training programme for fieldworkers was designed and delivered, and the questionnaire revised and reviewed. A data management system was established, and preliminary analyses carried out. At the end of the year, there was a follow-up round of visits to respondents.

The pilot phase aimed to produce generic research instruments which were intended to form the core of Round 1 of the child and household survey in the four study countries. The child and household questionnaire was the main instrument developed during this phase. The most important objective was to include key measures of outcome variables, such as children's health, nutritional status and cognitive development, and the factors likely to affect them. A multidisciplinary team including epidemiologists, anthropologists, social scientists, statisticians, economists, and child rights and welfare specialists worked on formulating questions with a strong theoretical basis that would also be understood in practice. As the aim was to develop a sound and reliable survey, the challenge was to simultaneously include multiple perspectives on child poverty while also keeping the questionnaire short enough to use effectively in the field.

Full documentation for the pilot questionnaire survey was prepared, including a survey manual, an interviewer manual

and justification documents for questions. Even at this early stage, the culture-specific nature of some questions, terms and variables was flagged as an area of potential difficulty in carrying out the study and a potential limitation for comparative data analysis.

The survey was presented to fieldwork teams as work in progress to encourage discussion and feedback. Fieldworker training, which included role-play exercises and practice sessions, generated debates about the precise meaning of complex questions and translation into local languages. The survey was piloted with children in rural and urban sites. Learning from the experiences of using the questionnaire and feedback from fieldworkers led to further revisions, including reordering questions into more logical sequences, redesigning the layout, highlighting the skip patterns more clearly, and amending questions that the fieldworkers found confusing or culturally inappropriate.

When the pilot survey sites were revisited a year later, an important lesson was that the much higher mobility of the urban population made tracking in urban areas more difficult, and that having contacts outside the child's household was essential for effective tracking.

Piloting Rounds 2-5

It was anticipated from the outset of the study that every Young Lives questionnaire would consist of core and country-specific elements. In addition, in each successive survey round the [questionnaire has also included new questions, and in some cases new sections](#), partly as a response to the changing circumstances and capacities of the children as they grow older.

After Round 1 had established a baseline of core panel data, significant changes and additions were made to the questionnaire for Round 2. For most country teams, this meant splitting the piloting into three or four phases. In Ethiopia, for example, the household questionnaire was piloted first, then the community questionnaire, then the child development section. This phasing allowed staff from different disciplines to be brought in at each stage, using their expertise to develop and test the validity of each section.

Each team also took different approaches to training fieldworkers in applying the questionnaire during the pilot phase. The Peru team, for example, paid particular attention to selecting and training supervisors, working with psychologists to develop training that involved decision-making games and role playing. Experience also showed that significant time should be dedicated to training fieldworkers in the challenging areas of applying cognitive skills tests and completing the income and consumption sections of the household questionnaire.

Having followed slightly different pathways for testing and adapting the Round 2 survey, all four teams carried out a two-week pilot study with the full questionnaire. These pilots followed a rolling schedule, so that each could be attended by a two-person team from the Young Lives UK office which

travelled from country to country. The aim of having some staff working on all four pilots was to ensure continuity of information reported back within and between countries.

While the main aim of the two-week pilot was to test the whole questionnaire with fieldwork staff, it was also an important opportunity to ensure that fieldwork supervisors were provided with adequate information and skills to deliver similar training to their fieldworkers.

The piloting process familiarised supervisors and teams with the study, instruments, and manuals before going to the field. It also allowed spotting any mistakes, checking if questions made sense, and monitoring the skills of supervisors while giving them practical experience in the kind of challenges fieldworkers may encounter. Lessons learned included:

- Feedback sessions are essential to maximise learning from pilot studies, but may need to be carefully facilitated to allow all feedback to be presented.
- The presence of senior research staff in the field helps both team building and capacity building.
- Making sure administrative and logistical tasks are carried out in advance means better use can be made of time spent in the field.
- Managing the changes to the questionnaire that arise from pilot studies needs careful coordination with other processes of translating, editing and formatting, and clear cut-off points beyond which no further changes can be made.

Piloting the Round 3 survey took a slightly different approach, not least because electronic data collection tools were being tested at the same time. Like its Round 2 predecessor, the pilot of the whole Round 3 survey incorporated fieldworker training, but this time there were two periods in the field. As before, it was preceded by a range of testing and training processes as different country teams got to grips with new tools, which included a self-administered questionnaire with its own survey manual, and cognitive tests for siblings.

A challenge in this pilot round was that with electronic data collection, small last-minute changes to questionnaire content created large programming changes which were more difficult to accommodate late in the process. The lesson learned from this was that plenty of time should be allowed between piloting and fieldwork to allow making changes to the programmes and revisions to other important documentation (such as manuals).

Building on the experience of previous rounds, in Round 4 and Round 5, piloting and training followed a three-phase approach. In the first phase, country workshops were conducted with field supervisors and members of the UK team to introduce the new modules (designed by country teams' researchers and other external specialists) and review the whole questionnaire. In reviewing the new modules, the main focus was on discussing their cultural appropriateness and finding better ways of phrasing new questions (including checking they were properly

translated). As part of this first phase, an initial pilot (called pre-pilot) was conducted at:

- the household and individual level (with 12 and 19 year olds in Round 4, and 15 and 22 year olds in Round 5), for household and individual-specific modules.
- the school (for the Younger Cohort) and university/college level (for the Older Cohort), for individual-specific modules that required larger samples to test their validity and reliability (such as cognitive tests and psychosocial scales).

The pre-pilot was carried out using paper questionnaires (previously amended based on feedback collected during the country workshop) since the main objective was to time the retained (core) sections and, for new modules, to test their content in terms of cultural appropriateness and clarity.

The second phase of training and piloting was planned to coincide with the training of trainers (ToT). At this point, the trainers – who are in most cases the field supervisors – reviewed each question of the final drafts of the questionnaires (prepared on the basis of the pre-pilot information) both in paper and in computer-assisted personal interviewing (CAPI), with the aim of mastering the content and checking the flow. Draft versions of the fieldworker manuals were reviewed in tandem so that they were amended with examples, or complemented with further information. The logistics of managing and sharing the data during fieldwork were also designed and established. A second pilot followed the ToT, with the sole purpose of checking the CAPI programmes. Feedback from both training and piloting at this point is crucial to make sure that the programmes were amended in time for the training of fieldworkers.

The third phase coincided with the training of fieldworkers. At this stage, supported by the fieldworker manuals, trainers went over the entire questionnaire with the fieldworkers, explaining the objective of each section and question. This training was complemented with CAPI practice sessions (fieldworkers interviewing one another) to make sure fieldworkers mastered the use of CAPI and the devices. In addition, every two weeks, fieldworkers piloted the questionnaires in CAPI by interviewing young people of the right age and their families in non-Young Lives sites on sections that were previously covered in their training sessions. At this point, feedback was reported back to the data managers in the countries and the UK office to finalise the programmes that will be then taken to the field for data collection.

Training and piloting in the qualitative research rounds

[Qualitative research](#) with a sub-sample of Young Lives children began in 2007, with a second round in 2008, a third in 2011, and a fourth in 2014. Research teams used a range of methods to develop detailed descriptions of the lives of case study children and of the dynamic processes and transitions that underlie their pathways through childhood.

Children's own views and understandings were the major source of qualitative data, but information was also gathered from important adults in their lives. Collecting this kind of data from children requires fieldworkers with very different skills and capacities from those needed by survey enumerators and supervisors, and a research design that balances flexibility and uniformity. Training and piloting reflected these requirements.

Young Lives qualitative research contains much that is new. There is little international experience of carrying out longitudinal qualitative research in developing countries, and of integrating qualitative and quantitative data in longitudinal studies. The need for reflection and consolidation was emphasised, and time for it built into piloting and training plans.

The first round of qualitative work was planned to complement the child-focused aspects of the Round 2 survey (2006), so the development of research questions was based on a thorough knowledge of relevant survey sections. Child-focused qualitative methodology and methods were reviewed and data-gathering techniques drawn from several disciplinary approaches were selected and piloted in Peru in 2006. This led to the selection of a set of useful methods which were fed into a research protocol, which was the basis of the first meeting with all the qualitative researchers from the UK and the four country offices.

This first meeting and a second four months later were significant milestones for both piloting and training. Both were used as opportunities for field research training that focused particularly on using participatory methods, building rapport and conducting fieldwork sensitively and ethically. Lead qualitative researchers and their assistants also received training in using qualitative software and accessing web-based resources.

The first meeting also allowed time for collectively consolidating core research questions and planning a round of pilots in all four countries. These were carried out between the first and second meetings, and teams tested different techniques and adapted them to local contexts and research priorities. Based on the results, each team collaborated with the UK-based qualitative researchers to finalise a country-specific methodology. Comprehensive fieldwork planning was carried out at the second meeting, which also generated a Memorandum of Understanding about the guiding ethical principles of the study which went on to be used in all in-country field team training. Country team members subsequently trained their own fieldwork teams, periodically supported by visits from UK-based researchers.

A lesson from these early pilot studies was that recruiting fieldwork staff could be difficult. Fieldworkers with training in qualitative or participatory research skills and experience of working with children are less numerous than survey enumerators. Gradually, however, most country teams have recruited a core team of fieldworkers who have participated in successive rounds of qualitative research.

By the third round, piloting and training aimed to review developments in the wider study and update teams on changes as well as to test new methods designed to capture information on youth aspirations and transitions, ensure that similar techniques were still age-appropriate for the sample children, and maintain and strengthen basic field skills. Piloting and training also provided an opportunity to engage teams in discussion of how to make the best

possible use of longitudinal qualitative data, and how to effectively manage the growing mass of case-level data.

By the fourth round, new questions emerged reflecting the ages of the young people, including new experiences of marriage and of parenthood. Piloting in this round was therefore focused on these potentially sensitive topics.

REFERENCES

Crivello, G., V. Morrow and E. Wilson (2013) *Young Lives Longitudinal Qualitative Research. A Guide for Researchers*, Technical Note 26, Oxford: Young Lives.

Fenn, B. (2006) 'A day in the life', unpublished reflections on piloting a longitudinal questionnaire, Oxford: Young Lives.

Seager, J. and T. de Wet (2003) *Establishing Large Panel Studies in Developing Countries: The Importance of the Young Lives Pilot Phase*, Working Paper 9, Oxford: Young Lives.

Section 11: Planning and Managing Fieldwork

May 2017

The process of collecting and recording data is the foundation of the Young Lives study. Managing fieldwork in any multi-country study in the developing world is intensive and demanding, and involves a set of basic considerations which include:

- **financial management.** Planning budgets, making resource decisions such as how many people to employ and whether and how much respondents should be paid, and establishing and following protocols for monitoring spending, are all key tasks in the fieldwork process.
- **ensuring fieldworkers are trained.** Young Lives has several different study components, each using different methods, and [fieldworker training](#) is needed for each.
- **advance logistical planning** of transport and accommodation. In all four countries this is particularly important in rural areas which are often remote and lack infrastructure.
- **obtaining ethics credentials and official permission** to operate in the field. This varies from country to country but is a vital part of preparation which often has to be started many months before fieldwork.
- **procuring and maintaining equipment.** A range of equipment is required to implement the Young Lives survey, including weighing machines, height-measuring rods, cameras, stopwatches, GPS instruments, Personal Digital Assistants (PDAs) and laptops.
- **seasonality.** In India, for example, most of the survey is done in the agricultural season between June and December because the majority of Young Lives families are casual agricultural labourers and this is the time of year when they are at home rather than on migration. In all four countries, seasonal weather conditions influence access to field sites.

Beyond these basic considerations, Young Lives has a set of unique characteristics which have a strong influence on the way fieldwork is planned and carried out. These include:

- working with children, which demands particular skills and characteristics from fieldworkers.
- working with both boys and girls, which makes gender considerations an important aspect of fieldwork planning, especially as the children approach adolescence.
- longitudinal work with two age cohorts, which lends particular importance to strict timing, as fieldwork must take place when the children are the right age.
- working in a range of geographic and linguistic settings, which presents logistical and staffing challenges for planning and managing fieldwork.
- tracking children even if they move, which increases the geographical spread of fieldwork and investment of researcher time.

- working with a range of research tools in the survey and qualitative components, which demands fieldworkers with a range of different skills, or the capacity to develop new skill sets.

Fieldwork planning

Fieldwork planning is carried out within the basic structure of the overall Collaborative Framework Agreements and annual Call Down Agreements made between Young Lives in Oxford and study partners in Ethiopia, India, Peru and Vietnam. While these plans outline deliverables on a year by year basis, detailed implementation plans and schedules are made at the country level. Box 1 outlines how this process is carried out in Ethiopia.

Box 1. Planning survey fieldwork in Ethiopia

The Principal Investigator (PI) and the Administration and Survey Officer (ASO) have central roles in planning fieldwork. The PI is involved in all stages, and is responsible for ensuring that all staff play their assigned roles. The PI and the ASO prepare a preliminary plan of the activities, schedule and budget for fieldwork, based on agreed deliverables and budgets. This plan is presented to researchers, field staff and the data manager for discussion and comment, and members of the Oxford team also provide comments and any technical support required. The process of preparing this plan includes several stages.

- A list of activities and inputs for fieldwork is made based on experience of the previous round of fieldwork.
- Additional activities and inputs are then added according to any new requirements for the current survey round.
- Time planning and scheduling is carried out for each stage of fieldwork, including time for travel, survey administration and data entry, verification and cleaning.
- Responsibilities are assigned for each activity, which helps determine what kind of personnel are needed for fieldwork, and is a key input for financial planning.
- The budget is planned, and sent to the Young Lives team in Oxford for approval.

By the end of this process, the team has an activity plan, a schedule and a budget which form the foundation of fieldwork activities.

As noted in Box 1, detailed planning of the activities and inputs necessary for carrying out the survey differs from round to round. This is due to both the longitudinal phasing of the study and the introduction of different components. Members of the Peru team, for example, observed that in Round 1, particular attention was given to developing and planning a methodology which ensured that children were enrolled in the study in a systematic way in each district, but that by Round 3, there was a strong focus on planning the introduction of [Computer Assisted Personal Interviewing \(CAPI\)](#) in the field using PDAs and tablets.

Recruiting fieldwork teams

In all four countries, fieldwork teams are needed for each survey round and study component which combine men and women, supervisors and fieldworkers, and people with different language skills. Efforts have been made to build field teams that can stay together through successive survey rounds, as this aids [cohort maintenance](#). Establishing and maintaining teams with the necessary mix of skills and qualities can be challenging when Young Lives is competing with other studies for a relatively small number of qualified staff. In all four countries, the first place to recruit is among fieldworkers who have already worked with Young Lives partner institutions.

The task of structuring and recruiting field teams is approached slightly differently in each country. In India, for example, seven teams were recruited for the Round

3 survey, each comprising one supervisor and six field investigators, three men and three women, one of whom had to be an anthropologist. In Ethiopia, by contrast, a set of minimum educational criteria were established for recruiting field supervisors and survey enumerators, in addition to which all fieldworkers had to be able to speak and write fluently in the local languages of their assigned field sites and to have experience of conducting surveys.

For the Round 2 survey, the Peru team took an innovative approach to recruiting and selecting field teams. The Round 1 survey was carried out by three teams, each formed of a supervisor, a data entry clerk and six interviewers. At the end of this first round, the study team concluded that they had underestimated the difficulties faced by supervisors, whose multiple and varied roles included maintaining the integrity of the fieldwork team. They decided that for Round 2 they would design a process that would not only select supervisors with the necessary qualities, but also lay the foundations for their training. Box 2 describes what they did.

Preparation and fieldwork

Once selected, fieldwork teams participate in [training, which often takes place during the piloting process](#) for the study component being carried out. Training processes vary between countries, between different study components, and between survey rounds. Some aspects, such as preparing fieldworkers to complete the complex income and consumption component of the household survey, are formal,

Box 2. Designing supervisor recruitment and training in Peru

The recruitment and training programme designed for Round 2 of the Young Lives survey aimed to recruit supervisors who not only had the skills needed to implement fieldwork and maintain data quality, but also the organisational capacities to lead a team through the challenges of fieldwork. The programme took into account qualities that included intellect and people skills.

Instead of advertising in newspapers, which experience showed led to an overwhelming response from people not necessarily qualified, publicity was directed at university social science departments. As well as being asked to provide a CV and detail their Quechua language skills, applicants were also invited to answer three key questions about what supervision means, their views on roles with power, and the importance of characteristics such as honesty and self-confidence. Successful applicants were invited for interview, and those who passed this stage were asked to participate in a three-day selection and training course facilitated by the study coordinator, principal investigator and a psychologist with experience in personnel selection.

This course aimed to identify of specific qualities in the candidates, partly through self-assessment of personal strengths and weaknesses with regard to teamwork and leadership qualities, and partly through role play and practice interviews. Candidates had to sit a written exam covering questions ranging from anthropometry and age calculations to responding to theoretical fieldwork scenarios, and psychological tests to evaluate their cognitive skills and personality traits. During the course, candidates were also trained in the processes of securing consent from respondents, applying structured interviews and completing questionnaire forms. Many of the sessions used participatory methods which are designed to support participants in taking a proactive role in learning.

This process, while relatively costly in terms of time and resources, was successful in that it identified some excellent supervisors who formed strong teams that were able to successfully undertake all the challenges of fieldwork. Some of the candidates who were not selected as supervisors became fieldworkers. As well as identifying these individuals, the training and selection process also strengthened and consolidated their skills. The rigorous and diverse nature of the process established very high standards and expectations for the project, as well as creating a team spirit among the group.

Source: Oré, Penny and Madrid (2012)

structured and intensive. Other aspects can be much more informal. In Peru, for example, once a draft of the survey has been prepared, fieldworkers are encouraged to apply the questionnaire as often as possible to familiarise themselves with its flow and format, practising with each other and at home and with their neighbours.

Preparing teams for the qualitative research has an important additional step that is not needed for the survey fieldwork. Unlike a questionnaire with its pre-ordained questions and anticipated responses, qualitative research is flexible and iterative. Because of this it is essential that fieldwork teams collectively examine all the existing evidence available about each community and case study child from previous data. This involves teams in reviewing community information from previous survey rounds and sub-studies, and examining the mosaic of data available on each child. Through this process, fieldworkers re-familiarise themselves the main features of each child's story – their circumstances, concerns and expectations for the future – and note anything that may need to be followed up. This activity, although time-consuming, is a vital part of preparation for qualitative fieldwork and should involve the whole team.

Pre-fieldwork logistical preparation involves team members in tasks that range from translating and editing questionnaires, to preparing databases, to ensuring that all equipment and permissions are in place. Once fieldwork teams have departed, they often do not return for several months. During fieldwork they face long hours, lengthy journeys and very basic accommodation, witness harrowing situations and encounter frequent fatigue. Throughout this process they have to maintain a high standard of scientific rigour in whatever circumstances they encounter.

In each study country, processes have been established to maintain contact with fieldworkers and preserve the security and integrity of the data they are collecting. In India, survey workers report to supervisors at the end of each day, and the supervisor checks each questionnaire for data consistency, submitting questionnaires to headquarters before moving to the next research site. In Ethiopia, field coordinators are responsible for reporting and gathering completed questionnaires, which are periodically collected or backed up and taken from the field to the office. Fieldworkers are

phoned every other day so that they can report on any challenges they are encountering. Research staff from the office also travel periodically to the field to check the validity of data and work to resolve any unanticipated problems.

All teams have also established financial management processes for fieldwork through which fieldworkers are paid and fieldwork expenses are reported, checked and submitted.

Lessons learned

The process of planning and managing complex rounds of fieldwork raised challenges from which Young Lives and the study countries have learned a great deal. Some of most important points include:

- The need to take into account that local language skills greatly reduces the pool of potential fieldworkers from which to recruit, and that there is sometimes a trade-off between linguistic skills and fieldwork experience.
- Planning for translation and back-translation of questionnaires into local languages is crucial to ensure that the purpose of each question is accurately understood by the respondent.
- Working with national government offices could entail dealing with very lengthy procurement and regulation procedures. Taking them into account at the outset is necessary to avoid delays in the implementation of the survey.
- Implementing the survey can be very stressful, and needs an effective inter-personal communication scheme in place to prevent the build-up of problems within teams.

Overcoming challenges and learning from the different survey rounds of the study has resulted in building robust fieldwork teams with members who have stayed with Young Lives through more than one fieldwork round. Their local knowledge and the long-term relationships they build with the children and their families and communities is invaluable in producing high-quality fieldwork, and an important factor in maintaining low attrition rates.

REFERENCE

Oré, B., M. Penny and S. Madrid (2012) *Selection and Training of Supervisors for Fieldwork: Experiences from the Young Lives Study in Peru*, Technical Note 24, Oxford: Young Lives.

Section 12: Cohort Maintenance – Tracking and Attrition

May 2017

Repeated interviews with the same subjects are a defining feature of longitudinal studies. Ensuring that the Young Lives children, their families and communities continue to participate in the study is essential to providing a picture of how the children's experiences of poverty change over time and across generations. It is also essential to ensuring that attrition bias is kept low, maximising the possibilities for precise analysis of the survey data, and strengthening its statistical validity.

In Round 1, there were nearly 12,000 children in the survey sample, in two cohorts, one aged 6 to 18 months and a second aged 7 to 8 years, spread across the four study countries. By the fifth round in 2016, many of these children have moved to seek education or work; some have left home, married and had children of their own; others have died. A few have decided that they no longer want to participate in the study. Keeping track of such a widely dispersed and mobile group of young people and minimising the numbers who drop out of the study presents logistical, administrative and managerial challenges.

Why track?

Young Lives is unusual in that it aims to keep track of all children in the cohort, even if they change location. Tracking is costly and time-consuming. It is a priority because:

- Tracking children between rounds reduces the amount of time spent looking for people while the survey is being carried out.
- Tracking maintains continuity of social contact and trust between researchers and respondents.
- The cohort is relatively small for a longitudinal study, and this makes minimising attrition rates particularly important for reducing attrition bias and keeping the statistical validity of the data.
- The study period is relatively long. Minimising attrition will ensure that the findings from later survey rounds are not biased.

Attrition is inevitable. The number of respondents who have not participated in each round of data collection has cumulated over time ('wave non-response'). While tracking the children aims to minimise non-response, it also aims to explain it when it does happen. This means that attrition can be analysed to ensure that it does not lead to biased inferences being drawn from survey data.

Tracking Young Lives children and maintaining response rates

After the first survey round, a tracking system was established with the aim of updating basic information about each child between survey rounds. This included household location and the names and addresses of two contacts for the child within the community but outside the household. Tracking rounds not only update information about location, but also serve as an early warning system for potential challenges during the survey, and as a mechanism for maintaining connections between researchers and respondents.

After Round 2, a set of follow-up protocols was implemented to increase the efficiency of tracking rounds. These now usually take place about a year before survey rounds, but this varies from country to country, according to both seasonal considerations and the age of Younger Cohort children when tracking is scheduled.

Wherever possible, researchers trace the new location of children who have moved and visit them at their new address. Patterns of migration differ across countries, and this has an impact on response rates. In Peru, for example, where migrants in the sample are widely dispersed, attrition rates are higher than in Vietnam, where internal migration is restricted.

Minimising attrition rates is about more than being able to locate children. Equally important is talking to respondents who want to leave the study to understand their reasons and perhaps retain them by addressing their concerns. In some cases, respondents are unhappy about the length and complexity of the survey, and the amount of time needed to complete it. Ensuring that the survey is well-paced and contains a range of different methods for engaging respondents is an important aspect of keeping attrition rates low.

All field teams produce and carry printed information about the study, explaining what data will be used for, and in some countries field teams also give photographs to respondents and their families. Pilot rounds for the survey always include training for enumerators on how to reduce refusal rates. Perhaps most important, however, is that across all four countries, many of the same fieldworkers have been retained for several rounds of the survey and often visit the same households in each round. This continuity has been helpful in keeping refusal rates low.

Patterns of attrition in the early rounds of Young Lives

Young Lives had an attrition rate of 5 per cent across the whole sample between Rounds 1 and 4. This is not only low in absolute terms, but also when compared with attrition rates for other longitudinal studies in developing countries (Outes-Leon and Dercon 2008). Table 1 illustrates an example of the causes of this attrition between Rounds 1 and 2, showing the distribution of non-response across three categories of attrition for each of the study countries.

Table 1. Attrition rates by category and country, Rounds 1 to 2

	Sample size	Attrition categories			Attrition rate (including deaths)
		Child died	Refused to answer	Untraceable	
Ethiopia	2,998	67	11	31	1.43
India	3,019	35	14	25	1.31
Peru	2,766	6	64	33	3.51
Vietnam	3,000	13	3	16	0.64
Total	11,783	121	92	105	2.7

Source: Outes-Leon and Dercon 2008: 5.

Feedback from the tracking process suggests that the relatively high rate of refusals in Peru compared with the other countries was in some cases linked to poor community understanding of the study's purpose. A member of the Peru team reflected that not enough work had been done in the first survey round to avoid giving the impression that the study was a project from which respondents would gain direct benefit. The researcher also attributed the relatively high attrition rate to family break-ups, and to a relatively empowered urban population who did not face cultural barriers in refusing to continue participating if they chose not to.

Child death accounts for a significant proportion of attrition, especially in Ethiopia, which is to be expected. In particular, the Younger Cohort can be expected to experience higher death rates than the Older Cohort. Attrition rates become more similar across cohorts when child deaths are excluded. Table 2 shows attrition rates excluding deaths for the Younger and Older Cohorts across all three survey rounds. It illustrates both wave non-response and similar levels of attrition across the two cohorts once deaths have been excluded.

Table 2. Attrition as a percentage of the whole sample, excluding deaths

	Round 1 to 2	Round 2 to 3	Round 3 to 4	Round 1 to 4
Older Cohort	1.8	1.2	5.4	8.2
Younger Cohort	1.9	0.8	0.9	3.6
Both cohorts	1.9	1.0	2.3	5.1

Attrition bias arises when sample attrition is non-random. Attrition between Round 1 and Round 2 has been assessed for attrition bias using two attrition probit tests, statistical

processes which search for patterns in outcome variables and household characteristics of attrition households (Outes-Leon and Dercon 2008). This analysis showed that there were some non-random patterns across most countries:

- child deaths correlated with households in lower wealth index, and in rural areas
- refusing and untraceable households were mostly at the higher end of the wealth index
- untraceable households were more likely to be in urban areas.

Despite following these non-random patterns, the probit tests show that attrition between Round 1 and Round 2 was an overwhelmingly random phenomenon.

Future challenges for cohort maintenance

One of the major challenges of maintaining the cohort of a longitudinal study was the need to follow the life-course events of the participants. For this reason, tracking between Rounds 3 and 4 was particularly important for the Young Lives teams, as the Older Cohort began to reach the age when many were likely to leave home and school. Among the Young Lives children there were culturally diverse ways of meeting this milestone, and some have particular implications for tracking. In Ethiopia, for example, there were instances of older children making several short migrations between rounds, during which they could easily become lost to the study.

Round 4 took place in 2013, when the Younger Cohort were the same age that the Older Cohort children were in Round 2. A longer period between the Round 3 and 4 surveys added to the possibility for attrition and demanded increased vigilance in tracking.

A key challenge for the study as it approached Round 5 was the issue of potential respondent fatigue. While all children in the sample participate in the survey at regular intervals, others are also part of the sub-sample for the qualitative research rounds, and yet others are involved in a range of sub-studies focused on particular issues ranging from social protection to early childhood care and education. The Principal Investigator in each country is responsible for ensuring that no single respondent is overloaded by the study. Maintaining effective tracking systems is essential to this.

REFERENCES

- Outes-Leon, I. and S. Dercon (2008) *Survey Attrition and Attrition Bias in Young Lives*, Technical Note 5, Oxford: Young Lives.
- Solon, A. (2016). 'Where are the children? The challenge of tracking in a longitudinal study of children.' Data Impact Blog, UK Data Service.

Section 13: Computer-assisted Personal Interviewing

May 2017

For the first two rounds of the Young Lives survey, data were collected using paper-based questionnaires. Once completed, the questionnaires were manually checked, entered and then re-entered into databases, cross-checked to highlight inconsistencies, cleaned and then transferred into data analysis software. Since the study started in 2002, various computer-assisted personal interviewing (CAPI) options have become available for replacing much of this process with electronic data collection, which allows fieldworkers to go to research sites with questionnaires loaded onto a tablet, laptop, notebook or personal digital assistant (PDA) onto which they can directly enter responses in a format that can be immediately uploaded into data files. This greatly increases the speed with which survey data can be accessed, and means that some data inconsistencies can be addressed while fieldworkers are still with respondents in the research site.

In 2008, after Round 2 was completed, the many potential advantages of CAPI led Young Lives to explore the possibilities of using it in subsequent rounds. A fundamental decision made at this stage was that the complicated nature of the survey and lack of off-the-shelf software on the market at the time required bespoke software which would allow Young Lives staff to develop their own programmes for different questionnaire components. Young Lives UK team hired a consultant to review the available software and hardware, and their findings were presented to the country study teams. It was decided to pilot PDAs and bespoke software in all four countries alongside the pilot of the Round 3 paper questionnaire. The pilot indicated that, overall, the technology was viable and the research teams willing to use it.

After this pilot stage, the Peruvian and Vietnamese study teams were keen to use electronic data collection for a proportion of the full Round 3 survey. PDAs were used to collect over 70% of the Round 3 data in Vietnam and 50% in Peru, and were also used for some questionnaires of the school component in Ethiopia. After the success in Round 3 and further developments with off-the-shelf software, Young Lives was able to roll CAPI out to all four countries in Round 4. Young Lives embarked on a process of selecting the software and hardware needed to implement this. In Round 4 the data management team was able to develop the programmes internally. CAPI was then used across all four countries in Rounds 4 and 5

Alongside its benefits, CAPI also presents challenges. CAPI technologies were developed for interviewing respondents at shopping malls and trade shows in developed countries, and their transfer to often remote study sites with poor infrastructure demands careful consideration. In addition, the size and complexity of the Young Lives survey, with

both country-specific components and multiple versions in different languages, presents particular challenges for software and programming.

Key considerations for implementing CAPI

Several different types of hardware – from hand-held devices such as mobile phones and PDAs to laptops and tablets – can be used for electronic data collection. Diverse factors had to be taken into account when considering the best hardware to carry out the survey:

- **Cost.** Balancing the requirements to run the software against our budget.
- **Size.** One researcher reflected that considering “the reality of sitting on a stone in the middle of a farmyard, balancing everything on your knee” was centrally important to evaluating the advantages and disadvantages of different kinds of hardware.
- **Battery use and ease of battery charging** in areas lacking reliable mains electricity.
- **Impact on interview dynamics.** Some fieldworkers voiced concerns about respondent’s perceptions of different types of hardware in remote areas where technology can be viewed with extreme suspicion and fear by local people. Others suggested that laptop screens would form a barrier between fieldworkers and respondents that might influence responses.
- **Security** and threat of theft.
- **Software compatibility.** Some of the software can only be used with a single type of hardware.

Having decided that the Young Lives survey would need programmable software, several other considerations needed to be taken into account when selecting software for electronic data collection:

- **Internet connectivity.** Some electronic data collection software is designed for use with the internet, and cannot function without a reliable, fast connection, which is not available in all study sites. Internet-based software does offer several advantages as well, including joint online working on programmes across countries and the opportunity to administer the questionnaire online to children who have moved and have email addresses. Internet connectivity also facilitates the development of a version control system that tracks changes in files across the whole survey.

- **Built-in checks.** Different types of software have various levels of built-in checks which alert fieldworkers when there are possible inconsistencies in survey responses. Country teams had diverse views about the level of built-in checks they would find useful in the field.
- **Opportunities for notes and comments.** Some researchers emphasised the utility of being able to make written comments, particularly to qualify data which does not fall easily into existing codes or standard consistency checks.
- **Language.** A central feature of Young Lives is that it applies questionnaires in local languages. Different software present a range of approaches towards translation and surveys in different languages, ranging from translation functions to building separate versions of the programme for each country.
- **Version control.** As each survey round is designed, piloted and revised, multiple versions of the questionnaire are in circulation. Most software includes a version control system to provide a central coordinating function, and some will automatically update a change made in one country programme in other country programmes.

Planning electronic data collection for the Round 4 and 5 surveys involved balancing these diverse considerations. Options for programmable software had increased since the pilot stage in Round 3, and three companies were invited to present their software to the Young Lives data management team. Each had both benefits and drawbacks, so making a final choice was a question of evaluating trade-offs. Surveybe, the software selected, offered the simplest user interface and the most straightforward programming language, as well as dedicated technical support from a company experienced in using the software in developing countries.

Disadvantages included no facility for internet-based version control, and the need to build five separate versions of the survey – a master programme with core questions, and four programmes for country-specific components in Amharic, Telugu, Spanish and Vietnamese. These drawbacks meant that special attention was paid to developing and implementing management protocols about programming changes and consistency between versions.

For Round 4 we switched from PDAs to tablets and laptops in all countries. While data was collected electronically in most sites, security considerations meant that data could still be recorded on paper in a small minority of locations.

Challenges in implementing electronic data collection

Implementing CAPI for Round 4 began with training in the UK for all programmers, assistants and data managers. As questionnaire sections were finalised, they were built into programmes, and the programming and questionnaire were piloted together. Challenges in rolling out CAPI across the whole study, many of which intersect with broader issues of team management and training, included:

- **Participation in programming.** Having selected software that allows Young Lives data management staff to build programmes for data collection, there are trade-offs involved between strengthening ownership and capacity by involving multiple staff in programming, and minimising errors and confusion by involving fewer staff.
- **Balancing checking and validation.** Some teams want more automatic checks built into the data collection software which quickly alert fieldworkers to anomalies in the responses they are recording. The challenge is to find a balance between challenging the answers of respondents when they fall outside expected norms and addressing the tendency for some interviewers to want to provide clean-looking data that meets expectations.
- **Potential bias.** The possibility that using electronic data would have an influence on the quality of data in comparison to paper-based questionnaires. The Peru team tested a sample of the data collected during Round 3 for such bias, and found that electronically collected data had a high level of agreement with data collected using paper-based questionnaires (Escobal and Benites, 2013).
- **Meeting country needs.** Ensuring that programmes are built which meet the needs of each country team, especially as regards local languages, has to be balanced with centralised coordination of the whole survey across all four countries and rounds.
- **Changing existing practices.** During the first two rounds, when the questionnaires were still on paper, teams had become accustomed to being able to make changes and revisions until just before the start of fieldwork. This was much harder to achieve with CAPI, and it was necessary to front load the timeline to take this into account.

REFERENCES

Escobal J. and S. Benites (2013) 'PDAs in Socio-economic Surveys: Instrument Bias, Surveyor Bias, Or Both?' *International Journal of Social Research Methodology* 16.1: 47-63.

Solon, A. (2016) '[Paper was invented in 206 BC.... It's time for a change](#)', Data Impact Blog, UK Data Service.

Section 14: Methods for Analysis

May 2017

As a longitudinal and policy-directed study comprising several distinct components, Young Lives entails many layers of analysis which rely on a variety of tools and approaches. Young Lives researchers use many different methods of analysis to discover and interpret patterns in the data collected in the field.

Units of analysis vary. Analysis of individual children allows researchers to build up detailed, contextualised life histories comprised of several types of data collected at different times, while analysis of groups of children, defined according to factors such as wealth, age, gender or location, is used to evaluate the impact of social differences on child outcomes. Although data cannot be directly compared between countries because of sampling differences, some analysis takes place at this level, highlighting how change over time is affecting the children in different countries, and comparing and contrasting the processes which are having most impact on their lives.

The structure of Young Lives, with its successive, linked cycles of survey and qualitative data collection, presents opportunities for cross-sectional and longitudinal analyses, and analysis often proceeds in both dimensions. Cross-sectional data are examined at a particular point in time, for example when a round of data is collected, and panel data are examined longitudinally across time points. Some variables are only suitable for one type of analysis.

The longitudinal character of Young Lives is a defining feature of the study. Longitudinal analysis focuses on the temporal dimensions of childhood poverty – change, process, continuity, transitions and turning points – and has many challenges. It is time-consuming and complex, involving very large quantities of data. There are particular difficulties in establishing causal relationships between different variables. Methods and software for longitudinal analysis are relatively undeveloped, especially for complex statistical models. Researchers also face the challenge of conducting longitudinal analysis of both quantitative and qualitative data in a way that produces policy relevant findings.

The complex design of Young Lives lends particular significance to defining the key terms and concepts used in the study, as these frame analysis, setting the boundaries of what can and cannot be concluded from the data. The most important are:

- **Poverty**, which is defined as multi-dimensional, complex and dynamic, subject to contextual specificity and multiple, interacting causes (Boyden and Dornan 2011). It is seen as comprising diverse material deficiencies, but also susceptibility to risk and constraint on choice.
- **Childhood**, which is seen as a distinct and universal life phase. Mirroring the holistic definition of poverty, children's development and well-being are seen as encompassing the material, physical, cognitive, psychological and socio-cultural domains. Children are viewed not as passive recipients of experience or victims of circumstance, but as social actors who have some agency in shaping their own lives.
- **Households**, which are understood as groups of people who live together, usually pool their income and eat at least one meal together when they are at home. Young Lives also recognises, however, that households are not necessarily cohesive units and that household members may act autonomously of – or in conflict with – each other.
- **Policy**, which is seen as non-linear and dynamic, and includes a full cycle of activities from conceptualisation to planning and implementation. Young Lives considers three broad types of policies: those that are specifically child-focused; general poverty reduction, development and social policies that have an impact on children; and those that shape the effectiveness of institutions that are responsible for child-related issues.

Statistical analysis of survey data

Establishing the reliability and validity of data is a keystone of good quantitative research. An early priority for statistical analysis of quantitative Young Lives data was therefore to carry out preliminary analyses of certain aspects of the whole dataset. These included:

- Analyses of the reliability and validity of the psychometric tests administered to Young Lives children. The reliability analyses examined the accuracy and stability of test scores using methods derived from Classical Test Theory and Item Response Theory. The validity analyses relied on estimating correlations between test scores and other variables such as child age and level of education, and checking these against empirical evidence reported in the wider literature on psychometric tests (Cueto et al. 2009).
- Attrition probit tests, which were used to analyse estimated models for anthropometric scores at age 1 and school enrolment at age 8 in order to enumerate any biases due to sample attrition. Further [statistical tests also examined whether attrition biases](#) might lead to biased inferences (Outes-Leon and Dercon 2008).
- Tests to compare [the Young Lives sample](#) with larger, nationally representative samples in order to be

transparent about intended and unintended sample biases. Methods included graphic comparisons of poverty rates, t-tests for the statistical significance of differences in living standards, and post-stratification, a technique used in survey analysis to incorporate population distribution of key characteristics into survey estimates (Escobal and Flores 2008).

Beyond these analyses of validity and reliability, the main purpose of statistical analysis of quantitative data is to interrogate the relationships between variables. In Young Lives, an important primary step towards the analysis of variables has been the definition and calculation of several composite variables from the raw data. These include anthropometric and other indices, but the most important is the Wealth Index, a composite, asset-based measure of economic well-being that is not production or location specific. The power of the Wealth Index is that it gives a measure of wealth which can be treated as a continuous variable and therefore be used for regression analyses and statistical modelling. Constructed from three other indices, the Wealth Index is used across countries and across survey rounds. It comprises the average of:

- a housing quality index, the simple average of types of flooring, roofing and walls, and the number of rooms per person.
- a consumer durables index, the scaled sum of ownership of a set of consumer durables, which includes up to three locally-specific items per country.
- a services index, the simple average of access to drinking water, electricity, toilet and fuel.

Classifying both children and households according to basic categories of difference such as wealth, location, age and gender is an important stage in generating simple statistics about the Young Lives survey data. These statistics form the foundation of country reports, produced for each round of the longitudinal survey to provide descriptive information about the whole sample and a starting point for cross-country comparisons and more in-depth analysis of particular thematic areas. The main method used to produce descriptive statistics is cross-tabulation, usually carried out using SPSS. The analysis presented in country reports has developed iteratively round by round:

- Round 1 reports provided baseline survey data in a standardised format. They contained descriptive statistics about the whole sample, and broken down by wealth and location.
- Round 2 reports presented analysis that was used to construct a profile of child poverty showing who and where poor children were, and the factors which were either keeping them in poverty or offering them a route out. They also presented preliminary analysis of the trends emerging between rounds.
- Round 3 reports were less uniform across countries, focusing more explicitly on patterns and themes that were emerging as important in each country. All, however, presented preliminary insights into trends over time, and the nature and extent of inequalities between children of different groups.

Since Round 2, country reports have also presented more complex statistics in the form of analysis based on regressions, mostly carried out using Stata software. Regressions are equations which represent the relationship between the values of two variables and allow the prediction of the most probable values of the outcome variable given the value of the observed variable. Regression analysis has the benefit of incorporating several relevant variables simultaneously, giving a better understanding of the processes at work. In Young Lives, different types of regression equation have been used to analyse different kinds of variable:

- A commonly used form of regression is Ordinary Least Squares (OLS), which is suitable for use when the outcome variable is continuous, such as a child's height or weight. For example, OLS was used to demonstrate that maternal education has a major positive effect on child nutrition in India, even when controlling for the significant effect of household resources (Galab, Reddy and Himaz 2008).
- A probit or logic model is used when the outcome variable can only be one or zero, such as whether or not a household is in poverty. For example, a probit model was used to show that parental education levels, mothers' ethnicity and household size are all statistically significant factors in explaining poverty in Vietnam (Le Thuc et al. 2008).
- An ordered probit regression is used when the observed variable has an ordinal value, for example when child well-being is measured by rating the child's place on a five-point response ladder. An ordered probit regression was used with this variable to show that children in rural Ethiopia who had a male head of household in Round 1 felt better off in Round 2, and that children in urban areas felt worse off compared to those in rural areas (Woldehanna, Mekonnen and Alemu 2008).

In some cases, numerous regression equations have been used to examine patterns emerging across all the Young Lives survey data. An analysis of the differences between boys and girls, for example, investigates gender gaps in all four countries across 13 determinants of child development including nutrition, education, subjective well-being and parental aspirations (Dercon and Singh 2011). The analysis used 156 regressions to analyse the determinants, each of which contained socio-economic and demographic controls, including the gender of the child.

While regression equations illuminate patterns and trends in the data, they do not explain why these phenomena happen. Establishing causality statistically first requires making a hypothesis that one variable causes a change in another variable, then testing this hypothesis through fitting an explanatory model to empirical data. Such models, composed of complex sets of equations, are used to make valid predictions reflecting causality. In a study examining a multi-dimensional phenomenon like poverty, however, such predictions are subject to biases due to unmeasured variables which may be influencing or producing the measured effect. Nonetheless, some analysis of causality has been undertaken using econometric models and

innovative approaches to controlling for bias. These include one study which finds that early childhood stunting has a significant effect on cognitive development in Vietnam (Le Thuc 2009), and another that finds that shocks and adverse events have more effect on the nature and amount of work done by children than levels of household poverty in Ethiopia (Heissler and Porter 2010).

Systematic analysis of qualitative data

Qualitative data are analysed using very different methods from the positivist, statistical approaches inherent in quantitative data analysis. The analysis of longitudinal qualitative data is a multi-stage process that relies on abductive and inductive methods of reasoning. Rather than focusing solely on either cross-sectional findings or differences between two points in time, it aims to illuminate the trajectories or pathways individual children are following, and understand how different factors interact to shape them.

Most Young Lives qualitative data take the form of text, although they also include the pictures, diagrams and photographs that are the [outputs of some data collection methods](#). The majority of the text comprises transcripts of interviews, but also includes reports on group activities and texts such as diaries produced by the children themselves. Each text file is transcribed and recorded in a common, collectively agreed format which allows it to be identified by date, research method and participants.

The principal method for organising these data to facilitate systematic longitudinal analysis is the coding and categorisation of text sections according to a mutually agreed framework. Once the framework for each round is finalised, each section of text is coded according to its content using the qualitative data analysis software Atlas.ti. Once coded, the whole dataset can be searched according to different codes and intersections of code, facilitating systematic analysis. The first stage of analysis in each round tends to be descriptive, summarising the data and highlighting emerging themes, patterns and connections, while subsequent stages reflect the specific focus and research questions developed by each country team.

The principal divisions in this framework, known as ‘super-families’, follow the core themes of the longitudinal qualitative research component, which are services, transitions and changes, and child well-being. Each super-family is sub-divided into ‘families’, topics which are derived from the main category. In the case of the transitions and changes super-family, for example, the family categories are expectations, relationships, educational transitions, and other transitions.

All country teams use the same framework of super-families and families but a third level of the framework, comprising detailed sets of codes within each family, is unique to each country. In Peru, for example, the educational transitions family was further sub-divided into communications and exchange between teachers, communications and exchange between teachers and parents, school organisational arrangements, school material resources, teacher training, adaptation, attitudes, expectations, support, children’s

readiness for school, children’s feelings, importance of preschool, and difficulties (Crivello et al. 2013).

Analysis can be approached either by theme or by individual:

- A ‘horizontal’ reading of the data is undertaken through the themes and sub-themes identified in the coding framework, enabling the identification of trends, similarities and differences between and within the research sites.
- A ‘vertical’ reading of the data entails starting with selected individual children and gathering all the available information about them in order to construct a detailed, composite image of their lives, and to triangulate data collected at different times.

Making such horizontal and vertical readings was relatively straightforward with a single round of data, but has become successively more complex as subsequent rounds of data have been collected. Research teams agree strategies for summarising longitudinal data. Tables and matrices constructed to track changes in the lives of case study children over time are an important tool for condensing data, guiding researchers to extract coded material from previous rounds. Case histories are constructed by examining all interviews with one child over successive data collection rounds, dividing the data into different domains such as education, work or risk, and creating a narrative or storyline for each domain.

Analysis of mixed methods data

One of the strengths of Young Lives is that it approaches the study of childhood poverty using a variety of research methods which generate both qualitative and quantitative longitudinal data. One of the major challenges faced by researchers is how these different types of data can be analysed and used together, given their foundations in different theories of knowledge and different understandings of validity and credibility. In practice, this challenge is met in different ways at different levels of the study:

- At the level of the two main longitudinal data collection components. For example, the research framework and [design of the longitudinal qualitative research component](#) were shaped by the findings from the two survey rounds that had already been carried out. The sub-sample for qualitative case study research was derived from the full survey sample to facilitate a clear linkage between the two datasets.
- At the level of country-specific and thematic analyses. For example, in a study of healthcare user fees in Ethiopia (Barnett and Tefera 2010) researchers analysed Round 2 survey data using cross-tabulations, chi-square tests and t-tests to produce descriptive statistics about user fees. They then used this analysis to develop a question guide for household interviews about different aspects of their experiences with user fees, which was piloted and modified in one community before being applied more widely. Once the interviews were transcribed and translated, content analysis methodology was used to develop a framework of topics

and sub-topics which formed the basis for coding, cross-referencing and analysis.

- At the level of meta-analysis for policy and communication purposes, a policy paper on gender inequalities (Pells 2011) draws together survey and qualitative data from across the study to analyse the gendered differences between boys and girls in education, domestic life and subjective well-being. The paper first identifies key questions in the policy debate on gender. It goes on to present descriptive statistics as the foundation of a narrative about how gender dynamics differ when children are at different ages, and how they

accumulate over time. This narrative uses quotations from children and condensed life histories drawn from the qualitative research, to illustrate processes and practices of everyday life that lie behind and help to explain the numerical patterns. It also, however, offers some analysis of qualitative data at country level, describing the six principle obstacles children face in meeting their goals that emerge from analysis of the qualitative data collected in India and Ethiopia, and draws on multiple regression analysis of gender bias in outcome indicators (Dercon and Singh 2011). Finally, each type of data analysis contributes a composite policy message that responds to the opening question.

REFERENCES

- Barnett, I. and B. Tefera (2010) *Poor Households' Experiences and Perception of User Fees for Healthcare: A Mixed-method Study from Ethiopia*, Working Paper 59, Oxford: Young Lives.
- Boyden, J. and P. Dornan (2011) *Putting Children at the Centre of Poverty Debates*, Policy Brief 12, Oxford: Young Lives.
- G.Crivello, V. Morrow and E. Wilson (2013) *Young Lives. Longitudinal Qualitative Research: A Guide for Researchers*, Technical Note 26, Oxford: Young Lives.
- Cueto, S., J. Leon, G. Guerrero and I. Muñoz (2009) *Psychometric Characteristics of Cognitive Development and Achievement Instruments in Round 2 of Young Lives*, Technical Note 15, Oxford: Young Lives.
- Dercon, S. and A. Singh (2011) *From Nutrition to Aspirations and Self-Efficacy: Gender Bias over Time among Children in Four Countries*, Working Paper 71, Oxford: Young Lives.
- Escobal, J. and E. Flores (2008) *An Assessment of the Young Lives Sampling Approach in Peru*, Technical Note 3, Oxford: Young Lives.
- Galab, S., P. Prudhvikar Reddy and R. Himaz (2008) *Young Lives. Round 2 Survey Report: Andhra Pradesh, India*, Oxford: Young Lives.
- Heissler, K. and C. Porter (2010) *Know Your Place: Ethiopian Children's Contributions to the Household Economy*, Working Paper 61, Oxford: Young Lives.
- Outes-Leon, I. and S. Dercon (2008) *Survey Attrition and Attrition Bias in Young Lives*, Technical Note 5, Oxford: Young Lives.
- Pells, K. (2011) *Poverty and Gender Inequalities: Evidence from Young Lives*, Policy Paper 3, Oxford: Young Lives.
- Le Thuc, D. (2009) *The Effect of Early Age Stunting on Cognitive Achievement among Children in Vietnam*, Working Paper 45, Oxford: Young Lives.
- Le Thuc, D., N.P. Nguyen, M.C. Tran, V.T. Nguyen and T.S. Von (2008) *Young Lives: Vietnam Round 2 Survey Report*, Oxford: Young Lives.
- Woldehanna, T., A. Mekonnen and T. Alemu (2008) *Young Lives: Ethiopia Round 2 Survey Report*, Oxford: Young Lives.

A Guide to Young Lives Research

Young Lives has been the first multi-disciplinary longitudinal study of childhood poverty to be carried out in more than one developing country.

Drawing on the practical experiences of our research teams, the Guide to Young Lives Research provides an overview of how the Young Lives study has been carried out, giving insight into the diverse methods and processes involved in a complex longitudinal study made up of many different components. It offers lessons on building and managing research partnerships, designing and conducting multidisciplinary research, managing and analysing data, and using research to influence policy.

Reflecting on 15 years of experience, each section summarises a different part of the research process, considering:

- what has been done in each area
- what challenges have arisen
- how ways of working have changed and developed
- what lessons have been learned.

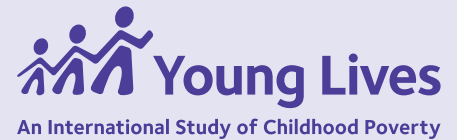
Contents

1. Young Lives: Study Overview
2. What Can Comparative Country Research Tell Us About Child Poverty?
3. What Can Longitudinal Research Tell Us About Children's Life-chances?
4. Research Ethics
5. Sampling
6. Child, Household and Community Surveys
7. Qualitative Longitudinal Research
8. School Surveys
9. Sub-studies
10. Piloting: Testing Instruments and Training Field Teams
11. Planning and Managing Fieldwork
12. Cohort Maintenance: Tracking and Attrition
13. Computer-assisted Personal Interviewing
14. Methods for Analysis

Each section of the Guide is also available to download separately.

As well as producing academic and policy-related publications, Young Lives communicates research findings through numerous platforms, including a series of illustrated mixed methods books profiling children's biographies over time, social media, video, podcasts, photography, and data visualisations.

For further details, visit www.younglives.org.uk



An International Study of Childhood Poverty

About Young Lives

Young Lives is an international study of childhood poverty, involving 12,000 children in four countries over 15 years. It is led by a team in the Department of International Development at the University of Oxford in association with research and policy partners in the four study countries: Ethiopia, India, Peru and Vietnam.

Through researching different aspects of children's lives, we seek to improve policies and programmes for children.

Young Lives Partners

Young Lives is coordinated by a small team based at the University of Oxford, led by Professor Jo Boyden.

- Ethiopian Development Research Institute, Ethiopia
- Pankhurst Development Research and Consulting plc, Ethiopia
- Centre for Economic and Social Studies, Hyderabad, India
- Save the Children India
- Sri Padmavathi Mahila Visvavidyalayam (Women's University), Andhra Pradesh, India
- Grupo de Análisis para el Desarrollo (GRADE), Peru
- Instituto de Investigación Nutricional, Peru
- Centre for Analysis and Forecasting, Vietnamese Academy of Social Sciences, Vietnam
- General Statistics Office, Vietnam
- Oxford Department of International Development, University of Oxford, UK

Contact:

Young Lives

Oxford Department of
International Development,
University of Oxford,
3 Mansfield Road,
Oxford OX1 3TB, UK

Tel: +44 (0)1865 281751

Email: younglives@younglives.org.uk

Website: www.younglives.org.uk