This is the final pre-publication version of a book chapter from:

Michael Bourdillon and Jo Boyden (eds) (2012) *Childhood Poverty: Multidisciplinary Approaches*, London: Palgrave Macmillan

Doing Longitudinal Research: Opportunities and Challenges in a Study of Childhood

Karen Brock and Caroline Knowles

Young Lives has been introduced as a longitudinal study that follows two cohorts of children in poor communities in Ethiopia, India, Peru and Vietnam as they grow into young adults. In this chapter, after pointing to the importance of longitudinal studies in the social sciences, we use the experience of designing and implementing Young Lives to reflect on some issues surrounding the process of such studies, and on the use of their findings to inform policy.

[A] The power, potential, and challenges of longitudinal research

The power of longitudinal research lies in its capacity to illuminate patterns of change in the lives of selected groups of people. Making repeated, structured observations about the same group over time allows the exclusion of unobservable individual characteristics that don't change over

time (for example, an adventurous personality with a tendency to take risks); and the identification of short- and long-term patterns of change. A classic example of this can be found in the British Doctors Study, a longitudinal study which surveyed 40,000 British doctors six times between 1957 and 2001, and found the first statistical proof that tobacco smoking increases the risk of lung cancer (Doll et al. 2004).

In the social sciences, longitudinal research can be divided into repeated cross-sectional and cohort studies. The former sample a cross-section of the population and survey it at given points in time. The latter track a group of people (a 'panel') selected because they have experienced the same event – typically birth – during a specified time period. Studies of these types can provide a glimpse into both the life histories of the individuals who make up a segment of the population, and the broader patterns of change that make up the social landscape.

Cohort studies have been particularly useful and important for understanding children and childhood across many disciplines and have used a range of qualitative and quantitative methods. An early example of a cohort study from the USA (Elder 1974) tracked the lives of children born in 1920 and 1928 through the years of the Great Depression, and had a profound influence in changing the way that the impact that unemployment, parental absence and death have on children was understood (Bronfenbrenner 1999). The UK has three national birth cohort

studies[1] and their findings have contributed to debates in many areas of social policy that are crucial for children.

Cohort studies can also be important in influencing popular understanding of life-cycle issues, including childhood. Two contemporary British TV series – *Seven-Up* and *Child of our Time*[2] – use cohort studies as the basis for periodic documentaries. *Seven-Up* set out to explore the assumption that children's futures are pre-determined by their social class, while *Child of our Time* posed the question, 'Are we born, or are we made?' The success of these two series is an example of the compelling nature of the information that can be generated by cohort studies, and how it can be used to stimulate debate on issues of child development and its social context.

The large datasets collected in developing countries have tended to focus on producing nationally representative quantitative data using surveys. They are usually cross-sectional rather than cohort studies, with the population covered changing between rounds and relatively little information about individuals. The World Bank's Living Standards Measurement Study, for example, has since 1980 supported the production of national household survey data across the developing world (Grosh and Glewwe 1995), while the USAID-supported Demographic and Health Surveys project has since 1984 collected data on population, health, HIV, and nutrition through more than 200 surveys in more than 75 countries.[3] Initiatives like these reflect the power of long-term research to provide

policymakers with quantitative data that measure not only rates of phenomena like unemployment, poverty, or HIV infection, but also shed light on the determinants of these outcomes.

In contrast to these nationally representative survey-based studies, some longitudinal research makes considerable use of qualitative research methods, including participatory and action research approaches, alongside survey techniques. Inherent in high-quality participatory research is an ethical obligation to feed findings back to participants and their communities. Inherent in action research is the notion that enquiry, learning, and action are equal parts of a research process. One example of a study design with these features is the Busselton Health Study, which has since 1966 involved the residents of an Australian town in a series of health surveys. The study was established with the intention of carrying out health research in a community setting that would provide epidemiological data, but also with the explicit intention of empowering participants to take an active role in their own health and well-being. This study alerts us to the possibility that longitudinal research can have an intended, direct effect on the lives of its participants. Challenges may arise from longitudinal studies, however, if the effects on participants are either unintended or negative, or go unnoticed.

As even this briefest of overviews illustrates, the power of longitudinal studies and their potential to provide information about change and how it happens are considerable. Longitudinal studies are also, however, both

expensive and difficult to conduct, and bring with them a host of methodological, logistical, and analytic challenges. Despite these challenges, and the ethical issues around working with children and bringing well-resourced and educated researchers into poor communities,[4] Young Lives was designed as a longitudinal cohort study because the major funder (the UK Department for International Development) wanted a study that could examine the factors and policies that make a difference in children's lives over time. Their intention was to provide evidence to help policymakers analyse and address the challenges they face in alleviating childhood poverty in developing countries. The following section discusses the methodology of the Young Lives study and outlines some of the challenges emerging.

[A] Young Lives research design and emerging methodological challenges

The twin objectives of Young Lives are to improve understanding of the causes and consequences of childhood poverty, and to inform the development and implementation of future policies and practices that will reduce childhood poverty. The first step in designing a longitudinal study to meet these objectives was to identify a conceptual and analytical framework for the enquiry.

Young Lives has adopted a multidimensional view of poverty in which income is one aspect among many others (as explained in the

Introduction). This view emphasizes that poverty is a complex, dynamic phenomenon subject to both contextual specificity and multiple, interacting contributory factors. Providing a comprehensive picture of the experience of poverty thus demands many variables, and an analytical framework encompassing many components. The Young Lives analytical framework integrates the outcomes of child poverty, the means by which poverty shapes children's transitions and trajectories through childhood, the intergenerational dynamics of poverty, and the impact of public policies and programmes designed to protect children and facilitate their progress.

As mentioned in the Introduction, this framework generated three research themes: the dynamics of poverty and uncertainty; children's experiences of poverty and its outcomes in their lives; and children's learning, their changing responsibilities, and their use of time. To provide information on these themes, Young Lives is tracking 12,000 children in the four study countries. In each country, there are 2,000 children who were born in 2000–01 and 1000 born in 1994–95, with roughly equal numbers of boys and girls in both cohorts.

A multi-stage sampling methodology known as the sentinel site surveillance system was used to select a sample of poor children in each country (Wilson, Huttley and Fenn 2006). The concept of a sentinel site comes from health studies and is a form of purposive sampling in which the site or cluster is deemed to represent a certain type of population or area, and is expected to show early signs of trends affecting those particular

people or areas. For example, monitoring a typical slum area of a given city may detect events and trends which will have an impact on most slums in that city.

In India, Ethiopia and Vietnam sites were purposively sampled according to predetermined criteria which differed from country to country. In India, selection criteria were agro-climatic areas and a number of development indicators, while in Ethiopia and Vietnam they were regional and rural/urban diversity and poverty ranking. The Peru study was slightly different, randomly selecting clusters of equal population across the country, but then excluding districts located in the richest 5 per cent of a national poverty map. Once study sites had been selected, in all four countries households with children in the right age group were randomly sampled. The result of this process is a pro-poor, clustered sample.

The study comprises five cohort survey rounds, the first carried out in 2002 and the last in 2016, the structure of which is shown in Table 1.1. Between these survey rounds, four rounds of qualitative research involving a mix of collective and individual methods are being carried out with a sub-sample of children and adults.

[Table 1.1 here]

The cohort survey at the heart of Young Lives consists of community, household, and child questionnaires. The community questionnaire is designed to provide background information about the social, economic, and environmental context of each community. The household and child

questionnaires gather information on topics such as household composition, livelihood and assets, food and non-food consumption and expenditure, socioeconomic status, social capital, economic changes and recent life history, childcare, child health and access to basic services, parental background, and the child's education. We also collect detailed time-use data for all family members, and information about the children's weight, height, and comprehension of maths and language. The survey further asks the children about their daily activities and experiences, attitudes to work and school, likes and dislikes, and hopes and aspirations for the future.

The qualitative research that supplements the survey data is building a set of 200 case studies, the main focus of which is children's own experiences and the circumstances of their daily lives. In each country, 50 case study children have been selected, and five rounds of data collection will document the changing trajectories of their individual lives in the contexts of their families and communities. These children were enabled to express their views individually in interviews and in group discussions; carers and teachers were also interviewed; and several activities observed. In addition to this, there are also shorter enquiries focused on context-specific issues, for example orphanhood in Ethiopia, and the impact of the National Rural Employment Guarantee Scheme in India.

The survey work and the qualitative research feed into each other. Analysis of survey data indicates research subjects and topics for more

intensive research, and provides patterns into which more detailed information is likely to fit. Particularly when data do not conform to established academic models, qualitative work is necessary to identify further research issues that could be included in later surveys, and possible new models for analysing survey data (see Chapter 17 by Kate Orkin for an example of this process). Further survey analysis can test the extent to which we can generalize qualitative findings. This interaction between different types of data and different kinds of analysis is a continuous process throughout the Young Lives programme.

The design of Young Lives has been shaped by the recognition that childhood poverty is a complex issue best described and analysed by integrating multiple disciplines and approaches. This multidimensional and multi-method approach is producing a unique dataset which situates children's experiences of poverty in relation to the people around them, and the sociocultural context, institutions, services, and policies that shape their lives and opportunities. At the same time, it is producing a series of challenges arising from the collection, analysis, and use of data.

In common with all longitudinal studies, Young Lives has collected, archived, and analysed huge quantities of data. Implementing this kind of study is a long-term process, and its centre is the integrity of the cohort. Ensuring low attrition rates requires building strong and lasting relationships of trust between researchers and children with their families, and investment of time in tracking children in the sample. To achieve this,

wherever possible Young Lives has supported field teams to stay together, and placed a strong emphasis on ethical issues in fieldworker training.

[A] Using Young Lives findings to influence policy and practice

Informing policies and practices that will reduce childhood poverty is the underlying driver of Young Lives. A large and diverse literature however suggests that using research to influence policy is far from straightforward, especially when the research subject is complex and multidimensional (Jones 2009).

An important step in using research findings to influence policy is to be clear about what they can and cannot be used for. Young Lives data cannot be used to compare communities or countries, or to monitor poverty. The Young Lives children and their communities are not a statistically representative sample, and this shapes the kind of narratives of childhood poverty that researchers can construct from the emerging data.

Although the Young Lives data cannot indicate, for example, the percentage of children in Peru who attend a poor-quality school, it is possible to compare the outcomes for children in the Peru sample who live in a poorer household and attend a poor-quality school with the outcomes of those children who live in a better-off household and attend a poor-quality school. It is this possibility of disaggregation according to factors of social difference – whether wealth, gender, ethnicity, or caste – that makes the Young Lives dataset particularly valuable for policymakers who aim to

deliver equitable development outcomes and effectively to meet the needs of different kinds of children. For example, one broad narrative to have emerged from the research is that poverty persists in families with poorly educated parents, and that children of better-educated parents are more likely to escape poverty. This can be further refined at the country level: in Vietnam for example, maternal education below primary completion is increasingly linked to extreme poverty among the children in the sample (Le Thuc et al. 2008). Such findings can be used both internationally and within countries to justify continued investment in girls' education as an effective intervention to break intergenerational poverty cycles.

The Young Lives data can also make a valuable contribution to understanding the impact of particular policies on children. Analysis of the effect of three different social protection schemes on children in Andhra Pradesh, Ethiopia, and Peru suggests that while social protection brings benefits for many children and is an important part of anti-poverty strategies, children benefit unevenly (some hardly at all), and there can be unintended consequences such as children having to do more work, which in turn can affect their school attendance and performance (Porter and Dornan 2010). Findings like these have resulted in recommendations to decision-makers about refining the design of ongoing programmes to make them more supportive of children.

The timing of data collection, as well as its subject and structure, influences how it can be used. As Young Lives runs from 2002 to 2016, the

final round of data collection will take place the year after the Millennium Development Goals (MDGs) will have succeeded or failed to meet their targets. Young Lives findings are well placed to contribute detailed sub-texts to unpack widely publicized headlines, particularly regarding education. For example, the second and third MDGs concern the achievement of universal primary education and the elimination of gender disparities in education, towards which considerable progress has been achieved on a global scale. Indeed, among the Young Lives children, levels of enrolment in primary school were typically high and rising, and the difference in enrolment rates between boys and girls at 12 years old was encouragingly small. However this positive progress has not necessarily been matched in terms of either attendance or quality. Data on the ability of 12-year-olds to read a simple sentence show clear differences in literacy levels between richer and poorer households in each country (Dornan 2010). These findings resonate with those on social protection in that they point to the importance of the quality of services as a policy issue. For services that support children to break out of poverty, part of this quality can be understood as interventions that match the realities of children's lives.

As well as running in parallel with the final years of the MDG target period, the period covered by Young Lives research has seen modest economic growth replaced with a world-wide economic instability. Comparing data from the first three rounds – 2002, 2006, and 2009 – will result in a detailed picture of how the food price crisis is affecting children

and reveal changes in the frequency of different kinds of shocks experienced by Young Lives families. The character of current economic shocks draws our attention to the power of the unforeseen in longitudinal surveying. It presents Young Lives researchers with the constant challenge of creating narratives about child poverty that are relevant to contemporary policy questions in a period of rapid change, and prioritizing pathways of data analysis that will provide the most policy-relevant information quickly enough for it to be put to the best possible use.

References

Elder, G. (1974) *Children of the Great Depression: Social Change in Life Experience*, Chicago IL: University of Chicago Press

Bronfenbrenner, U. (1999) 'Foreword' in G. Elder, *Twenty-fifth Anniversary Edition of Children of the Great Depression*, Boulder CO: Westview Press

Doll, R., R. Peto, J. Boreham and I. Sutherland (2004) 'Mortality in Relation to Smoking: 50 Years' Observation on Male British Doctors', *British Medical Journal* 328: 1519

Dornan, P. (2010) *Children and the Millennium Development Goals: Fragile Gains and Deep Inequalities*, Policy Paper 2, Oxford: Young Lives

Grosh, M. and P. Glewwe (1995) *A Guide to Living Standards Measurement Study Surveys and Their Datasets*, Living Standards Measurement Study Working Paper 120, Washington DC: World Bank

Jones, H. (2009) *Policy-making as Discourse: A Review of Recent Knowledge-to-Policy Literature*, IKM-Emergent/ODI Working Paper 5, London: Overseas Development Institute

Le Thuc, Duc, Ngyuen Phuong Ngoc, Tran Minh Chau, Nguyen Van Tien and Vo Thanh Son (2008) *Country Report 2: Vietnam Round 2 Survey*, Oxford: Young Lives

Porter, C. and P. Dornan (2010) *Social Protection and Children: A Synthesis of Evidence from Young Lives Longitudinal Research in Ethiopia, India and Peru*, Policy Paper 1, 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* 9.5: 351–65

Table 1.1 The structure of the panel in the study countries

	Younger Cohort (2,000 children)	Older Cohort (1,000 children)
Round 1 (2002)	6 to 18 months	7 to 8 years
Round 2 (2006)	4 to 5 years	11 to 12 years
Round 3 (2009)	7 to 8 years	14 to 15 years
Round 4 (2013)	11 to 12 years	18 to 19 years
Round 5 (2016)	14 to 15 years	21 to 22 years

[1] The National Child Development Study (1958), the British Cohort Study (1970), and the Millennium Cohort Study (2000), all housed at the Centre for Longitudinal Studies.

[2] *Seven-Up* has interviewed 14 children (selected to represent different social classes) every seven years since 1964, when the children were 7. *Child of our Time* is following 25 children (selected to represent a range of genetic, social, geographic and ethnic backgrounds) who were born around 2000 until they are 20.

[3] http://www.measuredhs.com

[4] See chapter 2 by Virginia Morrow for a full discussion of the ethical challenges of working with children and families in the Young Lives research.