



MIXED METHODS RESEARCH IN POVERTY AND VULNERABILITY

Sharing Ideas and Learning Lessons

Edited by

Keetie Roelen and Laura Camfield



Mixed Methods Research in Poverty and Vulnerability

Mixed Methods Research in Poverty and Vulnerability

Sharing Ideas and Learning Lessons

Edited by

Keetie Roelen

Research Fellow, Institute of Development Studies, University of Sussex, UK

Laura Camfield

Senior Lecturer, University of East Anglia, UK

palgrave
macmillan



Editorial matter, selection, introduction and conclusion © Keetie Roelen and Laura Camfield 2015
Individual chapters © Respective authors 2015

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6-10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The authors have asserted their rights to be identified as the authors of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published 2015 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC,
175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978-1-349-68681-0 ISBN 978-1-137-45251-1 (eBook)
DOI 10.1057/9781137452511

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Mixed methods research in poverty and vulnerability : sharing ideas and learning lessons / [edited by] Keetie Roelen (research Fellow, University of Sussex, UK), Laura Camfield (senior lecturer, University of East Anglia, UK).

pages cm

Papers from a workshop held in London, England, in July 2013.

Includes bibliographical references.

1. Poverty—Research—Methodology—Congresses. 2. Marginality, Social—Research—Methodology—Congresses. 3. Mixed methods research—Congresses. 4. Mixed methods research—Evaluation—Congresses. I. Roelen, Keetie, 1981– II. Camfield, Laura.

HC79.P6.M595 2015
305.5'69072—dc23

2015014920

Contents

<i>List of Tables and Figures</i>	vii
<i>Acknowledgements</i>	ix
<i>Notes on Contributors</i>	x
1 Introduction <i>Keetie Roelen and Laura Camfield</i>	1
Section I Poverty Measurement	7
2 Mixed Methods in Poverty Measurement: Qualitative Perspectives on the 'Necessities of Life' in the 2012 PSE-UK Survey <i>Eldin Fahmy, Eileen Sutton and Simon Pemberton</i>	9
3 Deprivation and Social Citizenship: The Objective Significance of Lived Experience <i>Daniel Edmiston</i>	33
4 Bringing Context to Poverty in Rural Rwanda: Added Value and Challenges of Mixed Methods Approaches <i>Neil Dawson</i>	61
5 Measuring the Resilience of Rural Households in Burkina Faso <i>Lucrezia Tincani and Nigel Poole</i>	87
Section II Evaluation Research	117
6 Assessing Rural Transformations: Piloting a Qualitative Impact Protocol in Malawi and Ethiopia <i>James Copestake and Fiona Remnant</i>	119
7 Evaluating the Impacts That Impact Evaluations Don't Evaluate <i>Stephen Devereux and Keetie Roelen</i>	149

Section III From Research to Policy	171
8 An Inclusive Proposal for the Use of Mixed Methods in Studying Poverty: An Application to a Colombian Municipality <i>María Fernanda Torres Penagos and Edna Bautista Hernández</i>	173
9 Challenges and Insights from Mixed Methods Impact Evaluations in Protracted Refugee Situations <i>Sally Burrows and Marian Read</i>	197
10 Competing Interpretations: Human Wellbeing and the Use of Quantitative and Qualitative Methods <i>J. Allister McGregor, Laura Camfield and Sarah Coulthard</i>	231
11 Conclusion <i>Laura Camfield and Keetie Roelen</i>	261
<i>Index</i>	271

List of Tables and Figures

Tables

2.1	Public perceptions of the necessities of life in Britain, 2012 (percentage agreement)	16
4.1	Household sampling and socio-ethnic diversity by village and district	69
4.2	Categories of land size, livestock, shelter and occupation used for hierarchical cluster analysis (based on interview data for 165 households)	73
4.3	Results of hierarchical cluster analysis displaying different groups identified from 165 households in rural Rwanda (greater detail on methods can be found in Dawson, 2013)	74
5.1	Overview of quantitative and qualitative tools used to address the research questions	93
5.2	Overview of the household composition of the eight polygamous families of the study	96
5.3	Seasonal variation livelihood diversification for polygamous families in the northern and southern field sites	102
5.4	Yearly average of livelihood diversification for monogamous households and households with more than one wife	102
5.5	Examples of decision-making observed by members of different household typologies	104
5.6	Seasonal variation of cash income and cash expenditure streams for all four households of the Tao family (all currencies indicated in CFA, with £1 = 827 CFA)	107
6.1	Summary of pilot projects, impact indicators and confounding factors	123
6.2	Responses to closed questions: Malawi projects	126

6.3	Responses to closed questions: Ethiopia projects	127
6.4	Frequency of causal statements and responses to closed questions compared from semi-structure interviews (first number) and focus groups (second number)	128
6.5	Most widely cited positive and negative drivers of change, Malawi projects	131
6.6	Most widely cited positive and negative drivers of change, Ethiopia projects	132
7.1	Typology of social relations around development programmes	158
8.1	Construction of indicators to measure the Multidimensional Poverty Index	180
8.2	Prioritising mixed diagnostic variables	191
9.1	EQAS, mixed methods and contribution analysis approaches applied to the WFP/UNHCR Evaluation Series in protracted refugee situations	203
9.2	Data collection methods and samples, by country	211

Figures

2.1	Example of 2012 ONS Opinions Survey Necessities module question format (adult items)	15
5.1	Map of Burkina Faso indicating major rivers and towns, the two field sites (triangles), rainfall gradients (dotted lines) and the area inhabited by the Mossi ethnic group (dashed circle)	94
7.1	Theory of change: development interventions	151
7.2	Theory of change: school feeding schemes	152
7.3	A matrix for assessing intended and unintended impacts	166
10.1	The OECD wellbeing framework: better life initiative	245
10.2	A wellbeing evaluation framework for social protection	250

Acknowledgements

We would like to thank all those who have contributed to the development and writing of this book. This edited volume follows a workshop held in London in July 2013 with the chapters representing the diverse presentations and rich discussions from over 40 participants. We are grateful to all workshop participants for sharing their experiences and insights, thereby laying the foundation for this book. We would like to thank the contributors who have developed their presentations into full chapters and the reviewers who have provided constructive and helpful comments.

The workshop and edited volume was made possible with financial support from ESRC grants ES-K001833-1 and RES-239-25-0006.

Notes on Contributors

Edna Bautista Hernández is an Advisor in Territorial Development at the National Planning Department, Colombia. She has a BA in Political Science and MSc in Public Policy from Universidad de Los Andes in Colombia. She was the leader of the qualitative methodology for the research presented in this volume.

Sally Burrows is Senior Evaluation Officer and Deputy Head of the Office of Evaluation at the United Nations World Food Programme (WFP), where she has been leading development of the approach to impact evaluation since 2009. She has worked on a variety of complex evaluations, including for various United Nations and government agencies and for non-governmental organisations as team member or leader. She has developed a practice in evaluation innovation during more than 25 years of experience in development cooperation and humanitarian action, with a wide range of types of organisations and in more than 25 countries.

Laura Camfield trained as an anthropologist, but now works collaboratively using qualitative and quantitative methods and trains others in their use, most recently with the DFID-funded Young Lives longitudinal research study. Her current ESRC-funded research is on enhancing the quality of cross-national methodologies used to generate data on poverty and vulnerability throughout the life course. At the University of East Anglia (UEA) she directs postgraduate research within the School of International Development (DEV), convenes the main research training module for postgraduate students and teaches on courses in impact evaluation and ethnography. She has published widely on methodology, specifically in relation to mixing methods to improve the quality of surveys and measures.

James Copestake is Professor of International Development at the University of Bath. His research experience and interests include the political economy and evaluation of poverty, international aid, development finance and agrarian change.

Sarah Coulthard is Senior Lecturer in International Development at the University of Northumbria, UK. She has an ongoing interest in environment-livelihood conflicts where sustainability and human wellbeing

trade-offs are acutely played out. Her research focus is on the impacts of social and environmental change within small-scale fisheries, mostly in South Asia and East Africa, where she has used mixed methods in her application of a wellbeing framing to study conflict between marine conservation and the wellbeing of fishing-dependent peoples.

Neil Dawson is a senior research associate in the School of International Development at UEA. His research explores the perspectives of rural inhabitants on social justice, wellbeing and poverty, with the goal of applying the methods and findings to complex issues surrounding development, agriculture and environmental management. Between 2013 and 2015 his work has been based in Laos and funded through the Ecosystem Services for Poverty Alleviation Programme. Prior to becoming a social scientist he studied and practised as both economist and ecologist, a multi-disciplinary background which has been fortuitously well-suited to integrating different methods.

Stephen Devereux is a research fellow at the Institute of Development Studies and Director of the Centre for Social Protection. He works on social protection and food security in Africa.

Daniel Edmiston is a post-doctoral researcher at the University of Oxford. He is currently working on an EU-funded project exploring the effects of social innovation policy on the most marginalised and disempowered citizens in Europe. He has previously worked for the Economic and Social Research Council, the UK Prime Minister's Strategy Unit and the Centre for Analysis of Social Exclusion, and London School of Economics and Political Science (LSE). His research interests include poverty and inequality, social security, comparative welfare reform, welfare state futures and the political economy of social citizenship.

Eldin Fahmy is senior lecturer in the School for Policy Studies at the University of Bristol, where he heads the Centre for the Study of Poverty and Social Justice. His main research interests include the measurement and analysis of poverty and social exclusion, and the applications of mixed methods designs in poverty research. Fahmy was a member of the 2012 UK Poverty and Social Exclusion Survey research team with lead responsibility for qualitative development work.

J. Allister McGregor is a professor at the Institute of Development Studies, UK. He works on issues of poverty and vulnerability and is an anthropologist and economist by training. For the last 15 years he has been advocating the use of human wellbeing as an approach to

understanding why people stay poor. He has been involved with colleagues in the development of a workable conception and methodology for studying wellbeing in diverse societal and cultural contexts.

Simon Pemberton is a Birmingham Fellow in the School of Social Policy at the University of Birmingham. His main research interests include the study of state and corporate harm, poverty, inequality and human rights and social harm, crime and criminalisation. Pemberton was a member of the 2012 UK Poverty and Social Exclusion Survey research team and contributed to qualitative development work.

Nigel Poole is Reader in International Development Policy at the Centre for Development, Environment and Policy, School of Oriental and African Studies (SOAS), University of London. He originally trained in natural sciences and agricultural extension. He switched into socio-economics and has experience in diverse countries in Latin America, the Mediterranean Basin, Sub-Saharan Africa and South and South-East Asia. His teaching is mainly in marketing, agribusiness and knowledge management. He is currently researching agrifood health and nutrition value chain linkages with the Department for International Development (DFID)-funded multi-institutional LANSAs research programme in South Asia, covering Bangladesh, India, Pakistan and Afghanistan, for which he is leading the Afghanistan Working Group.

Marian Read has more than 30 years of experience in international development with the Canadian government and United Nations WFP. As WFP's senior evaluation officer, she managed mixed methods impact, portfolio and policy evaluations of development and emergency operations in Rwanda, The Gambia, Nepal, Timor-Leste, Haiti, South Sudan, Ethiopia and Pakistan. Read pioneered the use of impact evaluation in humanitarian contexts. As WFP's deputy country director in Kenya, she led multi-disciplinary teams in implementing development and humanitarian operations that reached millions of people. As senior performance officer, she conceptualised an innovative annual report on global performance used by WFP's governing body.

Fiona Remnant is the research and communications officer on the Assessing Rural Transformations project at the University of Bath. She has previously worked with Oxfam Campaigns, and as a senior researcher at the Centre for Poverty Analysis in Sri Lanka.

Keetie Roelen is a research fellow at the Institute of Development and Co-Director of the Centre for Social Protection. In the past 10 years she

has performed research and policy advice in South-East Asia, Southern and Eastern Africa, and Central and Eastern Europe in partnership with international organisations including UNICEF, Save the Children and Concern Worldwide. Her current research interests include the dynamics of (child) poverty, social protection and the linkages between child protection and social protection. The use of mixed methods approaches is core to her research, evaluation and policy advice work.

Eileen Sutton is a research associate at the University of Bristol Biomedical Research Unit in Nutrition, Diet and Lifestyle. Sutton provides advice and support in qualitative research methods and has lead responsibility for patient and public involvement. Her research interests include poverty and social exclusion, health inequalities, and health services research and evaluation. Sutton was a member of the 2012 UK Poverty and Social Exclusion Survey research team and contributed to qualitative development work.

Lucrezia Tincani has a background in natural resource management and holds a PhD from SOAS, focusing on the socio-economics of livelihood construction in drought-prone Burkina Faso. She now works for the consultancy company Oxford Policy Management, where she focuses on issues of sustainability and inclusion in access to water and sanitation facilities.

María Fernanda Torres Penagos is a PhD candidate in Public Policy at Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM) in Mexico. She holds a BA in Economics and MSc in Public Policy from Universidad de los Andes, Colombia. She was the leader for the quantitative methodology of the research presented in this volume.

1

Introduction

Keetie Roelen and Laura Camfield

Background

The use of mixed methods in researching poverty and vulnerability and evaluation of interventions in this field has expanded rapidly in the last few years. The added value of mixed methods research in analysing poverty and vulnerability has now been widely acknowledged (see Shaffer 2013, Stern et al. 2012). Much work has been undertaken with respect to meaningfully combining methods at various stages in the research process – from generating data to analysis and reporting – and reflections thereon have led to mixed methods not only having become more ‘mainstream’ but also more robust and of greater quality. Despite an exponential growth of studies using mixed methods research in the last decade, gaps and challenges remain.

A workshop on mixed methods research in poverty and vulnerability held in London in July 2013 brought together academics, practitioners and consultants from developing and developed countries to share ideas and learn lessons about the use of mixed methods approaches in this particular area of study. A number of themes emerged in terms of where more advances are to be made, namely *credibility*, *complexity* and *usability*. This edited volume provides reflections on various issues within these themes, largely based on practical applications in research and evaluation. The collection includes contributions from different disciplinary perspectives and holds considerations on the process of data collection as well as the use of data for analytical and policy purposes.

In this introduction, we will discuss each of the three emerging themes and how they are covered in the contributions in this volume.

Credibility

Although mixed methods research in poverty and vulnerability may have firmly established itself as a valuable contribution to development studies, it still lacks *credibility* in many areas of academia. This holds particularly true for academics studying poverty and vulnerability from a singular disciplinary perspective such as economics (Shaffer 2013). Underlying this scepticism might be the epistemological clashes when trying to combine data and methods grounded in different disciplinary backgrounds. The field of impact evaluation has been particularly liable to such a divide, where quantitatively focused ‘randomistas’ often find themselves on the opposite side of heterodox quantitative, qualitative and Participatory Rural Appraisal (PRA)-influenced researchers (Bamberger et al. 2010). There may also be concerns relating to the rigour of mixed methods research given that few people are expert in both qualitative and quantitative data generation and analysis. This renders conventional guidelines for assessing quality insufficient. For example, Camfield (2014) notes that the mixing of methods requires an engagement with the metanarratives’ underlying assumptions about the topic under investigation and, therefore, with the epistemological understandings that shape those assumptions.

Considerations for improving credibility in mixed methods research, as well as the ways in which mixed methods approaches can make research more credible, are central to many contributions in this volume. From an epistemological perspective, a more explicit consideration of how different disciplinary backgrounds enter mixed methods approaches and shape the subsequent research design allows the researcher to extend beyond the implicit assumptions and methodological choices that are rooted in such disciplinary backgrounds. At the same time, greater reflection on disciplinary considerations that feed into the design of mixed methods approaches may allow users of research to overcome their own epistemological qualms. Edmiston (Chapter 3, this volume) shows how distinct citizenship theories and concepts of relative deprivation can be meaningfully and credibly combined through the study of lived experience, furthering our understanding of poverty and vulnerability in light of social, economic and cultural relations. Fahmy, Sutton and Pemberton (Chapter 2, this volume) highlight how consensus about ‘necessities of life’ is interpreted differently from quantitative and qualitative perspectives and that more deliberative methods are required for understanding public views on necessities.

Methodological opportunities for making analysis and presentation of findings more robust across a spread of methods, grounded in different disciplinary backgrounds, includes the role of methodological bilingualism by ensuring that the research team has experts from each of these backgrounds (Torres Penagos and Bautista Hernández, Chapter 8, this volume). The importance of combined use of methods and assigning equal weight to such methods is also considered imperative in overcoming epistemological and methodological divides and for adding credibility to the overall findings (Dawson, Chapter 4; Torres Penagos and Bautista Hernández). Finally, an issue often overlooked yet crucial for collecting credible and high-quality data is that of the positionality of researchers and the research-respondent relationships (Dawson). With respect to impact evaluation, Copestake and Remnant (Chapter 6, this volume) consider issues that tend to undermine its credibility, including the challenges of attribution and establishing external validity and systematic biases such as confirmation and pro-project bias. They conclude that greater emphasis on qualitative methods and the use of mixed method approaches might be most appropriate in addressing such issues.

Complexity

The use of mixed methods in research on poverty and vulnerability grounded in complexity frameworks is limited. This is despite growing recognition that pathways out of poverty are anything but linear, forcing us to think beyond direct impacts from single interventions and acknowledge ‘the multiplicity of contributions to development outcomes’ (Befani et al. 2014, p. 3). In a longitudinal mixed methods study in Bangladesh, Davis and Baulch (2011) found that household wealth can follow a range of trajectories, most of which are non-linear. Similarly, in a longitudinal study of 20 Ethiopian communities, the use of case-based methods for investigating changes over time showed communities to be ‘dynamic open complex systems’ (Bevan 2014).

Yet many studies appear to adopt the view of livelihood systems being clearly demarcated and delineated structures, and of poverty dynamics being linear processes. Evaluation studies appear particularly prone to such over-simplification. But as indicated by Picciotto (2014), while experimental impact evaluations may be able to attribute impact to an intervention, they are not able to answer questions about whether the intervention was appropriate, relevant or efficient. Pradel et al. (2013) argue that an outcome evaluation approach – focusing on proximate outcomes rather than impacts – is better suited to reflect both the complex

contexts in which interventions take place and the many factors that lead to change. Mixed methods research can make an important contribution to studies that analyse poverty and vulnerability as complex and overlapping states, as opposed to delineated and linear processes. This includes evaluations of programmes and their contributions to poverty reduction and improved livelihoods.

The notion of complexity and the role of mixed methods in addressing such complexity is a key theme in contributions throughout this volume. Copestake and Remnant argue for the pursuit of realism in impact evaluations and therefore for a more balanced integration of methods, realising that confounding factors in such evaluations are too plentiful and change too rapidly for purely experimental quantitative evaluation designs. The contribution by Devereux and Roelen (Chapter 7, this volume) is based on precisely this premise: that programme impacts are non-linear, particularly when considering social dynamics and ‘true impacts’ over a longer period of time. They argue that mixed methods approaches are crucial in unpacking that complexity.

Relationships form an important element of this complexity. The importance of relationships as an inherent but often overlooked factor in understanding poverty and wellbeing, and the role for mixed methods approaches in unpacking such relationships, is emphasised in several contributions in this volume. Edmiston reveals how an integrated study of deprivation and citizenship arrangements by combining quantitative data on objective and subjective measures of deprivation with lived experiences is crucial for unpacking the complex dynamics of deprivation at the micro-level within the context of macro-level socio-economic relations. McGregor, Camfield and Coulthard (Chapter 10, this volume) argue for the importance of using human wellbeing as a measure of development, partly on the premise that relationships are core to human wellbeing and a neglect of this dimension would obscure the complexities underlying the process of development. The use of mixed methods is considered vital for moving beyond simplistic and static understandings of wellbeing and thereby development.

Usability

Despite the additional insight and texture that mixed methods studies offer to the issues of poverty and vulnerability, policymakers often remain sceptical of such studies (for reasons discussed above) (Shaffer 2013) and subsequently make limited use of them. However, as pointed out by Sorde Marti and Mertens (2014), social scientists not only have a responsibility to identify problems and provide insight into them and

the processes leading up to them, but also to offer suggestions on how to respond to or solve those problems – to work towards ‘transformative social change’. Usability of mixed methods studies could be increased by introducing action research elements or by responding more directly to the information needs of policymakers when choosing methods and presenting findings.

Various chapters in this volume show how mixed methods studies can be user-friendly and meaningfully contribute to scientific and policy debates. Dawson shows how ‘conventional poverty measures’ provide a picture of development that may not necessarily resonate with those experiencing the effects of these policies. The mixed methods study juxtaposes findings following ‘conventional poverty measures’ and people’s own perceptions. Clear reference to information from both types of data and their contrasting insights makes the study more policy amenable. Torres Penagos and Bautista Hernández illustrate how the integration of data and methods at the municipal level generates information that is relevant for policymakers at that level, thereby facilitating policy uptake and shortening the linkages between evidence and policy impact. Finally, Burrows and Read (Chapter 9, this volume) discuss how an organisation-wide evaluation protocol can ensure that findings from country-specific mixed methods studies lead to greater policy uptake by the organisation’s managers.

This volume

As the three themes discussed above appear across the different chapters, this volume is structured around three main topics of study, clustering chapters into (i) poverty measurement, (ii) evaluation research and (iii) from research to policy. Within these three sections, individual chapters link to the knowledge gaps and challenges with respect to mixed methods research in poverty and vulnerability as discussed above. Contributions present case studies from developed and developing country contexts and applications of different approaches to mixed methods research, offering substantive findings and reflections following their use.

Section I pertains to studies regarding poverty measurement, including how mixed methods research can contribute to interpreting measures of ‘necessities of life’ in the United Kingdom (Fahmy, Sutton and Pemberton, Chapter 2), understanding deprivation and social citizenship in the UK (Edmiston, Chapter 3), contrasting pictures of development and poverty reduction in Rwanda (Dawson, Chapter 4), and vulnerability and resilience in Burkina Faso (Tincani and Poole,

Chapter 5). Section II offers reflections from impact and evaluation research, following the proposition of a new evaluation protocol for examining the impact of rural interventions (Copestake and Remnant, Chapter 6) and elaboration on an alternative evaluation framework in the area of social protection (Devereux and Roelen, Chapter 7). Section III covers contributions concerned with research for policy, sharing reflections from a multidimensional poverty study in a small municipality in Colombia (Torres Penagos and Bautista Hernández, Chapter 8), a cross-organisation evaluation of interventions in protracted refugee situations by the World Food Programme (WFP) (Burrows and Read, Chapter 9), and deliberations about the place of mixed methods in a human wellbeing approach to development (McGregor, Camfield and Coulthard, Chapter 10).

References

- Bamberger, Michael, Vijayendra Rao, and Michael Woolcock. 2010. Using Mixed Methods in Monitoring and Evaluation. Experiences from International Development. In *Policy Research Working Paper*. Washington, DC: World Bank.
- Befani, Barbara, Chris Barnett, and Elliot Stern. 2014. Introduction – Rethinking Impact Evaluation for Development. *IDS Bulletin* 45(6): 1–5. doi:10.1111/1759-5436.12108.
- Bevan, Philippa. 2014. Researching Social Change and Continuity: A Complexity-Informed Study of Twenty Rural Community Cases in Ethiopia in 1994–2015. In *Methodological Challenges and New Approaches to Research in International Development*, edited by Laura Camfield, 103–136. Basingstoke: Palgrave Macmillan.
- Camfield, Laura. 2014. Conclusion. In *Methodological Challenges and New Approaches to Research in International Development*, edited by Laura Camfield, 309–324. Basingstoke: Palgrave Macmillan.
- Davis, Peter, and Bob Baulch. 2011. Parallel Realities: Exploring Poverty Dynamics Using Mixed Methods in Rural Bangladesh. *Journal of Development Studies* 47(1): 118–142.
- Picciotto, Robert. 2014. Is Impact Evaluation Evaluation?. *European Journal of Development Research* 26(1): 31–38.
- Pradel, W., Cole, C., and G. Prain (2013). Mixing Methods for Rich and Meaningful Insight: Evaluating Changes in an Agricultural Intervention Project in the Central Andes. In *Better Evaluation*, downloaded from <http://betterevaluation.org/sites/default/files/Mixing%20Methods%20for%20Rich%20and%20Meaningful%20Insight.pdf> (30 April 2015).
- Shaffer, Paul. 2013. Ten Years of “Q-Squared”: Implications for Understanding and Explaining Poverty. *World Development* 45: 269–285.
- Sorde Marti, Teresa, and Donna M. Mertens. 2014. Mixed Methods Research With Groups at Risk: New Developments and Key Debates. *Journal of Mixed Methods Research*. doi:10.1177/1558689814527916.
- Stern, Elliot, Nicoletta Stame, John Mayne, Kim Forss, Rick Davies, and Barbara Befani. 2012. *Broadening the Range of Designs and Methods for Impact Evaluations*. London: DFID.

Section I

Poverty Measurement

2

Mixed Methods in Poverty Measurement: Qualitative Perspectives on the 'Necessities of Life' in the 2012 PSE-UK Survey¹

Eldin Fahmy, Eileen Sutton and Simon Pemberton

Introduction

Recent years have witnessed growing interest in the applications of mixed methods research strategies and specifically in the integration of qualitative and quantitative perspectives within social research. As a result, advocacy of mixed methods strategies has become increasingly accepted in research on the international analysis of poverty and vulnerability. However, despite its growing appeal in global poverty research within the United Kingdom, poverty research mixed methods designs remain rare with limited dialogue between proponents of qualitative and quantitative approaches. This partly reflects the persistence of long-standing methodological controversies in the applications of mixed methods approaches in poverty research. Combining data derived from multiple sources and generated using different data collection methods therefore continues to raise important conceptual, epistemological and methodological challenges in poverty measurement. In this chapter we illustrate some of these issues by drawing on qualitative development work undertaken as part of the 2012 UK Poverty and Social Exclusion Survey (PSE-UK) comprising a series of 14 focus group discussions in different locations in the UK. In doing so, we seek to illustrate the potential applications of qualitative evidence on poverty in assessing the credibility of evidence derived using large-scale survey methods.

The 2012 PSE-UK study is the latest and most comprehensive in a series of household surveys conducted since the 1980s adopting a 'consensual' approach to poverty measurement based on public perceptions of minimally adequate living standards. Consensual approaches to poverty measurement are now widely adopted in large-scale survey research both in the UK and internationally. However, determining the extent

and nature of public agreement on the items and activities constituting 'necessities of life' is not straightforward. In this chapter we consider the contribution of qualitative perspectives in understanding the public's views on this issue in the UK today and discuss the implications of our findings for empirical poverty measurement using social survey methods. Our findings suggest that public understandings of the term 'necessity' are diverse and may not always be consistent with researchers' interpretations, or with current usage in survey-based measurement. These findings have important implications for how we should interpret 'consensus' within survey-based consensual poverty measures, and we conclude by considering the wider methodological and epistemological implications of these findings in relation to research on poverty and vulnerability.

Mixed methods approaches in poverty research: Bridging the methodological divide?

The logic of mixed methods enquiry

Recent years have witnessed an increasing rapprochement between advocates of qualitative and quantitative methods in the practical conduct of social research, including in international research on poverty and development. Advocates of multi-method approaches have long argued that a tendency to view research methods in terms of polarised opposites (i.e. qualitative vs quantitative) encourages a methodological parochialism which frustrates our attempts to adequately address substantive research problems (e.g. Bryman, 1988; Brewer and Hunter, 1989; Tashakkori and Teddlie, 1998). As Hammersley (1992: 52) argues, the idea of a fundamental methodological divide exaggerates the differences between qualitative and quantitative approaches, and underplays the diversity of assumptions, strategies and techniques which underpin social research – a diversity which does not correspond closely with the qualitative/quantitative distinction.

Various writers have advocated the application of multiple methods as a means of overcoming the inherent weaknesses of 'mono-method' approaches. The concept of 'triangulation' as proposed by Denzin (1970) is perhaps the most widely cited rationale for multi-method approaches, and dominated early discussion of multi-method research strategies. As formulated by Denzin, this implied combining research methods to address the same research problem thereby enhancing the validity of resulting inferences. In this view, different data sources are seen as

essentially commensurate, and thus amenable to integration, in terms of the truth claims they make. However, whilst advocates of triangulation propose the combination of different methods as a means of minimising measurement error (e.g. Brewer and Hunter, 1989), the assumption that combining approaches in itself safeguards validity has been widely and effectively attacked (Fielding and Fielding, 1986; Bryman, 1992) leading to a significant reformulation of this concept in recent years (e.g. Denzin, 2012). Insofar as qualitative and quantitative methods reflect rather different concerns, and contrasting strengths and weaknesses, it is unlikely that the resulting data can be combined in the unproblematic fashion originally proposed by advocates of triangulation.

Brannen (1992: 16; see also Brannen, 2005) thus rightly refers to the complementarity of different approaches in multi-method research, in which rather than addressing the same aims methods are combined 'in order to study different levels of enquiry and in order to explore different aspects of the same problem'. Fielding and Fielding (1986: 33) similarly argue that combining methods may not necessarily enhance the accuracy of measurement but that it can produce a fuller, more multidimensional (but not more objective) account of social phenomena. Advocates of complementarity thus stress the relative merits of different methods in addressing different aspects of research problems. Combining qualitative and quantitative methods can potentially give both depth and breadth to research findings by drawing upon the different strengths of these approaches. In this chapter we seek to illustrate the applications of this interpretation of multiple method research in research on public perceptions of minimally adequate living standards.

Mixed methods in poverty research

As documented by Shaffer (2013: 269), a more systematic approach to the integration of qualitative and quantitative approaches and evidence has been a key characteristic of contemporary 'Q-squared' approaches in poverty research in the Global South, for example, in relation to the definition and social meaning of poverty, and understandings of its causes, dynamics and effects (see also Kanbur, 2005). Nevertheless, within UK research on poverty the language of dichotomy continues to pervade discussion of the methodological assumptions of research practice. Qualitative and quantitative approaches are typically taken to denote not just different methods and techniques, but also conflicting ontological and epistemological assumptions. One objective of this chapter is therefore to consider the philosophical and epistemological implications of combining qualitative and quantitative strategies in

research on poverty, and on this basis to examine the extent to which these approaches can be usefully combined in advancing the understanding of poverty. We illustrate the potential of this approach by considering the potential contribution of qualitative methods in addressing one basic question in contemporary poverty research, namely, 'what are the necessities of life?'. In doing so we hope to identify some key points of convergence and divergence between qualitative and quantitative approaches, and the issues and challenges they raise in assessing the degree of public consensus which may exist on this basic question.

One claimed advantage of mixed methods designs lies in their ability to address the 'identification problem' in research on poverty associated with difficulties in specifying the relevant dimensions of poverty, their weighting and poverty thresholds. Given that poverty is a social relationship, our definitions should reflect the meanings ascribed to the term within contemporary societies. Drawing upon Giddens' (1976) epistemology, Shaffer (2013: 270) thus observes that

[s]ocial phenomena are 'intrinsically meaningful,' in the sense that their significance and/or existence depends on the meanings ascribed to them. Understanding a concept such as 'poverty,' entails a 'double hermeneutic' of interpreting a concept which is pre-interpreted by social actors.

One claimed strength of consensual approaches to poverty measurement has been its capacity to incorporate public perceptions on poverty in the scientific measurement of this concept using survey methods (e.g. Gordon, 2006). However, as we shall see, determining the public's views on the 'necessities of life' is far from straightforward, and qualitative methods have an important role to play in better understanding public views on this.

Quantifying the 'necessities of life' in Britain

What is the consensual approach to poverty measurement?

In recent decades, consensual approaches to poverty measurement have been widely adopted in large-scale survey research both in the UK and internationally. This approach was pioneered in the UK in the 'Breadline' series of poverty surveys as originally implemented in the 1983 Poor Britain survey (Mack & Lansley, 1985). It has subsequently been developed and refined in the 1990 Breadline Britain survey

(Gordon & Pantazis, 1997), the 1999 Poverty and Social Exclusion of Britain (1999 PSE-GB) (Gordon et al., 2001), the 2002 Poverty and Social Exclusion in Northern Ireland (2002 PSE-NI) surveys (Hillyard et al., 2003) and most recently in the 2012 PSE-UK survey that is currently in progress. In recent years, this approach has also been more widely adopted in order to better measure living standards and social and material deprivation in the European Union (Guio et al., 2012) and in many EU member states including Sweden (Hallerod, 1995, 1998), Finland (Kangas and Ritakallio, 1998), Ireland (Nolan and Whelan, 1996; Layte et al., 1999), Belgium (van den Bosch, 1998) and The Netherlands (Muffels, 1993). It is also an approach increasingly widely applied further afield in both high-income countries such as Australia (Saunders, 2011; Saunders and Wong, 2011), Japan (Abe, 2010), Russia (Tchernina, 1996) and New Zealand (Perry, 2009), and in middle- and low-income countries including Bangladesh (Ahmed, 2007), South Africa (Wright, 2011) and Vietnam (Davies and Smith, 1998).

Conceptually, the consensual approach has its roots in Townsend's relative deprivation theory of poverty (e.g. Townsend, 1979, 1987). Within this perspective, poverty is viewed as an insufficient command of resources over time resulting in an inability to fulfil needs (i.e. deprivation). Crucially, needs are understood here as socially determined and relative to prevailing normative standards. However, in response to long-standing critiques of the limitations of expert judgement in determining the 'necessities of life', since the 1983 Poor Britain study, social survey methods have been used to ascertain the public's views on what constitute contemporary necessities and to incorporate these public judgements in the subsequent survey measurement of deprivation.

Nevertheless, whilst the conceptual rationale for this approach is now well established, ascertaining public attitudes towards the necessities of life in the UK today is less straightforward than might at first appear. Although the contention that a widespread public consensus exists on the items and activities needed to avoid poverty in the UK today has been central to the consensual approach, the nature and meaning of 'consensus' here is not currently well understood. Moreover, existing qualitative studies do not in general provide unambiguous empirical support for the existence of a public consensus regarding the meaning and definition of poverty itself. Rather, they demonstrate the plurality of public conceptions of poverty, for example, with regard to preferences for 'absolute' versus 'relative' interpretations (e.g. Beresford et al., 1999; Dominy and Kempson, 2006; Crowley and Vulliamy, 2007; Flaherty, 2008; Women's Budget Group, 2008). These studies reinforce survey

findings demonstrating the diversity of public views on the definition of poverty (e.g. Park et al., 2007; Clery et al., 2013). We will argue here that qualitative methods can make an important contribution in advancing the understanding of these issues. We will base our observations on the measurement approach adopted in the 2012 PSE-UK study, though a very similar methodology was implemented in the 1990 and 1999 surveys. We begin by briefly outlining the approach taken to measure the 'necessities of life' within the 2012 PSE-UK study and the role of the qualitative methods discussed in this chapter in relation to this wider study. Although consensual approaches have been increasingly widely adopted in poverty research, this approach has not been without its critics. Especial concern has focused both on the extent to which survey methods can in principle be informative about the nature of public deliberations on this topic (Walker, 1987), and the extent to which they in fact demonstrate the degree of consensus claimed by their proponents (McKay, 2004). In this chapter we therefore go on to consider what light qualitative methods can shed on these key controversies surrounding consensual approaches to measuring the necessities of life.

Quantifying the 'necessities of life' in the 2012 PSE-UK study

In the interests of methodological consistency, and in order to facilitate meaningful comparisons over time in public perceptions of the necessities of life, the same overall measurement approach and question wording were used in the 1999 and 2012 studies. In the 2012 PSE-UK study, a module on public perceptions of necessities was included within the Summer 2012 Office for National Statistics Opinions Survey in Britain, and within the June 2012 Northern Ireland Statistics and Research Agency Omnibus Survey. Based upon stratified random sampling methods, a total of 1,447 interviews were conducted in Britain and 1,015 in Northern Ireland, representing achieved response rates of 51% and 53%, respectively.

The selection of items for inclusion in the necessities survey was based upon: (a) analysis of existing survey evidence derived primarily from the 1999 PSE-GB and 2002 PSE-NI studies, as well as other relevant survey sources; (b) expert review of potential survey items conducted by the PSE-UK project team and project International Advisory Group comprising academic, policy and practitioner experts on poverty measurement and analysis, and; (c) a series of 14 focus group discussions with different population groups in England, Scotland, Wales and Northern Ireland. Our main focus here is on the role of these qualitative focus group discussions in informing understanding of the nature and extent of public

consensus on the ‘necessities of life’ in the UK today. Within the context of typologies of mixed methods research designs, for example, as proposed by Teddlie and Tashakkori (2006), the approach adopted here thus describes a sequential approach in which a primarily quantitative survey design is supplemented by qualitative data and methods that seek to provide complementary evidence to inform the selection and specification of survey instruments.

Survey respondents were asked to undertake a shuffle card exercise in order to determine those items and activities ‘which all adults should be able to afford and which they should not have to do without’. Respondents were asked to distinguish between those items and activities considered ‘necessary’ and those considered ‘desirable but not necessary’. An example of the overall procedure is provided in Figure 2.1 in relation to adult items (with separate survey tasks relating to adult activities, child items and child activities).

SHUFFLE SET A (PINK) CARDS AND SORT BOX				
[N1] <i>On these cards are a number of different items which relate to our standard of living. I would like you to indicate the living standards you feel all adults should have in Britain today by placing the cards in the appropriate box. BOX A is for items which you think are necessary – which all adults should be able to afford and which they should not have to do without. BOX B is for items which may be desirable but are not necessary.</i>				
	SET E (PINK) CARDS	A	B	Unallocated
		Necessary	Desirable but not necessary	Does not apply
		[SETGNEC]		[SETGDK]
(1)	Enough money to keep your home in a decent state of decoration			
(2)	Replace any worn out furniture			
(3)	Replace or repair broken electrical goods such as refrigerator or washing machine			

Figure 2.1 Example of 2012 ONS Opinions Survey Necessities module question format (adult items)

Table 2.1 Public perceptions of the necessities of life in Britain, 2012 (percentage agreement)

Heating to keep home adequately warm	96	Regular payments into an occupation/private pension	51
Damp-free home	94	Television	51
Two meals a day	91	Presents for friends or family once a year	46
Visiting friends or family in hospital or other institution	90	Replace worn-out clothes with new ones	46
Replace or repair broken electrical goods	86	Friends or family round for meal/drink monthly	46
Fresh fruit and vegetables every day	83	Car	44
Washing machine	82	A holiday away from home for one week a year	42
All recommended dental work/treatment	82	A small amount of money to spend each week on self	42
Celebrations on special occasions such as Christmas	80	Internet connection at home	41
A warm waterproof coat	79	Mobile phone	40
Attending weddings, funerals and other such occasions	79	Home computer	40
Telephone at home (landline or mobile)	77	Replace any worn-out furniture	39
Meat, fish or vegetarian equivalent every other day	76	An outfit to wear for social or family occasions	38
Curtains or window blinds	71	A roast joint (or its equivalent) once a week	36
A hobby or leisure activity	70	Hair done or cut regularly	35
Household contents insurance	70	Going out socially once a fortnight	34
Money to keep home in decent state of decoration	69	Attending places of worship	30
Appropriate clothes to wear for job interviews	69	Visits to friends or family in other parts of the country	27
A table, with chairs, at which all the family can eat	64	A meal out once a month	25
Taking part in sport/exercise activities or classes	56	Holidays abroad once a year	18
To be able to pay an unexpected expense of £500	55	Going out for a drink once a fortnight	17
Two pairs of all-weather shoes	54	Going to the cinema, theatre or music event monthly	15
Regular savings (£20 a month) for rainy days	52		

Within the context of quantitative consensual poverty measurement using survey methods, a public 'consensus' on the necessities of life is usually said to exist where on the basis of a representative sample survey design:

- (a) A simple majority of respondents agree that the specified items and activities are things which all adults, children or households should be able to afford and which they should not have to do without.
- (b) There are no significant social differences in respondents' perceptions of these items, that is, this consensus exists across social divisions such as those relating to class, gender, ethnicity and age.

Full results for all the items included in the 2012 ONS Opinions Survey are shown in Table 2.1.

The subset of items meeting this simple majority threshold (shown in bold) can then be subject to standard statistical methods in order to establish the degree of between-group consensus on these items, for example, between male and female respondents, old and young, and so on. Although some authors have questioned the extent of between-group consensus (see e.g. McKay, 2004), results for the 1999 PSE-GB and 2002 PSE-NI studies generally demonstrate a high degree of similarity across social groups in public perceptions of the items and activities which all people should be able to afford (Gordon et al., 2001; Pantazis et al., 2006). Initial findings from the 2012 PSE-UK also suggest a high degree of consensus between social groups on necessity items for example, with regard to gender, age group, marital status, ethnicity, education, housing tenure and income (Mack et al., 2013). These results informed the subsequent selection of items to be included in the 2012 PSE-UK main stage survey in order to establish how many households lack these 'necessities of life' because they cannot afford them, and to derive a consensual deprivation index on this basis.

Qualitative perspectives on the 'necessities of life' in Britain

What are the applications of focus group methods in the 2012 PSE-UK study?

As has been noted above, this general approach to consensual poverty measurement has been widely applied both in the UK and internationally in an increasing number of rich and poor countries. Nevertheless, this approach has not been without its critics, not least with regard to the nature and extent of the 'consensus' generated using survey

methods to establish the 'necessities of life'. As we will argue below, qualitative methods can advance our understanding of some of these issues. The overall aim of the qualitative fieldwork reported here was, therefore, to explore the public's perceptions of the necessities of life using focus group methods in order to inform the design of the 2012 ONS and NISRA modules, as well as the 2012 PSE-UK main stage survey itself. In addition, the focus groups also explored wider understandings of poverty, social exclusion and wellbeing.

A total of 14 focus group interviews were conducted in November and December 2010 in five different locations, including in each of the four territories comprising the UK: Bristol, Cardiff, London, Glasgow and Belfast. Three focus group interviews typically with between six to 10 participants were conducted in each location (two in Glasgow), with each group lasting approximately 2.5 hours in total. Participants were professionally recruited and all of them received a one-off gift payment of £35 plus travel expenses in recognition of their contribution to the research. Groups were facilitated by the research team and all sessions were audio recorded with the consent of participants in addition to the collection of contemporaneous notes during the session itself. Thematic analysis of the resulting transcripts was then conducted using proprietary software (QSR Atlas TI). Prior to delivery, research instruments and procedures were piloted in three preparatory group interviews with undergraduate students at the University of Bristol.

Since the aim of focus group discussions is generally to achieve consensus amongst participants, variability in public perceptions of the necessities of life needs to be taken into account at the design stage. The recruitment plan therefore aimed to promote homogeneity in group composition with regard to key social and demographic factors relevant to participants' views on necessities. From a mixed methods perspective, ensuring that the non-random samples recruited in qualitative development work are consistent with the desired sample profile in random sample surveys is also important (e.g. Ravallion, 2005). Separate group interviews were therefore conducted with low-income, non-low-income and mixed-income samples, and groups were also stratified by household type, and minority ethnic status. The profile of the achieved sample is described in the Appendix (Table 2A.1).

Research was conducted in two overlapping phases. In phase one groups, participants were asked to suggest potential indicators of deprivation in a relatively unstructured way. Our aim here was to generate a consensus within groups on possible indicators based primarily upon participants' own suggestions. Participants were asked to deliberate upon those items and activities which they considered to be

necessities for a 'typical' family with children in the UK today based upon the situation of a hypothetical family comprising a couple with two children. Sessions began by soliciting participant feedback on a selection of prompted items drawn from previous studies. Participants were then encouraged to add freely to and amend items as appropriate using brainstorming methods. The dynamics of participant interactions within focus groups meant that, in practice group decisions on many items were made on the basis of universal or near-universal agreement. However, where strong differences of opinion existed, a majority decision and individual respondent preferences were recorded.

Drawing on participants' suggestions in the phase one groups, the objective in the phase two groups was to 'test' the items agreed on by phase one groups, as well as to explore perceptions of wider living standards, including things which might be viewed as desirable but non-essential, or as 'luxuries'. To do so, some additional items and activities were selected for discussion based upon findings from earlier studies relating to items not widely viewed as necessities by the UK public. These items were added to the phase one results and participants were then asked to sort the combined items into three categories using card sort methods as follows:

- *Necessities*: things which are essential and which everyone should be able to afford if they want them in our society today.
- *Desirables*: things which many or most people have access to in the UK today but which are *not* essential in our society.
- *Luxuries*: things which are quite costly and exclusive and which fewer people have in our society today.

Our expectation was that a wider public consensus may exist where, using different research instruments, phase two groups independently classified broadly the same subset of items and activities as 'necessities' as those initially suggested in the more exploratory phase one groups. General consensus in this context therefore refers both to the extent of within-group agreement and the absence of substantial between-group differences in perceptions of the necessities of life in our society today – an issue to which we will return below.

What then can qualitative methods contribute in understanding public views on necessities?

Critics of the consensual approach have raised various methodological concerns about the nature of the 'consensus' achieved in sample surveys of perceptions of necessities, for example, by highlighting the

conceptual and methodological difficulties in establishing a 'valid' consensus on the basis of individualised survey responses. As Walker (1987) has argued, the nature of the 'consensus' generated using consensual survey methods is not currently well understood, and existing evidence derived using these methods reveals relatively little about the nature of public understandings of 'necessities' and the considerations that guide public deliberations on specific items. Moreover, as Walker notes, respondents in surveys on this topic are 'asked to provide immediate responses to tightly worded questions about complex and sensitive issues to which few of them will previously have given much thought' (1987: 213–214). Related to this prescient observation, survey evidence on public perceptions of necessities provides no information on the stability of respondents' views, for example, in the light of empirical evidence on the extent of deprivation or in the face of wider public discussion of these issues.

As Walker notes, survey methods tell us little or nothing about the evaluative criteria respondents employ in determining whether specific items are 'necessary', nor indeed whether the concept of 'necessity' is itself equated by respondents with those items which 'no-one should have to do without' as generally intended by researchers adopting consensual approaches. Walker's observations (1987: 219) lead him to describe the consensual approach as originally operationalised by Mack and Lansley (1985) as representing 'consensus by coincidence', and to conclude that on their own consensual survey methods:

... are unable to say anything about the criteria which people employ in judging whether or not items are 'necessary' nor, indeed, whether respondents felt equally strongly about each of the items assessed. Is the concept of 'necessary' really juxtaposed in people's minds with the notion 'should not have to do without' (which would seem to postulate some form of intervention) as the Mack and Lansley question implies? What reference groups do people use? How far are judgements grounded in experience or hearsay? How stable are people's responses in the light of information about living standards and on hearing the views of others?

Although these fundamental issues were first raised more than a quarter of a century ago, they remain pressing questions in evaluating the adequacy of existing approaches to consensual poverty measurement and in interpreting subsequent results.

Much more recently, McKay (2004) provides a more strident critique of the nature and extent of consensus generated using consensual methods. Most importantly for our purposes, based upon an analysis of the 1999 PSE-GB data, it is argued that the existing interpretation of 'consensus' by focusing exclusively on the extent of between-group social differences ignores substantial inter-individual variations in perceptions of necessities. As McKay observes, the wide distribution of responses to the necessity items suggests that more attention is needed to understand respondents' decision-making processes and interpretations of the survey task. Taken together these critiques suggest that more attention is needed to understanding respondents' decision-making processes in responding to survey items on necessities. We argue here that doing so involves taking seriously the opinions and perspectives of the public on the basis of more deliberative, qualitative methods of enquiry.

The following discussion explores these issues by examining participants' strategies for making decisions on the 'necessities of life' in the qualitative group discussions and by considering the light these issues can shed on the nature of survey response on this topic. We begin by examining the methodological challenges involved in examining public consensus on the 'necessities of life' using methods drawn from very different research paradigms and reflecting different epistemological and ontological assumptions. We then consider the contribution that focus group methods can make in better understanding respondents' decision-making processes, for example, with regard to issues of cognition, judgement and response.

Methodological issues in determining consensus

From a methodological viewpoint, it is important to emphasise that the inter-personal dynamics of focus group interactions tend towards consensus and this may in itself limit the diversity of participants' responses or the intensity of their views on specific items. It is difficult to make definitive statements about the impact of these forms of 'collective reasoning' on the selection of specific items. Nevertheless, focus group methods clearly result in quite different 'truth claims' with regard to the knowledge they create in comparison with survey-based approaches. In particular, this knowledge is assumed to be inter-subjective, negotiated and contingent rather than individualised, external and absolute.

We cannot therefore assume that the knowledge derived using these different methods is commensurate, nor is it in any real sense possible to 'triangulate' methods based on very different epistemological

assumptions in order to arrive at a better approximation of a singular social reality. Rather, they may at best be viewed as complementary in tapping different dimensions of experience. The question of whether ontological priority should be accorded to individualised or deliberative interpretations of 'necessities', thus, has no simple solution. In view of the preceding discussion on the nature of consensus, one might therefore also expect focus group methods to generate a much higher degree of within-group consensus in comparison with survey methods. As such, the process of deliberation differs markedly from the response process undertaken by individual survey respondents in selecting necessity items. Certainly, group dynamics were generally characterised by a process of convergence in participants' perspectives, both as a result of social conformity pressures, due to the impact of dominant individuals in shaping group dynamics and decisions, and most importantly as a result of the sharing of perspectives and experiences.

Nevertheless, even in a focus group context, individual participants' starting points in making decisions about necessities were often radically different. This was true both with regard to wider perspectives on the definition of poverty itself and with regard to perceptions of specific items and activities. The focus on group processes and outcomes means that the primary unit of analysis here is the group interaction rather than individual participants' preferences. It is therefore important to consider both the extent of differences in responses between groups (e.g. as a result of differences in social composition) and within-group differences in participants' views and perspectives. Overall, whilst our analysis of the interview transcripts suggested few substantial social differences in perceptions of necessities between groups, there was considerable within-group variation in these perceptions. Even after group discussion and negotiation it was not possible to reach widespread agreement on all items in many cases, and a majority verdict was therefore pursued. On this basis one might expect participants to independently derive quite different lists of necessities in the context of survey measurement, even where little social patterning of responses is evident. Although such variation may be unsurprising (especially, given McKay's 2004 analysis of the 1999 PSE-GB survey data), it was very questionable whether such variability could be construed as reflective of an underlying social 'consensus' on necessities.

Cognition issues in determining the 'necessities of life'

Participants' comments also suggested different interpretations of the survey task. A key distinction was evident between positive and

normative judgements concerning the items and activities that 'all people should not have to go without'. Some participants interpreted this phrase to denote those items that people in fact would be likely to prioritise if they were experiencing poverty, rather than the items all people ought to be able to afford (i.e. in a normative sense as intended by the research team). Thus, some participants' evaluations of specific items implied an assessment of whether the lack of an item was empirically associated with poverty vulnerability rather than any normative judgements about its social desirability.

Despite prompts from the researcher, some participants understood the task to involve selecting items necessary for people experiencing poverty. In some groups the terminology used in consensual research methods was also questioned. Some participants thus expressed concerns about the interpretation of the term 'necessity' to denote items or activities that all people should be able to afford, rather than to denote those items and activities people simply cannot live without. To this extent, researchers' interpretations of the term 'necessity' may not always be consistent with wider public understanding. The latter frequently referred to items and activities which were viewed as impossible to go without, implying a more restrictive interpretation of needs more consistent with basic needs and 'absolute' poverty measurement approaches.

Generally, references to what households and individuals 'should' be able to afford are potentially ambiguous in referring both to a normative judgement about entitlements, as well as to evaluative judgements concerning what households and individuals are in fact likely to be able to afford and need. The latter interpretation of 'necessities' as items that are both affordable and widely enjoyed, and perhaps also impossible to do without in our society today, was one widely supported within these discussion groups. Again this is not necessarily consistent with researchers' use of this term though it could be argued that such an interpretation is in principle consistent with Townsend's relative deprivation theory of poverty which underpins consensual approaches. However, participants' accounts of their decision-making sometimes also referred to estimations of how difficult it would be to do without the item in question regardless of issues of affordability, and therefore of the extent to which items and activities may be seen as 'luxuries' – however inexpensive they may be and regardless of how widely enjoyed such items are within our society today. For example, a DVD player may be something which all households 'should be able to afford' but few participants viewed it as a 'necessity'.

Int: Do you think fresh fruit or vegetables daily are important?

BRS2 RM: Not a necessity, no . . . In an ideal world yeah, everyone loves a bit of meat . . . but surely if you're on the poverty line a bowl of porridge would just see you through.

CDF3 RF: I know it's not a necessity but I think all children should experience some sort of a holiday . . . I think every family should be able . . . to afford to do that.

GLS1 RM: There's a difference between what that family should be able to afford and what a necessity is . . . I would say a TV is absolutely 100% this family should be able to afford, but it's not a necessity so it's difficult.

NI3 RF: The way I would have to look at necessity is, can you survive without it?

Need, entitlement and the abstract individual

Although the survey task refers to items and activities which 'all people should be able to afford', it quickly became apparent in the piloting of the focus group instruments and in their subsequent application that many participants felt they lacked the necessary contextual information to make a reasoned decision on necessity items. In the conduct of the focus groups, participants were provided with a hypothetical scenario or 'vignette' to facilitate group decision-making on 'necessities of life' items. Participants were therefore asked to deliberate upon those items and activities which they considered to be necessities for a 'typical' family with children in the UK today based upon the situation of a hypothetical family comprising a couple with two children, as detailed in Box 2.1.

Nevertheless, some participants had difficulty in making judgements on whether specific items should be viewed as necessities in the absence of further contextual information which might aid their decision-making, such as the family's level of income, or issues related to the contemporary costs of living for households in different circumstances. For example,

Box 2.1

Hypothetical family scenario for group deliberation

SCENARIO 1: Tom (aged 38) and Jenny (aged 35) are a married couple with two children, Jack (aged 12) and Lizzie (aged 8). They live in the suburbs of a large city. Tom works at a local hospital and is the sole wage earner within the household. Both parents are in good health.

in evaluating whether access to a car should be classified as a 'necessity', many participants felt they required information on the availability of affordable public transport, household composition (especially in relation to children and elderly or disabled residents), geographical location, working patterns and so on. Related to this point, such decisions also appeared to be difficult for some participants in the absence of further information on the full basket of goods, activities and services available to households and the extent to which different items may be substitutable (i.e. car vs affordable public transport, internet access vs telephone, etc.).

Such contextual information was also perceived to be important in shaping participants' normative judgements concerning entitlements often based upon underlying moral distinctions between the 'deserving' and 'undeserving' poor. Reference was made in many groups to the entitlements that were perceived to arise from fulfilment of social roles as workers and as good parents. There were, however, differing opinions expressed on issues of eligibility for people living on a low income or who were perceived to be 'welfare-dependent'. Whilst some referred to notions of universal entitlement, others made distinctions between the 'working' population and the 'poor', and sometimes between what we expect for ourselves and our families, and for the population as a whole.

GLS2RF: It depends how much he's earning first and foremost . . . it just really depends.

LDN2 RM: What sort of accommodation would he be able to afford? Are they social housing, are they private housing?

BRS1 RM: I'd say car only if public transport not available.

NI1 RM: I think it depends where you work and where your schools are.

LDN3 RM: If you're saying there's nobody working in the house then I'd say no way, but you've got a working household you would hope in this country that people could [go out for a meal].

CDF2 RF: I'm not being horrible to poor people but why should they be allowed to have double glazing when people who are working can't afford it.

Discussion and conclusions

What then can we conclude regarding the contribution of qualitative methods in informing the understanding of public perceptions of the necessities of life and of the definition and meaning of poverty more generally? The consensual measurement framework is now well established in poverty research practice, and survey methods have been

central to efforts to establish the extent of social consensus on the 'necessities of life' in the UK for more than 25 years. Nevertheless, little attention has been focused on the nature of the consensus generated using survey research methods; for example, by interrogating respondents' decision-making processes in reaching judgements on necessity items, in understanding respondents' own views on the concept of 'necessity' itself in this context or in considering the stability of respondents' views on these items. Although focus group methods were employed in the development of question items on the 'necessities of life' both in Mack and Lansley's ground-breaking 1985 Poor Britain study, and in the 1999 PSE-GB survey (see Middleton, 1998), little serious attention has been focused on these basic conceptual issues.

Nevertheless, as the above discussion illustrates, there are compelling grounds for questioning the extent to which existing survey approaches provide solid evidence of widespread public agreement on the necessities of life. Whilst there is very limited evidence of significant social patterning in perceptions of necessities (e.g. on the basis of income, household composition or ethnicity), even using focus group methods which tend towards consensus considerable individual level variations remain on this topic. These observations suggest that 'majoritarian' (rather than consensual) may be a better description of the overall methodology. Moreover, the focus group discussions reveal significant variability in participants' interpretations of the required task relating to issues of cognition, recall and judgement, for example, in relation to the term 'necessity' itself, and with regard to the information required to reach reasoned conclusions.

At the very least, this suggests that further question testing using methods such as cognitive interviewing and behaviour coding may be warranted in future work in this area in order to better understand the cognitive aspects of the survey task and to identify associated problems of cognition, recall, judgement and response. More generally, however, these findings also emphasise the importance of taking seriously participants' perspectives in the context of qualitative work on the necessities of life if we are to fully realise the potential of consensual approaches in the measurement of poverty. As Walker (1987: 221) concludes:

To be true to the consensual approach, people must be given scope to express their views. They need time to find their own words, to reflect on their own experience, and to grapple with the complexities of the subject. Researchers must equally be prepared to listen to their respondents and to work with their 'real-world' concepts. Similarly they should be willing to enter into a dialogue with their

respondents. Opinions grounded in ignorance, while interesting in themselves and sometimes valuable as predictors of behaviour, have little utility as a basis for policy not least because they are likely to be very unstable. Moreover they do not do justice to the intellect of the respondents or to their presumed commitment to the research exercise. Researchers are therefore obliged to provide respondents with the information which they need in order to make reasoned choices and, as far as possible, to provide feedback on the consequences of the choices made.

The above observations clearly raise some important questions regarding the credibility of existing survey evidence concerning the nature and extent of public consensus on the necessities of life in the UK today. Taking seriously participants' perspectives is therefore essential if consensual poverty measurement is to genuinely fulfil its potential as a research strategy for incorporating public views on the definition and meaning of poverty within scientific measurement. In recent years 'deliberative' methods have gained increased currency as a means of soliciting the public's reasoned and considered views on normative questions relating to poverty, for example, in order to ascertain the income needed for different households to meet minimally acceptable living standards (Bradshaw et al., 2008; Hirsch et al., 2009; Davis et al., 2010; Hirsch and Smith, 2010), and in operationalising notions of capability and well-being (Burchardt and Vizard, 2009, 2011). Nevertheless, the application of deliberative methods in understanding public views on necessities remains an under-researched topic and a pressing priority in relation to consensual poverty measurement.

The contention that poverty should be understood in relative terms with regard to prevailing societal norms is now widely accepted within poverty research. However, whilst this approach accords well with the 'double hermeneutic' associated with interpretive sociological perspectives (e.g. Giddens, 1987), too often poverty research in the UK and elsewhere remains the victim of sterile debates concerning the claimed relative merits of interpretive and positivist social enquiry resulting in limited dialogue between qualitative and quantitative research on this topic. Whilst the advantages of combining qualitative and quantitative approaches are now widely acknowledged in theory, in practice poverty research in the UK remains largely characterised by a continuing silo mentality with regard to shared methodological concerns with safeguarding the credibility of research evidence. Finding new ways to address this enduring problem should be a key priority within the emerging field of mixed methods poverty research in future.

Table 2A.1 Summary profile of focus groups

ID	Group Profile	N	Location
BRS1	Working age, no dep. children: non-low income. Older owner-occupiers living in detached homes, mixed sex group	8	Bristol
BRS2	Working age, no dep. children: non-low income. Mixed age group owner-occupiers, predominantly male	9	Bristol
BRS3	Pensioners: low income. Owner occupiers living in mixed dwelling types, predominantly female	9	Bristol
CDF1	Pensioners: low income. Owner occupiers living in mixed dwelling types, predominantly female	8	Cardiff
CDF2	Couples with dep. children: non-low income. Younger owner occupiers living in mixed dwelling types, mixed sex group	9	Cardiff
CDF3	Single parents: non-low income. Mixed aged group renters living in semi-detached homes, predominantly female	9	Cardiff
LDN1	Ethnic minority: mixed income. Mixed age group renters living in mixed dwelling types, mixed sex group	9	London
LDN2	Ethnic minority: low income. Mixed age group LA/HA renters living in terraced houses and flats, mixed sex group	8	London
LDN3	Ethnic minority: non-low income. Younger mixed tenure group living in varied dwelling types, mixed sex group	8	London
GLS1	Working age, no dep. children: mixed income. Younger mixed tenure group, all male group	3*	Glasgow
GLS2	Single parents: low income. Younger private renters living in mixed dwelling types, predominantly female	6*	Glasgow
NI1	Couples with dep. children: mixed income. Younger private renters living in semis and terraced dwellings, mixed sex group	9	Belfast
NI2	Single parents: low income. Mixed age group renters living in mixed dwelling types, predominantly female	9	Belfast
NI3	Couples with dep. children: mixed income. Mixed age group owner occupiers living in semis and terraced dwelling, predominantly female	10	Belfast

*Participant recruitment was affected by extreme weather conditions and transport disruption. As a result it was necessary to cancel one further group in Glasgow.

Note

- 1 This work was supported by the Economic and Social Research Council (Ref: RES-060-25-0052). The research materials described in this article have been deposited with the UK Data Service (SN 851404).

References

- Abe, A. (2010). Social Exclusion and Earlier Disadvantages: An Empirical Study of Poverty and Social Exclusion in Japan. *Social Science Japan Journal* 13(1): 5–30.
- Ahmed, M. (2007). Consensual Poverty in Britain, Sweden and Bangladesh: A Comparative Study. *Bangladesh e-Journal of Sociology* 4(2): 56–77.
- Beresford, P., Green, D., Lister, R., and Woodard, K. (1999). *Poverty First Hand: Poor People Speak for Themselves*. London: Child Poverty Action Group.
- Bradshaw, J., Middleton, S., Davis, A., Oldfield, N., Smith, N., Cusworth, L., and Williams, J. (2008). *A Minimum Income Standard for Britain: What People Think*. York: Joseph Rowntree Foundation.
- Brannen, J. (1992). Combining Qualitative and Quantitative Approaches: An Overview. In J. Brannen (ed.), *Mixing Methods: Qualitative and Quantitative Research*, 3–38. Aldershot: Avebury.
- Brannen, J. (2005). Mixed Methods Research: A Discussion Paper. *ESRC National Centre for Research Methods Discussion Paper*, December 2005, ESRC National Centre for Research Methods Review Paper, Southampton.
- Brewer, J., and Hunter, A. (1989). *Multimethod Research: A Synthesis of Styles*. Newbury Park, CA: Sage.
- Bryman, A. (1988). Quality and Quantity. In *Social Research*. London: Unwin Hyman.
- Bryman, A. (1992). Qualitative and Quantitative Research: Further Reflections on Their Integration. In J. Brannen (ed.), *Mixing Methods: Qualitative and Quantitative Research*, 3–37. Aldershot: Avebury.
- Burchardt, T., and Vizard, P. (2009). Developing an Equality Measurement Framework: A List of Substantive Freedoms for Adults and Children. *EHRC Research Report 18*. Manchester: Equality and Human Rights Commission.
- Burchardt, T., and Vizard, P. (2011). ‘Operationalizing’ the Capability Approach as a Basis for Equality and Human Rights Monitoring in Twenty-First-Century Britain. *Journal of Human Development and Capabilities* 12(1): 91–119.
- Clery, E., Lee, L., and Kunz, S. (2013). *Public Attitudes to Poverty and Welfare, 1983–2011*. London: NatCen/Joseph Rowntree Foundation.
- Crowley, A., and Vulliamy, C. (2007). *Listen Up! Children and Young People Talk About Poverty*. London: Save the Children.
- Davis, A., Hirsch, D., and Smith, N. (2010). *A Minimum Income Standard for Britain in 2010*. York: Joseph Rowntree Foundation.
- Davies, R., and Smith, W. (1998). *The Basic Necessities Survey: The Experience of Action Aid Vietnam*. London: Action Aid.
- Denzin, N. (1970). *The Research Act*. Chicago: Aldine.
- Denzin, N. (2012). Triangulation 2.0. *Journal of Mixed Methods Research* 6(2): 80–88.
- Dominy, N., and Kempson, E. (2006). *Understanding Older People’s Experiences of Poverty and Material Deprivation*. Norwich: Department for Work and Pensions.

- Fielding, N., and Fielding, J. (1986). *Linking Data: The Articulation of Qualitative and Quantitative Methods in Social Science*. Beverly Hills, CA: Sage.
- Flaherty, J. (2008). *Getting By, Getting Heard: Poverty and Social Exclusion in the Borders*. Glasgow: Report for the Scottish Borders Commission.
- Giddens, A. (1976). *New Rules of Sociological Method: A Positive Critique of Interpretative Sociologies*. London: Hutchinson.
- Giddens, A. (1987). *Social Theory and Modern Sociology*. Cambridge: Polity Press.
- Gordon, D. (2006). The Definition and Measurement of Poverty. In C. Pantazis, D. Gordon, and R. Levitas (eds), *Poverty and Social Exclusion in Britain: The Millennium Survey*. Bristol: The Policy Press.
- Gordon, D., Adelman, L., Ashworth, K., Bradshaw, J., Levitas, R., Middleton, S., Pantazis, C., Patsios, D., Payne, S., Townsend, P., and Williams, J. (2001). *Poverty and Social Exclusion in Britain*. Bristol: Policy Press/JRF.
- Gordon, D., and Pantazis, C. (1997). *Breadline Britain in the 1990s*. Aldershot: Ashgate.
- Guio, A.-C., Gordon, D., and Marlier, E. (2012). *Measuring Material Deprivation in the EU*. Luxembourg: Publications Office of the European Union.
- Hallerod, B. (1995). The Truly Poor: Indirect and Direct Measurement of Consensual Poverty in Sweden. *Journal of European Social Policy* 5(2): 111–129.
- Hallerod, B. (1998). Poor Swedes, Poor Britons: A Comparative Analysis of Relative Deprivation. In H. Andreß (ed.), *Empirical Poverty Research in a Comparative Perspective*, 283–312. Aldershot: Ashgate.
- Hammersley, M. (1992). Deconstructing the Qualitative–Quantitative Divide. In J. Brannen (ed.), *Mixing Methods: Qualitative and Quantitative Research*. Aldershot: Avebury.
- Hillyard, P., Kelly, G., McLaughlin, E., Patsios, D., and Tomlinson, M. (2003). *Bare Necessities: Poverty and Social Exclusion in Northern Ireland*. Belfast: Democratic Dialogue.
- Hirsch, D., Davis, A., and Smith, N. (2009). *A Minimum Income Standard for Britain in 2009*. York: Joseph Rowntree Foundation.
- Hirsch, D., and Smith, N. (2010). *Family Values – Parents’ Views on Necessities for Families with Children*. Department for Work and Pensions Research Report No. 641.
- Kanbur, R. (ed.) (2005). *Qualitative and Quantitative Poverty Appraisal: Complementarities, Tensions and the Way Forward. Q-Squared Working Paper 1*. University of Toronto, October 2005.
- Kangas, O., and Ritakallio, V. (1998). Different Methods – Different Results? Approaches to Multidimensional Poverty Measurement. In H. Andreß (ed.), *Empirical Poverty Research in a Comparative Perspective*, 167–203. Aldershot: Ashgate.
- Layte, R., Nolan, B., and Whelan, C. (1999). Targeting Poverty: Lessons from Monitoring Ireland’s National Anti-poverty Strategy. *Journal of Social Policy* 29(4): 553–575.
- Mack, J., and Lansley, S. (1985). *Poor Britain*. London: George Allen & Unwin.
- Mack, J., Lansley, S., Nandy, S., and Pantazis, P. (2013). Attitudes to Necessities in the PSE 2012 Survey: Are Minimum Standards Becoming Less Generous?. *PSE-UK Working Paper Analysis Series #4*.
- McKay, S. (2004). Poverty or Preference: What do ‘Consensual Deprivation Indicators’ Really Measure?. *Fiscal Studies* 25(2): 201–223.

- Middleton, S. (1998). Revising the Breadline Britain Questions: Relevant Findings from the Group. In J. Bradshaw, D. Gordon, R. Levitas, S. Middleton, C. Pantazis, S. Payne, and P. Townsend (eds), *Perceptions of Poverty & Social Exclusion, 1998: Report on Preparatory Research*, pp. 43–64. Bristol: University of Bristol.
- Muffels, R. (1993). Deprivation Standards and Style of Living Standards. In J. Berghman and B. Cantillon (eds), *The European Face of Social Security*, 43–59. Aldershot: Avebury.
- Nolan, B., and Whelan, C. (1996). *Resources, Deprivation and Poverty*. Oxford: Clarendon Press.
- Pantazis, C., Gordon, D., and Townsend, P. (2006). The Necessities of Life. In C. Pantazis, D. Gordon, and R. Levitas (eds), *Poverty and Social Exclusion in Britain: The Millennium Survey*, pp. 89–122. Bristol: The Policy Press.
- Park, A., Phillips, M., and Robinson, C. (2007). *Attitudes to Poverty: Findings from the British Social Attitudes Survey*. York: Joseph Rowntree Foundation.
- Perry, B. (2009). *Non-income Measures of Material Wellbeing and Hardship: First Results from the 2008 New Zealand Living Standards Survey*. Wellington, NZ: Ministry of Social Development.
- Ravallion, M. (2005). Can Qualitative Methods Help Quantitative Poverty Measurement?. In R. Kanbur (ed.), 'Qualitative and Quantitative Poverty Appraisal: Complementarities, Tensions and the Way Forward'. *Q-Squared Working Paper 1*. University of Toronto.
- Saunders, P. (2011). *Down and Out: Poverty and Exclusion in Australia*. Bristol: Policy Press.
- Saunders, P., and Wong, M., (2011). Using Deprivation Indicators to Assess the Adequacy of Australian Social Security Payments. *Journal of Poverty and Social Justice* 19(2): 91–101.
- Shaffer, P. (2013). Ten Years of 'Q-Squared': Implications for Understanding and Explaining Poverty. *World Development* 45: 269–285.
- Tashakkori, A., and Teddlie, C. (1998). *Mixed Methodologies: Combining Qualitative and Quantitative Approaches*. London: Sage.
- Tchernina, N. (1996). Economic Transition and Social Exclusion in Russia. *Research Series No. 108*. Geneva: International Institute for Labour Studies.
- Teddlie, C., and Tashakkori, A. (2006). A General Typology of Research Designs Featuring Mixed Methods. *Research in the Schools* 13(1): 12–28.
- Townsend, P. (1979). *Poverty in the UK*. London: Penguin.
- Townsend, P. (1987). Deprivation. *Journal of Social Policy* 16(2): 125–146.
- van den Bosch, K. (1998). Perceptions of the Minimum Standard of Living in Belgium: Is There Consensus?. In H. Andreß (ed.), *Empirical Poverty Research in a Comparative Perspective*, 135–166. Aldershot: Ashgate.
- Walker, R. (1987). Consensual Approaches to the Definition of Poverty: Towards an Alternative Methodology. *Journal of Social Policy* 16(2): 213–226.
- Women's Budget Group. (2008). *Women and Poverty: Experiences, Empowerment and Engagement*. York: Joseph Rowntree Foundation.
- Wright, G. (2011). Socially Perceived Necessities in South Africa: Patterns of Possession. *CASASP Working Paper No. 10*. Centre for the Analysis of South African Social Policy, University of Oxford.

3

Deprivation and Social Citizenship: The Objective Significance of Lived Experience

Daniel Edmiston

Introduction

All poverty research is underlain by a set of epistemic and ontological assumptions. Not only do these determine how poverty is operationalised and measured but also how institutions interpret and respond to poverty. As a result, poverty measurement is greatly contested. From the local to the international level, there are competing priorities determining the appropriate measure employed. Equally significant are the particularising factors that change the significance and utility of poverty measures from one context to the next. At its most general, poverty is not a fixed or isolated condition – it is a signifier of socio-economic and political relations within any given context. A methodological approach that can identify and analyse these relations goes some way towards an ‘objective’ measure that transcends contextual difference. This chapter explores the potential of a new analytical approach grounded in mixed methods to overcome some of the challenges faced in researching poverty and vulnerability in divergent contexts.

As a concept and practice, citizenship reproduces the structures and conditions that exist between communities, institutions and markets. Exploring the phenomenon of deprivation through the lens of citizenship ‘allows poverty to be analysed within a framework of institutional relationships’ (Jordan, 1996: 81). Citizenship, as defined by British sociologist T. H. Marshall, ‘is a status bestowed on those who are full members of a community. All those who possess the status are equal with respect to the rights and duties with which that status is endowed’ (1950: 28). According to Marshall, the social rights of citizenship help uphold the equality of status articulated through membership. These social rights include ‘the whole range from the right to a modicum of

economic welfare and security to the right to share to the full in the social heritage and to live the life of a civilised being according to the standards prevailing in society' (Marshall, 1950: 10–11). According to Marshall, social citizenship is principally concerned with moderating material inequalities to safeguard equality of status. Acknowledging the figurative and applied potential of social citizenship offers an analytical basis through which to understand poverty and the relations that result in inclusion or exclusion. Citizenship captures the relational dimensions of poverty that are often overlooked in poverty measures. In addition though, it reconceives the purpose and outcome of inclusion as 'equality of status':

In Britain, studies of poverty have long aimed at proving that, below a certain level of resources, individuals were excluded from participation in mainstream social activities. The evidence of exclusion was thus that they did not possess, consume or do certain things (with efforts to show that this was because they could not afford them) that many or most others did. As critics have pointed out, these methods bias such studies towards income as a measure of poverty; they also construct social exclusion mainly in terms of participation in market exchange. They make little direct attempt to study the social relations of poverty – how the poor relate to each other, and to members of other economic groups. (Jordan, 1996: 81–82)

Understood in this way, poverty and vulnerability can be seen as a relational condition – an artefact of systemic processes structured by citizenship.¹

Lived experiences are, by their nature, derived from individual interpretation. However, this chapter explores how subjective impressions could be of objective significance for deprivation measurement and poverty alleviation. In certain instances, social citizenship can be configured in such a way that it propagates rather than smooths out the material and status differentials between citizens (Wacquant, 2009; Jo, 2013). With this in mind, the phenomenology of deprivation is as much a reflection of citizenship arrangements as it is of lived experience. In this sense, the lived experience of citizens offers diagnostic insight into the reality of, but also the structuring of, deprivation. This chapter explores how mixed methods poverty research can be used to make micro-level qualitative data on lived experiences conceptually and empirically tractable in the analysis of macro-level socio-economic relations. Such an approach offers the possibility for lived experiences to concretely inform

institutional responses to poverty and vulnerability across a range of contexts.

The chapter draws on lessons learnt from a mixed methods research project employing secondary quantitative data analysis and in-depth qualitative interviews to explore the lived experiences of those negotiating poverty and affluence in day-to-day life in the UK (Edmiston, 2015). The research project and accompanying methodology have subsequently been replicated in New Zealand. The first section examines existing approaches to measuring poverty and then proposes a new analytical framework to overcome some of the challenges and incorporate some of the strengths of different approaches to poverty and vulnerability research. The second section outlines the dialectic between deprivation and social citizenship: it suggests that poverty is a manifestation of citizenship structures. In light of this, the third section outlines the objective significance of lived experience for understanding how poverty arises and how it can be addressed. The last section considers some of the analytical limitations of purely qualitative and quantitative methods when researching poverty and vulnerability. The chapter concludes by suggesting to what ends these lessons might be put.

Towards a new analytical framework for measuring deprivation

In this section, I consider a range of approaches to measuring poverty. In doing so, I explore the potential of a new analytical framework that seeks to mitigate some of the challenges and limitations faced by existing paradigms of poverty research. This framework offers a means by which different approaches can be integrated to more fully capture the dynamics of poverty and the social relations and institutional configurations that give rise to it.

Measures of poverty have traditionally been based on disposable financial resources and/or consumption patterns (Chambers, 1995). The history of modern poverty research has its roots in identifying an 'objective' or 'absolute' standard of poverty. Those pioneering this approach identified a constellation of fiscal or dietary minima: all those that fell below this threshold were considered to be living in poverty (Rowntree, 1901; Booth, 1903).² These researchers were able to quantify the extent of poverty in certain geographical areas. This method of poverty measurement has been advanced on the grounds that it is based on the actual needs of those subject to poverty rather than the behaviours, resources and consumption patterns of those that are not poor

(Joseph and Sumption, 1979). However, Peter Townsend contends that an 'objective' measure of poverty can be arrived at that is meaningfully constructed in reference to the rest of a given society:

Individuals, families and groups in the population can be said to be in poverty when they lack the resources to obtain the types of diet, participate in the activities, and have the living conditions and amenities which are customary, or at least widely encouraged or approved, in the societies to which they belong. Their resources are so seriously below those commanded by the average individual or family that they are, in effect, excluded from ordinary patterns, customs and activities. (Townsend, 1979: 31)

Measurements of relative poverty, deprivation and social exclusion all work from the understanding that the dynamics and experiences of poverty are determined by and in relation to the context within which they occur. This 'participation standard' has proven influential in UK poverty research. There are, however, a number of criticisms of this approach. Firstly, such a definition makes comparison difficult because the basic goods required for effective participation will vary across time and space. Secondly, the threshold of customary goods and activities is somewhat arbitrary (Piachaud, 1981). And thirdly, people have very different lifestyle expectations and behaviours that shape the experience and utility of the income they receive (Mack and Lansley, 1985). Doyal and Gough (1991) and Sen (1999) surmount some of these challenges by suggesting there are a range of 'needs' or 'capabilities' that are universal. How these 'needs' or 'capabilities' are met will vary. However, they suggest that it is possible to settle on an objective measure of common human 'needs' or 'capabilities' and say with some certainty and consistency that these are not met for certain groups.

Mack and Lansley (1985: 39) suggest that an objective measure of poverty can be arrived at by identifying those suffering 'an enforced lack of socially perceived necessities'. Often referred to as the consensus-based approach, this measure surveys the general public to ask which material and social goods they believe everyone should be able to have and which no one should have to do without. If 50% or more of the representative survey believe an item is necessary for an adequate standard of living, this good is classified as a 'necessity'. All those that desire and lack the means for these necessities are classified as living in poverty according to this measure. There is epistemic and instrumental value in measuring the necessities of life via a consensus-based approach. Poverty alleviation

and welfare policy design can be informed by a populist imperative to follow the minimum according to what the polity deem appropriate. However, a 50% threshold for a good to be considered a 'socially perceived necessity' has been criticised as being somewhat arbitrary (Halleröd et al., 1997). In addition, this approach may be susceptible to flawed judgements from the general public and a lack of expertise about what comprises 'the necessities of life' (as also discussed in Chapter 2 in this volume). The Minimum Income Standard project seeks to address this by combining professional expertise with public opinion and deliberation to arrive at a publicly and professionally endorsed standard of living for different demographic groups (Bradshaw et al., 2008). This multi-stage iterative process involves a series of discussion groups with people from a range of socio-economic backgrounds and the input of 'expert professionals' to design a set of minimum income standard budgets. However, such an approach still offers little explanation of why and how people experience poverty and what significance this has for their perceived and objective status within a given polity.

Poverty research that explores the lived experiences and opinions of individuals goes some way to remedy these limitations. With the exception of the Minimum Income Standard, approaches to poverty measurement cited thus far have tended to ground their definition at a level of abstraction that often fails to fully incorporate the views of the poor themselves. By exploring the subjective perceptions of those near to or in a position of material hardship, it is possible to overcome some of the paternalistic prescriptions that tend to occur in researching and tackling poverty (Atkinson, 1989; Hulme and McKay, 2005). The subjective poverty line is one such example of this. According to this approach, individuals define their needs and state whether they fall above or below a threshold they deem appropriate (Van Praag et al., 1980). Proponents of this approach 'assume that individuals themselves are the best judge of their own situation' (Van Praag et al., 1980: 462).

Poverty research that explores subjective perceptions and lived experiences tends to identify those living in poverty according to a set of a priori criteria such as those discussed above. These studies offer valuable insights into the experiences, dynamics and nature of poverty that enable poverty researchers to explore 'how social actors subjectively perceive (or sometimes fail to see) the unequal conditions of class and place that frame their biographies' (MacDonald et al., 2005: 874). However, these studies still depend on a pre-conceived understanding of socio-economic and political relations, rather than a set of a posteriori definitions and measures of poverty that are grounded in the subjective perceptions of individuals.

In *Relative Deprivation and Social Justice*, Runciman (1966) goes some way to defining poverty according to the views and experiences of individuals. He explores people's perceptions of their material and non-material welfare in relation to others. Runciman claims that people tend to compare their situation with a reference group similar to themselves rather than the whole of society. This tendency means some are less aware of the extent of income inequality and poverty and where they lie on the income distribution. Various studies have shown that people tend to make comparisons within their own social groups and networks rather than across abstract sociological categories (Evans and Kelley, 2004). As a result, people tend to inaccurately assess their socio-economic position and the relative value of their income and wealth in relation to the rest of the population (e.g. Toynbee and Walker, 2009). Those in a position of ostensible deprivation have been known to deny their own poverty (Flaherty, 2008; Shildrick and MacDonald, 2013). Equally, people in relative affluence have said that they 'struggle' to meet their individual or household needs (Hamilton, 2003). Issues of 'status anxiety' (de Botton, 2004) can be seen as confounding the credibility of subjective perceptions in defining and measuring poverty. Equally, the validity of an individual's judgement 'may be seriously limited by his or her social experience' (Sen, 2002: 860). However, some have suggested that people's exposure to and thus awareness of poverty and affluence has grown and as a result people have become better at making comparisons across different reference groups (Pradhan and Ravallion, 2000). Whilst limitations may remain in allowing the poor to define themselves, 'the views and experiences of people who consider themselves poor need not be neglected' (Roll, 1992: 21).

It is entirely possible and indeed desirable to incorporate subjective impressions into an 'objective' account and measure of poverty (Deleek and Van den Bosch, 1992). One productive way of undertaking this task is to scrutinise subjective impressions within their social context (Sen, 2002: 861). Chase and Walker (2013) recently did so by exploring how feelings of shame amongst poor people can be conceived as a 'social fact' in a consumerist society. The authors demonstrate how a 'self-conscious' emotion such as shame is constructed in reference to an individual's expectations and circumstance as well as the expectations and circumstance of others. Exploring lived experiences and attitudes through the lens of social citizenship makes it possible to do the same thing in relation to poverty and so minimise the limitations and maximise the strengths of various poverty measures.

It has long been argued that subjective impressions of circumstance can constitute an objective measure of lived experience (Bourdieu, 1977). A great deal can be gleaned when lived experiences of deprivation are incorporated into or used as the point of departure in poverty research. A growing body of evidence has demonstrated the objective significance of lived experiences and subjective perceptions. Ritterman et al. (2009) found that subjective perceptions of social status were as good a predictor of behaviours as objective socio-economic categories. Interestingly, there was an inverse effect on behaviours when asking people about their social status at a community and national level. The authors suggest that research into how people perceive themselves in relation to their reference groups is likely to increase understanding and thus efficacy of policy interventions. A range of other studies have found that subjective perceptions of one's social status are independently associated with health problems and, at times, a better predictor of health complaints than objective socio-economic status measures (Goodman et al., 2003; Cohen et al., 2008; Miyakawa et al., 2012). Goodman et al. (2005) found that social disadvantage increases stress irrespective of the sociological categories drawn upon. 'Objective' poverty measures then could be called a risk marker, rather than a risk factor, and the authors recommend researching the experience of adversity that is reflected by social characteristics and the processes by which it operates (Goodman et al., 2005). These studies demonstrate that personal views about material circumstances and social position within a given polity can profoundly affect outcomes, behaviours and social networks. By exploring the underlying causes and meanings of experience it is possible to analyse the relational and institutional features that give material expression to poverty (Green, 2006).

Some have suggested that it is important to separate the causes, consequences and experiences of poverty from its definition. According to this logic, it becomes possible to disentangle the significance of different dimensions of poverty. However, such an approach may obscure the multidimensional nature and constitution of poverty. For example, material deprivation has a number of foreseen and unforeseen consequences across time and space that can have compounding effects. These effects can influence the utility of steps taken to tackle the causes of poverty. Without a rounded conception of poverty, the efficacy of prevention and alleviation are potentially undermined (Tomlinson et al., 2008: 600). Consideration of poverty within a social citizenship framework helps overcome some of these challenges by ensuring interventions are designed to account for and tackle the systemic phenomenon of poverty.

Mindful of this, Ferragina et al. (2013) explore the psychosocial dimensions of poverty as well as its effect on the capacity of individuals to effectively participate in society. By acknowledging that 'poverty is more of a process than a state', the authors find that a range of factors, in addition to income, shape the attitudes, behaviours and outcomes of individuals (Ferragina et al., 2013: 36). Whilst this study did not adopt a citizenship framework, it was critical of equating 'participation' with 'inclusion'. By incorporating the attitudes and choices of individuals into their analysis, the authors were able to go some way in capturing the 'social relations of poverty' (Jordan, 1996: 82).

Mixed methods research in poverty measurement contributes towards 'innovative attempts to better determine who are the poor and why they are poor' (Shaffer, 2013: 270). Existing poverty research tends to draw on a set of a priori selection criteria to identify those cursorily defined as 'poor'. These criteria enable researchers to quantify the extent of poverty and additional qualitative research offers 'depth' in understanding its experience and nature. However, qualitative and quantitative methods are less commonly employed in conjunction to capture the relations of poverty that prove so significant in understanding its constitution. Using social citizenship as a framework, it is possible to explore the relational nature of poverty and the implications of this for policy design and coordination (Tomlinson et al., 2008).

In a recent mixed methods study, I identified those in a condition of material and symbolic deprivation according to three selection criteria: income, employment status and area deprivation. Those that had an equivalised net household income below 60% of the median were included because this is largely understood to be a reliable measure of relative poverty (cf. Gordon and Pantazis, 1997). Those that were unemployed were included because employment status is a significant determinant of one's income, as well as one's involvement and validation within a polity given the 'work-biased construction of citizenship responsibilities' (Lister, 2002: 107). Finally, those living in the top 30% of most deprived areas in England were included. The Index of Multiple Deprivation does not measure the deprivation of an area as such, rather the proportion of people experiencing deprivation within that area (DCLG, 2011). As a result individuals sampled are surrounded by a greater proportion of people and households that are in a similar material situation to themselves. Only individuals fulfilling all three selection criteria were included for the purposes of this study. Other studies have, through a citizenship lens, employed similar compound selection criteria to capture the multidimensional nature of exclusion

(e.g. Ferragina et al., 2014). Such studies have illustrated that a composite indicator of deprivation is able to track experiences across divergent contexts but also capture diversity across institutional settings.

Of course, employing pre-defined selection criteria (compound, or otherwise) may suffer the same limitations of other studies outlined above. However, the validity, and thus value, of selection criteria can be tested through their integrated use alongside other research methods that explore the attitudes and experiences of individuals. In my own study, secondary quantitative data analysis of the Citizenship Survey was undertaken to explore whether deprivation (as operationalised according to the three selection criteria) affected attitudes and experiences relating to social citizenship. As such, it was possible to identify what the dominant conceptions of social citizenship were amongst the general public and then explore the extent to which those in a condition of deprivation deviated from this. Rather than asking the general public what goods and activities they believe are necessary, the Citizenship Survey asks which social citizenship rights respondents feel everyone living in the UK should have. This included a range of social and economic rights such as the right to have a job and the right to be looked after by the state if you cannot look after yourself. Support for the rights of citizenship tends to be high amongst the general public (Vizard, 2010) but there is greater variation in the extent to which individuals feel they actually have these social and economic rights. There was a very strong association between the 'objective' condition of deprivation (according to the three selection criteria) and of individuals feeling that they did not have social and economic rights. This correspondence supports the validity of the selection criteria but also demonstrates the utility of employing a citizenship framework to understand the relational and institutional significance of deprivation.

Following this, qualitative interviews were undertaken with individuals who satisfied the same three selection criteria. Participants were asked about a range of aspects of their day-to-day lives: satisfaction with public services and their local area; how they saw their own material situation relative to others; whether they thought they had enough money to have a good quality of life; whether they felt they had social rights and responsibilities; what they understood a social citizen to be; and whether they felt like a social citizen according to their own definition.

With this research design it was possible to overcome some of the inherent limitations of existing paradigms of poverty research. Paternalistic and arbitrary judgements about the definition and extent of poverty were moderated by drawing on a range of selection criteria that were

subsequently explored in qualitative interviews. The validity of these selection criteria was interrogated from the point of view of those experiencing deprivation. 'Objective' measures of poverty at the household level have been criticised because they fail to account for the structural causes of difference between the lives of those subject to poverty (e.g. Millar and Glendinning, 1987). By capturing views about income, consumption patterns and behaviours, it was possible to explore the 'exercise of agency within structural inequality' (Orton, 2009) in people's day-to-day experience of poverty. For example, qualitative interviews with individuals living below the relative poverty line offered insight into the financial constraints faced by individuals. In addition, it was possible to explore how this affected people's capacity to participate in societally valued activities; whether this was a problem or concern for individuals; in what ways this affected their social relations and networks; and the coping mechanisms and strategies utilised to overcome or mitigate the effects of income poverty.

By exploring the subjective perceptions of those in a position of material deprivation, it was possible to explore how people negotiate a socio-economic and political landscape that has increasingly come to structure their marginality (Wacquant, 2009). Doing so in conjunction with 'objective' and 'socially determined' measures of deprivation gives credibility to the lived experiences and views of participants. It also adds greater depth and texture to the measures of poverty employed. Through a citizenship framework, participants were asked to think about all citizens in their social comparisons. This encouraged participants to think beyond limited reference groups to all people as social citizens, that is, equal in regard to rights and duties conferred in their status (Marshall, 1950). This appears to establish some sense of proportion when respondents are asked to consider their own position within society. Income, employment status and area deprivation were not only conceived on the basis of some abstract participation standard – they were understood as representing dominant conceptions of social citizenship amongst the general public in the UK. Lived experiences were therefore understood as a reflection of macro-institutional structures which established the policy significance, reliability and credibility of lived experience. By employing a mixed methods approach that incorporates 'objective' and 'subjective' measures, it is possible to capture the various dimensions of poverty and their significance at a conceptual and applied level. For example, the research found that many unemployed individuals interviewed not only felt excluded from the pecuniary benefits of engaging in paid employment, but also felt socially excluded due

to the valorisation of paid work in the UK (Lister, 2002). For these individuals, they felt their relative worth as citizens was somehow compromised as a result of their unemployment. In addition, many individuals suffering income poverty felt that their reliance on food banks or their inability to heat their houses undermined their common membership as a social citizen of the UK. Analysis of attitudes, experiences and behaviours through a citizenship lens offers critical insight into the socio-economic and political relations that result in deprivation.

The nexus between deprivation and citizenship

Whilst citizenship is not a concept immediately recognised as influential in daily life, the principles underpinning it prove pervasive in lived experiences (Dean and Melrose, 1996; Dwyer, 2002; Lister et al., 2003). Those experiencing deprivation in their day-to-day life are, to some extent, alienated from the common experiences that underpin and reinforce sentiments of collective belonging and mutuality. As a result, those experiencing deprivation appear to develop distinctive frames of experience and reference that 'should be understood within a dynamic of inclusion and exclusion' (Jordan, 1996: 111).

The presence of poverty within any polity corrupts the material and symbolic significance of social citizenship (Lister, 1990). The lived experience of deprivation emphasises the disjuncture between the dominant ideals of social citizenship and the socio-economic reality that exists. As such, sustained and intense forms of deprivation threaten the material and 'ontological security' (Giddens, 1991: 47) of social citizenship:

Someone who was living in absolute poverty could not be considered a citizen in any meaningful sense . . . their continued exclusion from many of the day-to-day practices that are taken for granted by the wider population indicates that the full promise of social citizenship remains a distant dream for many, and that Marshall's expansive vision of a 'civilised life' remains an illusion. (Dwyer, 2002: 84)

For those experiencing deprivation, their lives are often characterised by precarity, upheaval and vulnerability (Hooper, 2007). Such an experience makes it difficult for these individuals to reconcile their own situation with the apparent benefits of social citizenship. The findings from my own study suggest that the broken promises of social citizenship are materially and symbolically alienating and many struggle to identify with citizenship as a meaningful status given their own experience of it.

When asked, only a third of deprived respondents felt as if they were social citizens. These respondents recognised the exclusionary potential and reality of citizenship but nonetheless felt that they were able to challenge this in their interactions with public institutions. The other two thirds of deprived respondents did not feel as if they were social citizens. These individuals felt that their employment status, income poverty, surrounding environment or a combination thereof precluded them from the ostensive benefits of citizenship. Deprived and affluent respondents also differed in how they viewed the role of personal responsibility, the relationship between structure and agency, and what they felt the rights of social citizenship should be. As a result of their lived experiences, deprived respondents were much more likely to recognise exogenous factors impinging on individual agency. Owing to their experiences and material position, affluent respondents exhibited less awareness of the systemic features of deprivation but particularly struggled to grasp the 'exercise of agency within structural inequality' (Orton, 2009). It is clear that the topographies of social citizenship, resulting in a stratification of resources and status, feed into the lived experiences of different social groups. This results in differing 'ways of construing the self, perceiving the world, and relating to others' (Kraus et al., 2012: 561).

The condition of deprivation is meaningful and potent in isolation. However, the relativity of deprivation is perhaps most powerful for understanding the nature of poverty and how individuals make sense of citizenship as an ideal and practice. People seem to evaluate the worth of their resources, orientate their experiences and authenticate their position in relation to and against others so that the material resources at one's disposal are socially constituted (Dittmar and Pepper, 1994). Deprivation and affluence then are not only a material deficit or accretion, they are an expression of one's citizenship status and position (Beresford et al., 1999; Palmer et al., 2008). In this sense, deprivation and its consequences negotiated in day-to-day life undermine a common citizen identity. Other studies have found similar psychosocial effects with deprived individuals feeling like 'second class citizens' or 'excluded from citizenship' (Dean and Melrose, 1996; Dwyer, 1998; Lister et al., 2003; Humpage, 2008; Scanlon and Adlam, 2013; Roseneil, 2014).

A number of studies have attempted to explore how structured inequalities shape the divergent experiences and attitudes of citizens. However, studies that have attempted to incorporate lived experiences into an explanation of attitudinal difference tend to suggest that economic self-interest is a principal driver of attitudinal divergence (e.g. Evans, 1993; Brooks et al., 2006; Brooks and Svallfors, 2010: 208;

Evans and de Graaf, 2013). These studies fail to account for the complex ways in which citizens mediate their own experiences and interactions with citizenship structures. As Andrea Campbell (2003) puts it, public policies and institutions ‘help make citizens’. Invariably then, the lived experiences arising from structural inequality ‘influence the ways individuals understand their rights and responsibilities as members of the political community’ (Mettler and Soss, 2004: 61). The lived experience of those subject to deprivation can constitute an objective reality in this regard. These lived experiences can be broadly comparable across time and space whilst still accounting for differences across institutional settings. Irrespective of context, a nexus tends to exist between deprivation and social citizenship in most developed and developing countries. A methodology that explores how people navigate this nexus goes some way to researching poverty in a consistent and holistic way.

Inclusion and exclusion are a necessary function of all social formations. However, what is interesting is to explore the processes that lead to such a condition. Once poverty is operationalised as both a material condition and a corrosive social relation (Lister, 2004), it is possible to capture the material and symbolic significance of deprivation. Poverty research has tended to focus on the material features of deprivation. However, there is an increasing appreciation for the social and figurative facets of poverty (e.g. Chase and Walker, 2013). Jo (2013) suggests that ‘social needs are inherently connected to the broader context of the social, cultural and economic systems and institutions at work’ (Jo, 2013: 517). By grounding poverty research in the lived experiences of notionally equal citizens facing material hardship, it is possible to better understand these processes and undertake a thicker analysis of poverty.

The objective significance of lived experience

Whilst definitions and measures of poverty are often relative to the standards prevailing in society, the existence of poverty is frequently understood as ontologically independent. It has therefore been argued that objectivity in poverty and vulnerability research occurs when ‘a definition is independent of the feelings of poor people’ (Roll, 1992: 12). Whilst, a significant body of research has explored subjective perceptions of poverty (Danziger et al., 2000; Dunlap et al., 2003; McIntyre et al., 2003; Saunders et al., 2006; Hooper, 2007; Ridge, 2011), these studies are generally recognised as capturing qualitative experiences (rather than quantifying the nature) of poverty. These studies often

operate within epistemological confines that restrict their capacity to explore the relational features that are constitutive of poverty (Harriss, 2007). As a result, the phenomenological significance of deprivation has often been overlooked in mainstream poverty research (Charlesworth, 2000; Van der Merwe, 2006).

Psychosocial studies and wellbeing research,³ by virtue of their inter-disciplinarity, seek to 'rethink some key popular and disciplinary constructions of the welfare subject and the scope of social policy'⁴ (Stenner and Taylor, 2008: 443). Affective states that were once considered part of the private or emotional sphere are now being attended to in comparative poverty research to understand how institutional arrangements can affect individual experience and behaviour across a range of contexts in similar ways. Underlying research in this area is a basic recognition that affective states cannot be understood in isolation from the circumstances in which they arise. For example, in response to being bitten by a dog, a child may violently retaliate, cry, run away or do nothing. Irrespective of the (in-)action taken, the child's affective state, experience and action can only be understood in relation to the actions of the dog and the broader context within which they arise. Similarly, the affective states and experiences of poor people are significantly independent of socio-economic structures – but their meaning, cause and consequence can only be fully understood in relation to the social policies and market forces that shape contextual and material circumstance. Recent work on the psychology of scarcity demonstrates this point well. Mani et al. (2013) find that compared to better-off individuals, worse-off individuals are more likely to make bad decisions and mistakes. A reasonable conclusion may be that individuals are worse off as a result of their poor decision-making. However, Mani et al. (2013) demonstrate that scarcity arising from socio-political and economic systems greatly affects an individual's psychological capacity for good decision-making. Arguably, lived experiences are an objective reflection of institutional and structural arrangements as much as they are an individualised affective state or subjective perception. The fact that subjective perceptions and lived experience cannot be fully understood in isolation from structural or relational schema does not undermine their objective significance. Indeed this inter-dependency demonstrates the importance of theorising the relationship between subjective experience, material circumstance and institutional arrangements. This produces new forms of poverty knowledge that is 'more capable of taking account of the relational complexities of subjectivity' (Taylor, 2011: 789).

As previously stated, subjective perceptions do not always correspond with material circumstances or social positions (Flaherty, 2008; Shildrick and MacDonald, 2013) as:

people who are poor try not to admit it to themselves exactly because the admission of 'failure' adds to the misery they feel from their objective situation. Thus in interviews poor people, lower class people, are often able to characterise the situation of others like themselves in lively detail, but they often seek to distance themselves personally from that common plight. (Rainwater, 1990: 3)

Findings such as this do not undermine the credibility of subjective perceptions and lived experiences in poverty research. Neither do they necessarily represent a discord between the structural and affective realities of deprivation. As the quote above illustrates, poverty denial could very well be an articulation of the relational nature of poverty. Reference groups, cultural praxes, normative frameworks, agentic choice and other complex structural determinants of experience will shape affective responses to and lived experiences of material hardship. By exploring how people navigate and give meaning to a social, political and economic system, it is possible to see how the substantive and figurative significance of social citizenship is undermined for those experiencing deprivation. Citizenship is by no means the only framework through which to explore the significance of lived experience, but it does proffer three inter-related conceptual benefits. Firstly, the 'equality of status' premise of citizenship makes the analysis of structural inequalities meaningfully independent of normative critiques of poverty and vulnerability. Secondly, citizenship is based on principles of common membership that govern the relations between polity, state and market. In this way, poverty can be understood not as a peculiarity of market processes but as a systemic feature of citizenship arrangements. Finally, despite differences across domestic and international settings, citizenship offers a common conceptual and analytical framework through which to explore the distributional effects of institutions within and across national boundaries.

Grounding poverty analysis and debate in lived experiences can be transformative for those facing poverty and those seeking critical and efficacious policy solutions to address it (Dutro, 2009). One challenge is to make the rich micro-level data that is drawn from such research 'policy relevant' (Brock, 1999). The structural macro-level data that tends to treat the poor as a unit of observation needs to be refined and

complemented by poverty definitions, experiences and solutions that originate from poor people themselves (Marti and Mertens, 2014). By integrating disparate paradigms and methods of poverty research, it is possible to develop new understanding about the existence and nature of poverty, as well as effective poverty prevention and alleviation strategies.

Lived experiences of deprivation: An integrated mixed methods approach

This section considers how an integrated mixed methods approach can add credibility to the lived experiences of deprivation. This approach could help transcend pre-existing epistemic and disciplinary boundaries in a way that makes lived experiences tractable with institutional and non-institutional responses to poverty. In particular, this section considers how the novel use of mixed methods can make use of qualitative data to analyse socio-economic and political structures bound up in social citizenship (e.g. Dwyer, 2002).

There is often a false distinction between quantitative and qualitative methodologies. Quantitative research can adopt interpretivist approaches, just as qualitative research can take a realist approach (Bryman, 1984). Nevertheless, there are epistemological and ontological tensions that arise between the two paradigms of poverty research. These tensions are symptomatic of the particular strengths and limitations inherent in purely quantitative or qualitative methods.

Quantitative poverty research tends to focus on the measurement and extent of poverty within and across contexts. In many ways, quantitative studies are best suited to capture the structural manifestations and, at times, the dynamics of poverty. As a result, quantitative methods have a propensity of 'imposing analytical categories with little local relevance' on those experiencing disadvantage (Shaffer, 2013: 270). Indicators such as income and consumption have been used because they can be most readily compared but this raises a number of problems (Chambers, 1995). It goes without saying that there is great diversity in patterns of social differentiation, resource use, production and consumption across and within socio-economic and political communities. Overly positivistic studies of poverty fail to capture the multiplicity of ways in which poverty can be interpreted, understood and felt by those living with 'minimal' resources (Halleröd, 1996). The affective attachments and cultural images bound up in poverty are often coarsely misrepresented or absent in the interpretations of data from surveys or large-N poverty

studies. As a result, the features of significance in poverty measurement and research are regularly constructed with a western bias that impedes the local relevance of quantitative measures (Grech, 2009). Irrespective of the challenges faced in measuring poverty consistently and accurately, these studies run the danger of de-politicising the nature of poverty and the 'cures' deemed desirable or even feasible (Green and Hulme, 2005; O'Connor, 2009). Worse still, they potentially subjugate certain forms of knowledge over others and prioritise elite views of poverty rather than the views of those subject to it (Harriss, 2007).

Many qualitative studies attempt to address this imbalance. Participatory poverty research 'seeks to understand the experience and causes of poverty from the perspective of the poor themselves' (Robb, 2002: xi). Such methods have become increasingly prevalent in international development since the 1980s. One challenge is to make the rich micro-level data from participatory poverty research empirically compatible with policy responses. This is a challenge faced by most qualitative poverty research that is grounded in the voices and experiences of the poor. Resource requirements, experiences and behaviours differ within and across social groups. This makes it difficult to consistently and systematically measure poverty across contexts from purely qualitative research. For this reason, the use of qualitative data is often tokenistic, operating within the confines of existing policy paradigms, objectives and socio-structural inequalities (Cornwall, 2003). As a result, the extent to which qualitative research on lived experiences can be 'empowering' or 'transformative' has been questioned (Norton and Stephens, 1995; O'Meally, 2014). Qualitative researchers often claim that their methods are innovative and contribute something new to understanding poverty. However, it is not so much the methods of poverty research that have the capacity to be transformative, rather it is how the research design is formulated and to what ends 'the voices of the poor' are put (Brannen, 2005).

There is a tendency in some qualitative poverty research to focus on the effects of poverty without reflecting on its causes (e.g. McIntyre et al., 2003). This limits the utility of some qualitative poverty research because it does not fully capture 'the structures and relationships that give rise to the effects believed to define poverty' (Harriss, 2007: 2). In addition it can lead to development strategies and institutional practices that can have the perverse effect of propagating inequalities – principally because they fail to attend to power relationships and the structurally embedded nature of poverty (Ansell, 2014; O'Meally, 2014). This chapter suggests that, through effective integration of mixed methods, lived

experiences of deprivation offer something of more instrumental worth for social and development policy. The conceptual, structural and relational schema that shape poverty are grounded and reflected in lived experiences. As such, subjective perceptions of deprivation must be integrated into poverty research in a way that reflects both their ontological and epistemological values. O'Connor (2009) argues that it is necessary to reconsider the purpose and function of poverty research so that it captures and thus speaks to the structural causes and manifestations of inequality. She suggests that effective poverty research must

redefine the conceptual basis for poverty knowledge, above all by shifting the analytic framework from its current narrow focus on explaining individual deprivation to a more systemic and structural focus on explaining – and addressing – inequalities in the distribution of power, wealth, and opportunity. (O'Connor, 2009: 22)

Research that studies the poor must acknowledge that this is not the same thing as studying poverty (O'Connor, 2009). Detaching experiences and effects from the causes of poverty renders poverty measurement imprecise and policy responses ineffectual. By looking at the relationship between deprivation and social citizenship, the lived experiences of those subject to poverty offer significant explanatory power and insight into the causes and consequences of material hardship. In doing so, it becomes possible to connect macro socio-economic and political structures with micro processes and factors shaping the extent and nature of poverty (Carney et al., 1999).

My own study could have exclusively drawn upon quantitative methods to establish the prevalence of income poverty, unemployment and area deprivation. However, looking at the correlates of deprivation in isolation, or even in conjunction, says very little about the nature of poverty and the relational determinants of its existence and experience. Drawing upon a representative quantitative survey to explore how the aforementioned selection criteria coalesce to affect support for social citizenship and the extent to which people feel they have social rights does go some way to explore the nexus between deprivation and social citizenship. Whilst it captures something about the probabilistic determinants of difference in attitudes, experiences and outcomes, it fails to offer an explanation for this difference and the multiple ways in which it may occur. Without qualitative research with those in similar circumstances, it becomes impossible to explore the significance or meaning of quantitative differences observed in the first phase of the study.

As previously stated, qualitative interviews explored the lived experiences of deprivation in relation to structural and institutional schema. The data generated were used in conjunction with quantitative data and put towards a number of different ends. As Brannen (2005: 12) argues, 'data collected from different methods cannot simply be added together to produce a unitary or rounded reality or truth'.

The integration of mixed methods and data can serve a number of different purposes; and in this instance, lived experiences of deprivation were drawn upon for three key reasons. Firstly, exploring the lived experiences of deprivation helped partially explain attitudinal differences towards conceptions of social citizenship. For example, it was found that deprivation reduced support for certain social and economic citizenship rights. By exploring lived experiences of deprivation, it was possible to explain why these differences in attitudes occurred and what these indicated about the relationship between deprivation and citizenship. Secondly, subjective perceptions about the material and symbolic significance of deprivation were able to capture the relational nature of poverty. By attending to the relational and systemic features of poverty, it was possible to research lived experiences of deprivation in the UK but also replicate the study in New Zealand in a systematic way. Finally, and as a result, the findings of the study helped provide empirical support to citizenship theories about the relationship between deprivation and social citizenship (Lister, 2004).

Without sufficient integration of mixed methods, this would have not been possible. Integrated mixed methods help avoid a mismatch between categories and aspects of poverty measured, and facilitate a 'better understanding of the transition matrix in the study of poverty dynamics and greater appreciation of causal variables, and their interaction' (Shaffer, 2013: 280). Particularly in my study, integration of mixed methods facilitated an effective examination of the relational nature of poverty. Whilst integrated mixed methods add credibility to lived experiences, I would argue that this is primarily for applied rather than epistemic reasons. To say that qualitative research findings gain credibility as a result of a mixed methods approach undermines the objective significance and utility of lived experience in understanding poverty. As this chapter has suggested, lived experiences can proffer unique insight into 'the social and political-economic relationships that trigger poverty' (Harriss, 2007: 13). Indeed, lived experiences could, in this instance, be said to challenge the internal validity of certain quantitative poverty measures. Having said that, integrated mixed methods can equally strengthen the external validity of qualitative research findings.

The use of mixed methods helps contribute towards the instrumental use of qualitative research – particularly research that is grounded in or at least informed by lived experiences of deprivation. Due to their transgressive nature, qualitative research findings that challenge the dominant poverty and policy paradigm are often not compatible with existing forms of poverty knowledge. As a result, there may well be tensions that arise in definitions of poverty deemed appropriate and features of poverty deemed significant. Equally, explanations of poverty and policy recommendations derived from lived experiences are less likely to be compatible with existing institutional and structural mechanisms designed to tackle poverty. Institutional action depends on pre-existing poverty knowledge (Bradshaw, 2006) and as a result new forms of poverty knowledge that challenge the status quo are less likely to be compatible with frameworks of action. Mixed methods research is faced with the challenge of making lived experiences of deprivation epistemologically compatible with the existing poverty paradigm. In turn, this can inform innovative poverty action. Whilst this approach has the potential of being transformative in the long-term, it is necessarily iterative in the short-term. By moderating existing paradigms of poverty knowledge through mixed methods, lived experiences can come closer to affecting progressive social debate and action. Combining new forms of poverty knowledge with quantitative methods enhances the credibility and usability of research findings. These findings may then be able to more constructively engage with policy debates, and potentially shift the terms of deliberation in a direction that more effectively tackles poverty (e.g. Zea et al., 2014). If mixed methods poverty research is to be transformative, researchers have a ‘responsibility of creating novel scientific knowledge that do not just progress in the diagnosis of the social problems but which also inform policies or actions aimed at reversing them’ (Marti and Mertens, 2014: 2). I would argue that integrating lived experiences of deprivation into a mixed methods approach is one way of achieving this.

Conclusion

A number of studies have combined objective and subjective measures to arrive at a definition of poverty (Mack and Lansley, 1985; Deleeck and Van den Bosch, 1992; Eroğlu, 2007). However, fewer studies have accounted for the objective significance of lived experience. That is, qualitative and mixed methods studies of poverty often fail to see that

lived experiences and subjective perceptions are a manifestation and reflection of structural relations (e.g. Novak, 1996). Such an approach would offer new insights into measuring poverty and vulnerability but most importantly it would also offer a new methodological framework through which to assess the efficacy of poverty alleviation processes in a comparative context. Rather than assessing the efficacy of welfare regimes on their own terms (i.e. via different structural welfare arrangements), the lived experience of citizens helps uncover the extent to which the dominant praxis of social citizenship operates effectively to tackle or at least temper poverty and vulnerability. These benefits can only be reaped by integrating mixed methods to overcome the limitations of purely quantitative or qualitative research. In this sense, qualitative and quantitative methods

are fundamentally complements, not substitutes – and certainly not rivals. They mutually inform each other, to everyone’s benefit. . . . Developing powerful and effective diagnoses of the causes of poverty, and appropriate treatments to reduce poverty, requires both well-designed quantitative investigation and giving a genuine voice to poor people. (Robb, 2002: xii–xiii)

Mixed methods can strengthen the credibility and external validity of lived experiences. Whilst lived experiences tend to be part of the qualitative paradigm, this chapter has argued that researching lived experiences of deprivation could be seen as a methodology in its own right. A number of recent studies have made significant contributions in this regard (e.g. Chase and Kyomuhendo, 2014). Understanding the causes and consequences of poverty is contingent on capturing the socio-economic and political relations that shape the nature of poverty. A methodological approach that can examine these relations goes some way towards an ‘objective’ measure that captures the complexities of poverty in a systematic way. In spite of significant differences within and across social groups, this methodology makes it possible to consistently measure and explore poverty across a range of contexts. This is exactly because the organisational variations across different contexts give rise to very different types and thus experiences of poverty. Poverty prevention and alleviation strategies that incorporate such an approach have the potential to be more effective because they account for the mitigating and propagating determinants of poverty.

Notes

- 1 These systemic processes extend well beyond the social realm of citizenship to civil and political domains as well. The terms ‘citizenship’ and ‘social citizenship’ are therefore used interchangeably henceforth to reflect the inherently ‘social’ nature of all citizenship practices.
- 2 Of course, Rowntree went further to develop the notion of ‘secondary poverty’ to suggest that there were a group of people whose basic needs were not being met due to behaviour or lifestyle choices.
- 3 For example, Chase and Kyomuhendo (2014), Walker (2014), Tomlinson et al. (2008) and Goffman (1963).
- 4 Within this context, ‘welfare subject’ refers to an individual affected by but also psychologically responding to structural welfare arrangements.

References

- Ansell, N. 2014. Challenging Empowerment: AIDS-affected Southern African Children and the Need for a Multi-level Relational Approach. *Journal of Health Psychology*, 19, 22–33.
- Atkinson, A. B. 1989. *Poverty and Social Security*. Hemel Hempstead: Harvester Wheatsheaf.
- Beresford, P., Green, D., Lister, R., and Woodard, K. 1999. *Poverty First Hand: Poor People Speak for Themselves*. London: CPAG.
- Booth, C. 1903. *Life and Labour of the People in London*. London: Macmillan.
- Bourdieu, P. 1977. *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.
- Bradshaw, J., Middleton, S., Davis, A., Oldfield, N., Smith, N., Cusworth, L., and Williams, J. 2008. *A Minimum Income Standard for Britain: What People Think*. York: Joseph Rowntree Foundation.
- Bradshaw, T. 2006. Theories of Poverty and Anti-Poverty Programs in Community Development. *RPRC Working Paper No. 06-05*. Columbia: Rural Poverty Research Centre.
- Brannen, J. 2005. *Mixed Methods Research: A Discussion Paper*. ESRC National Centre for Research Methods.
- Brock, K. 1999. *It's Not Only Wealth that Matters – It's Peace of Mind Too: A Review of Participatory Work on Poverty and Illbeing*. Washington, DC: World Bank.
- Brooks, C., Nieuwbeerta, P., and Manza, J. 2006. Cleavage-based Voting Behavior in Cross-national Perspective: Evidence from Six Postwar Democracies. *Social Science Research*, 35, 88–128.
- Brooks, C. and Svallfors, S. 2010. Why Does Class Matter? Policy Attitudes, Mechanisms, and the Case of the Nordic Countries. *Research in Social Stratification and Mobility*, 28, 199–213.
- Bryman, A. 1984. The Debate About Quantitative and Qualitative Research: A Question of Method or Epistemology?. *British Journal of Sociology*, 35, 75–92.
- Campbell, A. L. 2003. *How Policies Make Citizens: Senior Political Activism and the American Welfare State*. Princeton: Princeton University Press.
- Carney, D., Drinkwater, M., Rusinow, T., Neeffjes, K., Wanmali, S., and Singh, N. 1999. A Brief Comparison of the Livelihoods Approaches of the UK Department

- for International Development (DFID), CARE, Oxfam and the United Nations Development Programme (UNDP). London: Department for International Development.
- Chambers, R. 1995. Poverty and Livelihoods: Whose Reality Counts?. *Environment and Urbanization*, 7, 173–204.
- Charlesworth, S. J. 2000. *A Phenomenology of Working-class Experience*. Cambridge: Cambridge University Press.
- Chase, E. and Kyomuhendo, G. (eds) 2014. *Poverty and Shame: Global Experiences*. Oxford: Oxford University Press.
- Chase, E. and Walker, R. 2013. The Co-construction of Shame in the Context of Poverty: Beyond a Threat to the Social Bond. *Sociology*, 47, 739–754.
- Cohen, S., Alper, C. M., Doyle, W. J., Adler, N., Treanor, J. J., and Turner, R. B. 2008. Objective and Subjective Socioeconomic Status and Susceptibility to the Common Cold. *Health Psychology*, 27, 268.
- Cornwall, A. 2003. Whose Voices? Whose Choices? Reflections on Gender and Participatory Development. *World Development*, 31, 1325–1342.
- Danziger, S., Corcoran, M., Danziger, S., and Heflin, C. M. 2000. Work, Income, and Material Hardship after Welfare Reform. *Journal of Consumer Affairs*, 34, 6–30.
- DCLG. 2011. *The English Indices of Deprivation 2010*. Department for Communities and Local Government. London: HMSO.
- de Botton, A. 2004. *Status Anxiety*. London: Hamish Hamilton (Penguin).
- Dean, H. and Melrose, M. 1996. Unravelling Citizenship: The Significance of Social Security Benefit Fraud. *Critical Social Policy*, 16, 3–31.
- Deleeck, H. and Van den Bosch, K. 1992. Poverty and Adequacy of Social Security in Europe: A Comparative Analysis. *Journal of European Social Policy*, 2, 107–120.
- Dittmar, H. and Pepper, L. 1994. To Have Is to Be: Materialism and Person Perception in Working-class and Middle-class British Adolescents. *Journal of Economic Psychology*, 15, 233–251.
- Doyal, L. and Gough, I. 1991. *A Theory of Human Need*. London: Palgrave Macmillan.
- Dunlap, E., Golub, A., and Johnson, B. D. 2003. The Lived Experience of Welfare Reform in Drug-Using Welfare-Needy Households in Inner-City New York. *Journal of Sociology & Social Welfare*, 30, 39–58.
- Dutro, E. 2009. Children Writing “Hard Times”: Lived Experiences of Poverty and the Class-Privileged Assumptions of a Mandated Curriculum. *Language Arts*, 87, 89–98.
- Dwyer, P. 1998. Conditional Citizens? Welfare Rights and Responsibilities in the Late 1990s. *Critical Social Policy*, 18, 493–517.
- Dwyer, P. 2002. Making Sense of Social Citizenship: Some User Views on Welfare Rights and Responsibilities. *Critical Social Policy*, 22, 273–299.
- Edmiston, D. 2015. *The Poverty and Riches of Social Citizenship in the UK: How Lived Experiences Affect Attitudes Towards Welfare, Rights and Responsibilities*. PhD Sociology and Social Policy, University of Leeds.
- Eroğlu, S. 2007. Developing an Index of Deprivation Which Integrates Objective and Subjective Dimensions: Extending the Work of Townsend, Mack and Lansley, and Halleröd. *Social Indicators Research*, 80, 493–510.
- Evans, G. 1993. Class, Prospects and the Life-cycle: Explaining the Association between Class Position and Political Preferences. *Acta Sociologica*, 36, 263–276.

- Evans, G. and de Graaf, N. D. 2013. *Political Choice Matters: Explaining the Strength of Class and Religious Cleavages in Cross-national Perspective*. Oxford: Oxford University Press.
- Evans, M. D. and Kelley, J. 2004. Subjective Social Location: Data from 21 Nations. *International Journal of Public Opinion Research*, 16, 3–38.
- Ferragina, E., Feyertag, J., and Seeleib-Kaiser, M. 2014. Economic and Social Outsiderism among Young People in the EU. In *Social Policy Association Annual Conference*, Sheffield.
- Ferragina, E., Tomlinson, M., and Walker, R. 2013. *Poverty, Participation and Choice – The Legacy of Peter Townsend*. York: Joseph Rowntree Foundation.
- Flaherty, J. 2008. *Getting by, Getting Heard: Poverty and Social Exclusion in the Borders. Listening to the Voices of Experience*. Glasgow: Poverty Alliance.
- Giddens, A. 1991. *Modernity and Self Identity: Self and Society in the Late Modern Age*. Cambridge: Polity Press.
- Goffman, E. 1963. *Stigma: Notes on the Management of Spoiled Identity*. Englewood Cliffs, NJ: Prentice Hall.
- Goodman, E., Adler, N. E., Daniels, S. R., Morrison, J. A., Slap, G. B., and Dolan, L. M. 2003. Impact of Objective and Subjective Social Status on Obesity in a Biracial Cohort of Adolescents. *Obesity Research*, 11, 1018–1026.
- Goodman, E., McEwen, B. S., Dolan, L. M., Schafer-Kalkhoff, T., and Adler, N. E. 2005. Social Disadvantage and Adolescent Stress. *Journal of Adolescent Health*, 37, 484–492.
- Gordon, D. and Pantazis, C. 1997. Measuring Poverty: Breadline Britain in the 1990s. In Gordon, D. and Pantazis, C. (eds), *Breadline Britain in the 1990s*. Aldershot: Ashgate.
- Grech, S. 2009. Disability, Poverty and Development: Critical Reflections on the Majority World Debate. *Disability & Society*, 24, 771–784.
- Green, M. 2006. Representing Poverty and Attacking Representations: Perspectives on Poverty from Social Anthropology. *The Journal of Development Studies*, 42, 1108–1129.
- Green, M. and Hulme, D. 2005. From Correlates and Characteristics to Causes: Thinking about Poverty from a Chronic Poverty Perspective. *World Development*, 33, 867–879.
- Halleröd, B. 1996. Deprivation and Poverty: A Comparative Analysis of Sweden and Great Britain. *Acta Sociologica*, 39, 141–168.
- Halleröd, B., Bradshaw, J., and Holmes, H. 1997. Adapting the Consensual Definition of Poverty. In Gordon, D. and Pantazis, C. (eds), *Breadline Britain in the 1990s*. Aldershot: Ashgate.
- Hamilton, C. 2003. Overconsumption in Britain: A Culture of Middle Class Complaint? *Discussion Paper No. 57*. Canberra: The Australia Institute.
- Harriss, J. 2007. Bringing Politics Back into Poverty Analysis: Why Understanding Social Relations Matters More for Policy on Chronic Poverty than Measurement. *Chronic Poverty Research Centre Working Paper No. 77*. <http://www.chronicpoverty.org/>
- Hooper, C.-A. 2007. *Living with Hardship 24/7: The Diverse Experiences of Families in Poverty in England*. London: Frank Buttle Trust.
- Hulme, D. and McKay, A. 2005. Identifying and Measuring Chronic Poverty: Beyond Monetary Measures. *CPRC-IIPA Working Paper 30*. New Delhi: Indian Institute of Public Administration.

- Humpage, L. 2008. Talking About Citizenship in New Zealand. *Kōtuitui: New Zealand Journal of Social Sciences Online*, 3, 121–134.
- Jo, Y. N. 2013. Psycho-social Dimensions of Poverty: When Poverty Becomes Shameful. *Critical Social Policy*, 33, 514–531.
- Jordan, B. 1996. *A Theory of Poverty and Social Exclusion*. Cambridge: Polity Press.
- Joseph, K. and Sumption, J. 1979. *Equality*. London: J. Murray.
- Kraus, M. W., Piff, P. K., Mendoza-Denton, R., Rheinschmidt, M. L., and Keltner, D. 2012. Social Class, Solipsism, and Contextualism: How the Rich are Different from the Poor. *Psychological Review*, 119, 546.
- Lister, R. 1990. *The Exclusive Society: Citizenship and the Poor*. London: CPAG.
- Lister, R. 2002. Work for Those Who Can, Security for Those Who Cannot: A Third Way in Social Security Reform or Fractured Social Citizenship?. In Edwards, R. and Glover, J. (eds), *Risk and Citizenship: Key Issues in Welfare*. London: Routledge.
- Lister, R. 2004. *Poverty: Applied Ethics and Social Problems*. Cambridge: Polity Press.
- Lister, R., Smith, N., Middleton, S., and Cox, L. 2003. Young People Talk About Citizenship: Empirical Perspectives on Theoretical and Political Debates. *Citizenship Studies*, 7, 235–253.
- MacDonald, R., Shildrick, T., Webster, C., and Simpson, D. 2005. Growing Up in Poor Neighbourhoods the Significance of Class and Place in the Extended Transitions of ‘Socially Excluded’ Young Adults. *Sociology*, 39, 873–891.
- Mack, J. and Lansley, S. 1985. *Poor Britain*. London: Allen & Unwin.
- Mani, A., Mullainathan, S., Shafrir, E., and Zhao, J. 2013. Poverty Impedes Cognitive Function. *Science*, 341, 976–980.
- Marshall, T. H. 1950. *Citizenship and Social Class: And Other Essays*. Cambridge: Cambridge University Press.
- Marti, T. S. and Mertens, D. M. 2014. Mixed Methods Research with Groups at Risk New Developments and Key Debates. *Journal of Mixed Methods Research*, 8, 207–211.
- McIntyre, L., Officer, S., and Robinson, L. M. 2003. Feeling Poor: The Felt Experience Low-income Lone Mothers. *Affilia*, 18, 316–331.
- Mettler, S. and Soss, J. 2004. The Consequences of Public Policy for Democratic Citizenship: Bridging Policy Studies and Mass Politics. *Perspectives on Politics*, 2, 55–73.
- Millar, J. and Glendinning, C. 1987. *Invisible Women, Invisible Poverty*. Hemel Hempstead: Harvester Wheatsheaf.
- Miyakawa, M., Hanson, L. L. M., Theorell, T., and Westerlund, H. 2012. Subjective Social Status: Its Determinants and Association with Health in the Swedish Working Population (the SLOSH Study). *The European Journal of Public Health*, 22, 593–597.
- Norton, A. and Stephens, T. 1995. Participation in Poverty Assessments. *Social Policy and Resettlement Division Discussion Paper*. Washington, DC: World Bank.
- Novak, M. 1996. Poverty: Facts and Feelings. *Družboslovne razprave (Sociological Essays)*, 12, 84–98.
- O’Connor, A. 2009. *Poverty Knowledge: Social Science, Social Policy, and the Poor in Twentieth-century US History*. Princeton: Princeton University Press.
- O’Meally, S. 2014. The Contradictions of Pro-poor Participation and Empowerment: The World Bank in East Africa. *Development and Change*, 45, 1248–1283.

- Orton, M. 2009. Understanding the Exercise of Agency within Structural Inequality: The Case of Personal Debt. *Social Policy and Society*, 8, 487–498.
- Palmer, G., Macinnes, T., and Kenway, P. 2008. *Monitoring Poverty and Social Exclusion 2008*. York: Joseph Rowntree Foundation.
- Piachaud, D. 1981. Peter Townsend and the Holy Grail. *New Society*, 10 September, pp. 419–421.
- Pradhan, M. and Ravallion, M. 2000. Measuring Poverty Using Qualitative Perceptions of Consumption Adequacy. *Review of Economics and Statistics*, 82, 462–471.
- Rainwater, L. 1990. Poverty and Equivalence as Social Construction. *The Luxembourg Income Study Working Paper No. 55*. Luxembourg: CEPS/INSTEAD.
- Ridge, T. 2011. The Everyday Costs of Poverty in Childhood: A Review of Qualitative Research Exploring the Lives and Experiences of Low-income Children in the UK. *Children & Society*, 25, 73–84.
- Ritterman, M. L., Fernald, L. C., Ozer, E. J., Adler, N. E., Gutierrez, J. P., and Syme, S. L. 2009. Objective and Subjective Social Class Gradients for Substance Use among Mexican Adolescents. *Social Science & Medicine*, 68, 1843–1851.
- Robb, C. M. 2002. *Can the Poor Influence Policy?: Participatory Poverty Assessments in the Developing World*. Washington, DC: International Monetary Fund.
- Roll, J. 1992. *Understanding Poverty: A Guide to the Concepts and Measures*, Family Policy Studies Centre Occasional Paper 15. London: Family Policy Studies Centre.
- Roseneil, S. 2014. On Meeting Linda: An Intimate Encounter with (Not-)belonging in the Current Conjunction. *Psychoanalysis, Culture & Society*, 19, 19–28.
- Rowntree, B. S. 1901. *Poverty: A Study of Town Life*. London: Macmillan.
- Runciman, W. G. 1966. *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England*. London: Routledge & K. Paul.
- Saunders, P., Sutherland, K., Davidson, P., Hampshire, A., King, S., Taylor, J., Australia, M., and Anglicare, N. 2006. Experiencing Poverty – The Voices of Low-income Australians: Towards New Indicators of Disadvantage Project. *Stage 1: Focus Group Outcomes*. Social Policy Research Centre, University of New South Wales.
- Scanlon, C. and Adlam, J. 2013. Knowing Your Place and Minding Your Own Business: On Perverse Psychological Solutions to the Imagined Problem of Social Exclusion. *Ethics and Social Welfare*, 7, 170–183.
- Sen, A. 1999. *Commodities and Capabilities*. Oxford: Oxford University Press.
- Sen, A. 2002. Health: Perception versus Observation: Self-reported Morbidity has Severe Limitations and Can Be Extremely Misleading. *British Medical Journal*, 324, 860.
- Shaffer, P. 2013. Ten Years of “Q-Squared”: Implications for Understanding and Explaining Poverty. *World Development*, 45, 269–285.
- Shildrick, T. and MacDonald, R. 2013. Poverty Talk: How People Experiencing Poverty Deny Their Poverty and Why They Blame ‘The Poor’. *The Sociological Review*, 61, 285–303.
- Stenner, P. and Taylor, D. 2008. Psychosocial Welfare: Reflections on an Emerging Field. *Critical Social Policy*, 28, 415–437.
- Taylor, D. 2011. Wellbeing and Welfare: A Psychosocial Analysis of Being Well and Doing Well Enough. *Journal of Social Policy*, 40, 777–794.

- Tomlinson, M., Walker, R., and Williams, G. 2008. Measuring Poverty in Britain as a Multi-dimensional Concept, 1991 to 2003. *Journal of Social Policy*, 37, 597–620.
- Townsend, P. 1979. *Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living*. Harmondsworth: Penguin.
- Toynbee, P. and Walker, D. 2009. *Unjust Rewards: Ending the Greed that is Bankrupting Britain*. London: Granta Books.
- Van der Merwe, K. 2006. The Phenomenology of Experiencing Poverty – An Exploration. *TD: The Journal for Transdisciplinary Research in Southern Africa*, 2, 131–144.
- Van Praag, B., Goedhart, T., and Kapteyn, A. 1980. The Poverty Line – A Pilot Survey in Europe. *The Review of Economics and Statistics*, 62, 461–465.
- Vizard, P. 2010. What Do the Public Think about Economic and Social Rights? Research Report to Inform the Debate about a Bill of Rights and a Written Constitution. *CASE Report 61*. Centre for the Analysis of Social Exclusion, London School of Economics.
- Wacquant, L. 2009. *Punishing the Poor: The Neoliberal Government of Social Insecurity*. Durham, NC: Duke University Press.
- Walker, R. 2014. *The Shame of Poverty*. Oxford: Oxford University Press.
- Zea, M. C., Aguilar-Pardo, M., Betancourt, F., Reisen, C. A., and Gonzales, F. 2014. Mixed Methods Research with Internally Displaced Colombian Gay and Bisexual Men and Transwomen. *Journal of Mixed Methods Research*, 8, 212–221.

4

Bringing Context to Poverty in Rural Rwanda: Added Value and Challenges of Mixed Methods Approaches¹

Neil Dawson

Introduction

Immense efforts and funds are directed towards trying to improve the lives of those in the world's least developed countries. Yet, as current arguments surrounding the Millennium Villages Project highlight (Sanchez et al. 2007, Wilson 2013), the methods for achieving and measuring the desired poverty alleviation impacts are subjects of intense debate. There are different standpoints about how poverty should be conceptualised, the pathways through which the lives of the world's poorest may be improved and how changes in their lives should be measured or assessed. Rwanda, a small landlocked state in Sub-Saharan Africa, provides a fascinating example of different perspectives on poverty and poverty reduction policies in developing countries. While politicians, practitioners and researchers hail Rwanda's progress in meeting national socio-economic targets as representing a model of development (IMF 2011, UN 2013), a smaller group of academics and activists take a quite polar stance (see, for example, Straus and Waldorf 2011). While national indicators show a rapidly declining incidence of poverty, studies exploring the control that individuals are able to exercise over land and property, their individual freedoms and ability to participate in decisions affecting them paint a picture of a very contrasting trajectory in their wellbeing (Des Forges 2005, Pritchard 2013).

Using Rwanda as a case study, and employing both qualitative and quantitative analyses based on household interview data, this chapter aims to: (a) show why and how conflicting standpoints about trajectories in poverty and wellbeing may occur; (b) critically analyse some of the different types of poverty analyses utilised to characterise trends and inform policy interventions; (c) present a mixed methods study

applying a holistic wellbeing approach to the lives of rural Rwandans; and finally (d) discuss some of the added and complementary values of applying such an approach, while also addressing some of the challenges and obstacles which must be overcome in bringing mixed methods approaches into common usage for the design and evaluation of development policy. Particular attention will be paid to issues of positionality and integration of methods at the time of data collection.

Bridging the divide in poverty research

Much of the disagreement about how to tackle issues of poverty stems from the lens through which people's lives are viewed and the methods used to describe them. Policymakers and analysts tend to seek consistent poverty indicators so that regions and time points may be compared with one another. Yet development policy may be most successful when paying attention to local knowledge and practices as well as more material or wider economic concerns (Tendler 1997). Contrasting approaches, often promoted by more locally focused actors, therefore favour attention to contextual issues. Centuries of academic debate has drawn simplistic pictures of two polar incompatible fields in academic research, often boiled rather crudely down to a question of methods: the use of quantitative methods to measure, classify and statistically analyse; and the use of qualitative methods to engage with people and interpret their perceptions, a debate sometimes referred to as 'the paradigm wars' (Kuhn 1970, Kanbur and Shaffer 2007). A similar divide can also be considered to occur in poverty research. The difference between the two approaches is not trivial as the implications to be drawn from them about poverty and its causes, even within the same population, can be quite different (Pogge and Reddy 2005, Camfield and Ruta 2007).

Hearteningly, these polarised approaches to poverty and views about policies to alleviate it are not beyond reconciliation. Both offer insights into the different dimensions of an individual's wellbeing, which are entirely compatible and complementary rather than contradictory. Mixed methods approaches have been widely recognised for offering opportunities to advance poverty methodologies and provide comprehensive policy recommendations (Bamberger et al. 2010, Stern et al. 2012). However, their fusion is not a straightforward exercise and obstacles exist to bringing mixed methods research into mainstream academic research and development practice, particularly regarding the view that they may compromise the consistency and comparability so valued in standard poverty measurement studies (Brannen 2005).

Rather than labelling the distinction between approaches to researching poverty as a simple question of methods, quantitative versus qualitative (a distinction which does not hold in all circumstances), the divide is represented here as being more nuanced based on research objectives and underlying methodologies. This chapter will instead distinguish between approaches aiming to achieve consistency of measurement and those seeking to reveal specificity of context (Ravallion 1996).

To achieve consistency of measurement and therefore comparability between times and places, poverty is represented through indicators considered to be reliably observed and recorded. The most commonly employed measurement of poverty is the consumption method through which data is used to establish whether or not individual households have the resources to buy a set basket of goods enabling them to acquire sufficient food to meet their basic needs. And if not, those households may be considered to be below the poverty threshold. This approach indisputably dominates poverty assessment among states and large development institutions (McKay and Lawson 2003). The consumption approach provides a consistent measure which can be repeated in different settings at different times to allow comparison to other social and geographic contexts and past and future time periods (Ravallion 1996). Although the prices of the bundle of goods required for achieving those functionings is adjusted to reflect local economic conditions, as a measure of poverty it seeks to maximise consistency at the expense of specificity (*ibid*). Broadly, such an approach assumes that the reality of an individual's poverty, and a person's capabilities and functionings, can be observed and recorded objectively.

In terms of Rwanda, trends in consumption poverty reveal great advances in poverty alleviation. Rwanda faces challenges in its development which include acidic soils lacking in nutrients and organic matter, the highest population density in mainland Africa and high dependence on subsistence agriculture (REMA 2009). Partially to combat these problems, Rwandans have been exposed to internationally supported far reaching development policies to promote economic development, rural modernisation, land reform and agricultural transformation. The development policies have been hailed as overwhelming successes, particularly for their impacts in alleviating poverty (IMF 2011, UN 2013). Statistics show there have been tremendous gains in Rwanda in recent years: Based on the national poverty line of around US\$3/UK£4.50 per adult per week, consumption-based poverty fell by 12% from 57% to 45% in the five years from 2006 to 2011 (NISR 2012). Despite the global downturn Rwanda has consistently achieved high economic growth of

over 5% each year since 2008 and aims to become a middle-income country by 2020 (UNDP 2007, IMF 2011). This represents an astonishing turnaround since the devastation that remained after the 1994 genocide and the following civil war.

The conclusions to be drawn from the Multidimensional Poverty Index (MPI) are very similar to those provided by consumption poverty rates: Authors of the MPI report for Rwanda indicate that approximately 17% of Rwandan households have moved out of multidimensional poverty from 2005 to 2010 (OPHI 2013) and have been quoted as suggesting that if such trends continue, multidimensional poverty may be eradicated in Rwanda in 20 years (Anonymous 2013). The MPI also seeks consistency of measurement but attempts to look beyond consumption as an indicator, to include additional dimensions linked to Millennium Development Goals (MDGs). It attributes weights to household poverty indicators for education, health, housing, sanitation, fuel, water and a number of assets to arrive at an aggregated weighted index (Alkire and Santos 2010).

Approaches to the study of poverty that seek consistent measurement provide valuable understanding of the demographic and geographic distribution of poverty and some of the material and human resources that those people are deprived of. However, research aiming to inform development interventions in developing countries has long recognised that economic indicators, though they usefully reflect some of the relative hardships faced by the population of a given country, are insufficient to understand the complexities of individuals' lives, actions and struggles (Easterlin 2003). Contextual and non-material factors can be crucial to the alleviation of poverty, which may persist not only due to material factors but also due to social processes that result from and lead to inequalities in power between people, which dictate the outcomes they are able to achieve (Cleaver 2005, Du Toit 2005). As Hulme and McKay (Hulme and McKay 2005: 2) succinctly conclude:

Chronic poverty has been studied almost exclusively in relation to income or consumption poverty, and using household panel survey data. Further, much of the focus has been on the identification of chronic poverty and finding correlates, without developing an understanding of the underlying processes by which some people are trapped in persistent poverty while others escape. A broader multidimensional – and multidisciplinary – perspective needs to be brought to the understanding of chronic poverty.

More context-specific approaches may therefore seek to explore social complexity and look not only at comparative measures but at the factors that characterise poverty or wellbeing in a given context. They may focus on: the experiences, values and perceptions of individuals and groups; plurality in ways of thinking and acting; how people's lives are influenced through interactions with other people and institutions; and the dynamics and drivers affecting their lives.

A key message put forward in this chapter is that consistent measures including consumption poverty and the MPI are not wrong in their representation of poverty, but rather incomplete and therefore represent insufficient evidence to draw conclusions about trends in wellbeing or the effectiveness of development policies. Approaches seeking consistency of measurement, alternative approaches incorporating context specificity and the resultant differing viewpoints about development in Rwanda do not actually contradict one another. Both may be equally robust and valid. Indeed, when an individual's wellbeing is considered more holistically, it becomes evident that studies seeking consistent indicators and those exploring social complexity and specificity each tend to measure different aspects of a person's wellbeing. Applying elements of both simultaneously has been shown to provide an additional understanding beyond single disciplinary approaches in the study of poverty (Bamberger et al. 2010, Shaffer 2013), particularly in adding some specificity or contextualisation to the consistency of, for example, consumption-based studies (McGee 2004).

In the case of Rwanda, there is a basis to suggest that life is not actually improving for many Rwandans, that their wellbeing is poorly represented by some of the simple development indicators used to assess policy and that claims of development success should be treated with more than just caution (Ansoms 2011). Critics point out that despite introducing many development programmes, political opposition has been greatly suppressed in Rwanda, the role of civil society is severely limited, policymaking is highly centralised with limited participation and policy monitoring is often restricted to broadscale indicators (Beswick 2010, Gready 2010, Reyntjens 2011), earning Rwanda the label of a developmental patrimonial state (Booth and Golooba-Mutebi 2012). This is not a controversial or disputed claim. It simply receives little attention regarding poverty measurement or alleviation efforts. World Bank governance indicators show Rwanda to be in the lowest 11% of countries in the world with respect to 'voice and accountability', defined as the perceptions of the extent to which the country's citizens are able to participate in selecting their government, as well as freedom

of expression, freedom of association and a free media (WB 2013). Many of the national scale, one-size-fits-all, blueprint development policies introduced (including villagisation of the entire rural population, housing modernisation, health insurance schemes, universal education, trade building improvements, land consolidation and crop specialisation programmes) appear to be negatively impacting large numbers of Rwandans (Newbury 2011, Pritchard 2013, Dawson 2013).

Mixed methods research for bridging the divide

Mixed methods research may be particularly well suited to balance issues of consistency and specificity. However, for several inter-linked reasons, mixed methods are far from the mainstream or most commonly employed in either academia or development practice. Mixed methods research incurs the limitations of and must meet the standards for rigour of both quantitative and qualitative methodologies: in practical research terms this involves promoting credibility or trustworthiness, contextual suitability and auditability as demanded in qualitative study at the same time as ensuring generalisability, validity, reliability and replicability for quantitative analyses (Brannen 2005).

Even if methods are mixed, it is important to note that many studies utilising both quantitative and qualitative methods, which may be considered to be mixed methods studies, do not explore additional dimensions of poverty or wellbeing in sufficient depth to achieve sufficient specificity. The nature of the methods being mixed is very important, particularly for the study of poverty. Studies exploring subjective, non-material and relational factors generally involve more attention to the way the researcher-participant relationship, the situation in which research takes place and the language used influence the answers provided and interpretation of the results (Bryman 2004). This attention to positionality can be a key contribution of the qualitative component of mixed methods studies to the more comprehensive picture they may provide, illustrated below through the shortcomings of one prominent mixed methods study of poverty in Rwanda.

In 2001 the Government of Rwanda undertook a participatory poverty assessment (PPA) under their 'Ubudehe' programme to understand local perceptions of poverty. The exercise was undertaken in one sector in each of the 30 districts in the country. The PPA highlighted the importance of land and livestock to Rwandans and their role in defining poverty. Howe and McKay (2007) took the common characteristics of poor households put forward by Rwandans as strong determinants of poverty

in the PPA (such as land and livestock) and then applied these criteria to the quantitative household survey in order to establish proportions of chronically poor households. In so doing, the study moved beyond normative definitions of poverty to utilise a subjective and context-specific definition put forward by people within that society.

However, one of the major potential contributions of mixed methods approaches is to elaborate the relational and political dimensions to wellbeing and development policy. In incorporating local perspectives into research, attention must be paid to relationships with participants and interpreting their motivation for answering questions in a certain way. The PPA for the 'Ubudehe' programme, the qualitative dimension to this study, relied on research undertaken by the national government (Holvoet and Rombouts 2008). Responses may differ considerably between research conducted by government representatives and researchers emphasising their own independence and motivations, and it is therefore understandable that people did not mention issues regarding ethnicity and politics during the PPA, particularly given the exclusive use of group-based research methods. When participants referred to culture it was in regard to their own 'ignorance, mismanagement of resources, and idleness', as being causes of poverty (Howe and McKay 2007: 201). This point is particularly relevant in Rwanda, where the government has sought both to de-ethnicise the population through reconciliation policies and to unify ethnic groups through powerful discourse surrounding Rwandan citizenship and identity and severe punishment for those who do not comply (Purdekova 2008). Indeed, 'the denial of politics' is a major issue when addressing the effectiveness of development efforts in Rwanda (Holvoet and Rombouts 2008). Through its lack of attention to positionality and reliance on group-based methods, the results of this poverty research lost some of the specificity required to reconcile the differing viewpoints of Rwandan development described earlier. The application discussed in the next section and reflection thereof considers the aspects of positionality and combination of methods with households in more detail.

The application of mixed methods to research rural wellbeing in Rwanda

The previous section described some considerations for undertaking mixed methods study to reveal context specificity and some of the shortcomings of poverty studies in Rwanda. The study presented in this section seeks to move beyond those shortcomings and constraints. To do so

it utilises a conceptual approach for the study of wellbeing in developing countries which builds upon the ideas of the capability approach (for more detail see Gough and McGregor 2007, Dawson 2013). Wellbeing is defined to consist of three inter-related dimensions: the material, relational and subjective dimensions (White 2009). Subjective wellbeing incorporates the idea that two individuals will apply different meanings to resources, will develop varying aspirations and may be differentially satisfied with the same objectively measured quality of life (Copestake and Camfield 2009). The relational dimension focuses attention on 'the rules and practices that govern "who gets what and why"' (White 2009). Groups who apply different (subjective) meanings to ways of living and acting often occupy different relative positions of power in society. This reflects and results in differences in the recognition of their needs and wants by people and institutions, their ability to participate in decisions affecting their lives and therefore in the outcomes they are able to achieve, material or otherwise (Mosse 2010). Poverty may therefore be seen in part as the consequence of social categories and unequal power relations between them, which is very difficult to capture through consistent and measurable rather than context-specific and complex representations (Green and Hulme 2005).

Mixed methods are well suited to wellbeing research as different aspects of an individual's wellbeing lend themselves to each, such that quantitative indicators can be presented alongside qualitative data (Camfield et al. 2009). For example, the study of cultural values and their influence on an individual's aspirations and actions requires concerted qualitative study. But basic human needs (which form an integral part of this wellbeing definition) can be considered as universal for every individual's survival, and they may therefore be studied and recorded more objectively (Doyal and Gough 1991).

This conceptual approach to the study of multidimensional wellbeing was applied to fine-scale research in three study areas in mountainous western Rwanda between October 2011 and May 2012. Of the three research sites, one was in the district of Nyamasheke in the southwestern corner of Rwanda, one in Nyamagabe district further to the east in southwestern Rwanda and one in Rutsiro district in the northwest of the country. The three study areas within those districts were selected because they varied along a gradient of remoteness and were also far enough apart that regional variation could be explored. Within each of the three study sites, a number of villages, adjacent to one another, were selected to provide a representation of the variety of social, economic and cultural groups present. The names and

locations of villages were not pre-selected but rather a decision was taken after several weeks had been spent at each site regarding the suitability of individual villages and the number of villages required to adequately represent the variation present in the population of that area. Each village often comprises just one or two hillsides of clustered houses, typically between 100 and 200 households, and the inhabitants may be quite distinct from the next hillside. Neighbouring villages may contrast strongly based on the history, religion, ethnicity, land use, economic activities and wealth of their inhabitants. The aim of sampling was not to choose a sample population which was deemed to represent proportionally the wider rural population, but to encompass the variation present in the population of the region to enhance the study's generalisability. In all sites, selection of only a single village would have provided a very poor representation of that variety. In total, eight villages were included in the study across the three study areas.

In placing a focus upon subjective experiences of wellbeing and poverty, it was an important first step of this study to consider local conceptions of wellbeing. Hence, focus groups were used to explore the important elements of wellbeing in that local context and to determine whether this differed between villages and groups, whether the importance of certain domains of life was contested or whether consensus could easily be reached regarding what it meant to live well. This initial qualitative exercise was conducted once in each of the eight villages with five to seven randomly selected participants.

The second research method involved conducting interviews with individuals from 165 different households across the eight villages. The number of interviews conducted in each village was derived from

Table 4.1 Household sampling and socio-ethnic diversity by village and district

District	Village	Number of households	Number of interviews	Long-term residents	Returnees from DRC	Twa	Percentage of households interviewed
Nyamasheke	1	126	20	20	0	0	16
	2	176	30	29	1	0	17
Nyamagabe	3	170	20	9	8	3	12
	4	120	20	20	0	0	17
Rutsiro	5	121	20	3	17	0	17
	6	133	20	18	2	0	15
	7	127	20	19	0	1	16
	8	107	15	2	0	13	14
Total	8	1,080	165	120	28	17	15

the overall number of households, which was determined from lists of households and occupants maintained by village chiefs or cell administrators. Households were selected at random from that list for interviews which all took place in the participant's house. To provide a representation of the variation present within the population of each village, interviews were conducted with over 10% of households (Table 4.1). In terms of the three socio-ethnic groups present, this random selection resulted in 73% of households sampled being long-term residents (120 households), 17% returnees from the Democratic Republic of Congo since the 1994 genocide (28 households) and 10% Twa, an indigenous group, many of whose lives were forest-based until recent decades (17 households, Table 4.1). Groups were identified through time spent in villages prior to formal research (see Dawson 2013 for further detail).

Quantitative and qualitative methods for the household level study were not conducted separately, but were instead integrated in the same interview. Interviews involved application of standardised, survey-type questions (to collect socio-economic data about measurable material resources, education and ability to meet basic needs) alongside open and flexible questions exploring the individual's values, priorities, land practices and how they felt about changes in their lives over the past 10 to 15 years.

Three integrated analyses were performed on the resulting data. Results from the analysis of initial focus group exercises were used to guide quantitative analyses of the survey-type data collected for the 165 households interviewed. That analysis, therefore, focuses upon the aspects of material wellbeing put forward by participants themselves as important for their wellbeing and assesses changes in those key resources over time, rather than analysing normatively selected indicators. In order to differentiate based on socio-economic status of households, a hierarchical cluster analysis was used to group households based on the main material and human resources. Clustering was agglomerative using between-group linkages and squared Euclidean distances with standardised values to account for the different scales of the four variables described below.

Finally, the qualitative data obtained from the open questions included in interviews, the unstructured components were used to explore people's values, factors which determine their ways of acting and people's feelings regarding changes occurring in their lives. The mixed methods analytical approach used in this study can be described as an integrated mixed methods study, taking a 'qual-quant-qual' analytical form.

Results and discussion

This section details results from the different stages of data analysis: firstly, focus groups exploring local conceptions of wellbeing; secondly, the quantitative analysis of material and human resources highlighted in those focus groups, collected through household surveys; and finally, further qualitative analyses of both material and non-material aspects of wellbeing and drivers of key changes in respondents' lives.

Initial qualitative exploration

Focus groups were conducted to elicit views as to a local conception of wellbeing. To achieve this only one simple question was asked: what does it mean to live well here, in this village? The facilitator sought only to encourage debate among the five to seven randomly selected participants in each village. The major themes were then extracted from these transcripts through qualitative analysis.

The responses from the eight focus groups exhibited a strong consistency across villages and there was clear agreement within those groups on the major components of wellbeing in that local context. They provided information on issues that people find important for assessing wellbeing, spanning subjective and relational dimensions alongside material ones. These consisted of: land; livestock; employment; health; housing; infrastructure; social relations and sharing; and also finally the autonomy of a household over its own land use and investment decisions. Some components of wellbeing commonly afforded high importance were absent, notably education.

Land was put forward as the priority element of wellbeing and a key resource required in seven of the eight focus groups, and livestock was in each case named as a key resource to enable people to use that land productively to produce food for their household or for trade opportunities. The importance afforded to land and livestock represents a notable departure from 'conventional poverty measures' and the values attached to these resources, and impacts of deprivation in them are therefore prioritised here for more detailed discussion. Without these paired resources people would be expected to struggle to find sufficient food for their household. This initial, very clear response by participants is quite unsurprising given the environmental constraints faced in food production and the proportion of the population who practise subsistence farming. Furthermore, the majority of households grew crops for subsistence rather than for trade and therefore income would not necessarily represent a good proxy for land in poverty assessments.

In addition, land use, particularly agricultural practices, formed an intimate link with social relations and cultural practices. Farming methods often become common to people across a wider area in response to environmental constraints like climate and terrain. The relationship between agricultural practices and wellbeing is therefore complex and often relates to subjective and relational wellbeing, as well as purely material aspects. In this context, as was revealed through qualitative exploration of values, people utilised knowledge which has built up over centuries to practise a complex polyculture to minimise their chances of suffering hunger, and around which patterns of trade and social relations had evolved (Dawson, 2013). The only village for which land was not put forward as the major contributor to their wellbeing was in a village consisting entirely of Twa households, who possessed negligible land and whose cultural attachments to forest appeared stronger than those for agricultural practices.

While education plays a major part in poverty indicators and MDGs, its contribution to wellbeing was highly contested in this particular context and it was considered to play a key role in only the two wealthiest villages. Many people placed education of their children as an aspiration for the future and valued the increased availability of schools closer to their homes. Yet its impact on the wellbeing of households was considered by many to be minimal. It is not difficult to illustrate why this should be the case: In rural Rwanda, the majority of people earn about 100 times less than the minimum wage in the UK on a day they can actually find work (most commonly 50 pence for a day of labouring in crop fields), which is seasonally available and for many very uncertain from day to day. Yet food costs around a quarter to a half of the equivalent cost in Britain. Higher-paid alternatives for work are scarce such that those with relatively high levels of education are often still reliant on manual labouring in rural areas.

Finally, two further elements of wellbeing considered to be important in this rural Rwandan context were less material in nature. People frequently put forward social relations and sharing within the village and autonomy, the freedom to act to pursue their own wellbeing goals, as key contributors to their wellbeing. The inclusion of these non-material elements supports the use of concerted study of subjective and relational aspects of wellbeing to assess the impacts of change or of policies on rural Rwandans. Autonomy was considered to be a recurring theme through expressions such as, 'We need for rural people's opinions to be heard by the government!' Respondents in every focus group highlighted policy examples through which their behaviour, practices,

earnings and expenditure were subject to control which they considered excessively restrictive and for which non-compliance could result in fines. These included policies governing housing materials, standards for household items, agricultural practices, crop selection and livestock husbandry.

Quantitative representation of heterogeneity in sample population

By disaggregating households, the variation within a population can be represented to highlight which types of people suffer which types of poverty, with implications for the design of poverty alleviation policies. In attempting to display variation in the more objective aspects of wellbeing in this case study, the results of focus groups (the objective indicators put forward as being important by participants themselves) were used to guide quantitative analysis. Therefore land size, livestock, employment and shelter were categorised to describe households (Table 4.2).

A hierarchical cluster analysis was then employed to look not only at which households suffer deprivations in each of those resources but to determine the extent to which socio-economic groups could be identified based on this group of resources, and particularly to highlight which and how many households in the sample could be described as suffering poverty in the local context. Clustering was agglomerative using between-group linkages and squared Euclidean distances with

Table 4.2 Categories of land size, livestock, shelter and occupation used for hierarchical cluster analysis (based on interview data for 165 households)

Land size	Livestock	House type	Occupation
< 0.1 hectare (17%)	No livestock (33%)	Basic constructions of earth and sticks (25%)	Agricultural labour only (17%)
0.1 to 0.25 ha (23%)	Small livestock or borrowed cow (31%)	3 rooms or less, constructed with large adobe or concrete blocks (42%)	Other labouring (e.g. tea labour, building, charcoal making or brewing (25%)
0.26 to 0.5 ha (22%)	Own one cow (22%)	Larger houses built with concrete or manufactured bricks (33%)	Own trade such as crops and shop owners (36%)
0.51 to 1 ha (19%)	Own two cows or more (14%)		Professionals (builders, teachers, administrators, mechanics or drivers) (22%)
1.1 to 2.5 ha (14%)			
> 2.5 ha (4%)			

standardised values to account for the different scales of the four variables detailed in Table 4.2. Cluster results are presented in Table 4.3.

The four groups identified in the cluster analysis presented in Table 4.3 were very clearly divided and the sample size enabled common groups and patterns to be drawn from it. Standard errors for each of the variables were very small and differences between each group were highly statistically significant ($p < 0.01$). The differences between landless labourers and relatively wealthy groups in terms of material resources reflect high levels of inequality and suggest that changes and development policies are experienced very differently by groups in rural Rwandan society. Most notably in terms of studying poverty dynamics, the group identified as 'landless labourers' is very similar to the groups defined as living in chronic poverty in Howe and McKay's study of rural poverty dynamics described above (Howe and McKay 2007). For landless labourers land and livestock holdings are negligible, their shelters very basic and small and they are entirely dependent upon sporadic labouring opportunities which yield uncertain and very small wages. This group represented 34% of the sample population (although this may not be representative of the wider population in the region due to the small sample size and potentially non-representative sample). The deprivation experienced by this group is illustrated by their inability to meet basic needs: 89% of this group collect firewood illegally and 75% suffer food scarcity, having to go at least an entire day a month with no food at all (this often occurred much more frequently).

Table 4.3 Results of hierarchical cluster analysis displaying different groups identified from 165 households in rural Rwanda (greater detail on methods can be found in Dawson, 2013)

	Landless labourers (34%)	Resource poor workers (38%)	Relatively wealthy, diversified farmers (24%)	Relatively wealthy professionals without livestock (3%)
Land	Average 0.13 ha	Average 0.56 ha	Average 2 ha	Average 2.25 ha
Livestock	Majority have none, 7% own cow	29% own a cow	Nearly all own cows	No livestock. All grow trees commercially
Occupation	Labourers earning 40p to £1 per day	Regular, low-paid work. 43% trade crops	Own business or professionals. 68% trade crops	All are professionals and trade crops
Housing	Small and basic houses	Mostly medium	Relatively large houses	All have large houses

Levels of poverty and inequality were also explored across socio-ethnic groups. During introductory and sampling phases of fieldwork, prior to focus groups and household interviews, clear differences had been observed between socio-ethnic groups based on their material wellbeing and most obviously based on their spatial segregation. Analysis was therefore performed to explore the proportion of people from each socio-ethnic group who were represented in the four socio-economic groups identified in Table 4.3. Results revealed that less than 10% of returnee households were classified as landless labourers compared to 30% of long-term resident households and a striking 88% of Twa households. Indeed not one Twa household belonged to the highest two categories. Households lacking a male household head were also skewed towards the poorer groups but the effect was slight compared to socio-ethnic difference.

Quantitative analysis, based on these objective indicators, can also inform us about trends in material wellbeing among the sample population. Recall, supported by triangulation of responses, is a common method to reconstruct data about past events (Bamberger et al. 2010). By asking respondents to recall to what extent their land holdings, livestock and occupations have changed over the preceding decade it is possible to ascertain changes in their socio-economic status over time. Responses were verified visually where possible, and also verified through repeated questioning at later stages of the single interview conducted in each sample household. This exercise revealed that 12% of the sample population had fallen into the landless labourer category in the last decade, having sold large areas of land in that time. Analysis of change between socio-economic categories also revealed that land holdings were not falling consistently across the categories identified. While 28% of all sampled households revealed falling land holdings, 16% were also able to increase holdings over the same 10 to 15 year time period. Those who were able to increase holdings were almost all in the two relatively wealthy socio-economic categories in Table 4.3. This indicates a redistribution of land away from the poorest households towards the wealthiest and suggests some very different trends to those indicated by standard poverty measures. National level measures of inequality suggest that inequality is extremely high in Rwanda. The Gini coefficient has consistently been greater than 0.5 since 2000 (<http://data.worldbank.org/indicator/SI.POV.GINI>, accessed 9 July 2013). This simple depiction of difference and change does not provide conclusions as to why this inequality between groups exists but does indicate that social and cultural difference and associated inequalities in power may

play a substantial role in the outcomes which people are able to attain. The nature of these social processes and inequalities therefore form an important component of further qualitative analysis.

Further qualitative analysis: Relational and subjective wellbeing and causes of change

In order to gain further insight into the impacts and effectiveness of development policy in rural Rwanda, further qualitative study was performed based on household interviews. Results are presented only briefly here. See Dawson (2013) for further detail.

Despite wellbeing gains in certain areas of rural lives (particularly security, health, housing standards and education provision), virtually all respondents expressed during household interviews that they perceived wellbeing as a whole to be decreasing and poverty levels to have increased over the previous 10 to 15 years. Furthermore, in many cases it was the poorest among them who indicated they had suffered the most from these changes. This trend has also been noted in other rural areas in Rwanda (WFP 2009, Ansoms and McKay 2010). Increases in poverty were on the one hand experienced through falling resource levels, particularly land and livestock holdings, a trend illustrated through the quantitative analysis of material wellbeing. On the other hand this was felt in ways not measured by conventional poverty indicators: through non-material aspects of people's lives such as impacts on cultural practices, social relations and a curtailment of people's individual freedom to act as they wish to pursue their goals.

The drivers of change mentioned during interviews included sporadic food price changes, environmental change in the form of soil degradation and social change through increased exposure to global markets (Dawson, 2013). However, one of the major themes to be drawn from household interviews was the primarily negative influence of what were ostensibly development policies on the wellbeing of rural households. The very policies which had been declared successes based on limited quantitative indicators, including poverty measurements (IFAD 2011, IMF 2011), were perceived by participants to have negatively impacted them. The development policies considered to have impacted poor households in this study negatively include enforced housing improvements, modernisation of trade buildings, villagisation of scattered rural settlements and agricultural intensification policies. I discuss villagisation and agricultural intensification below.

The villagisation policy aims to bring the entire rural population of the country into clustered settlements by 2020 (ROR 2007), ostensibly

in order to provide services and better housing to the rural population. With local targets passed down to local administrators by the central government, considerable pressure is applied on particular families living remotely by their fields to move home. The financial burden involved in moving to more costly areas with higher housing standards often involves selling land, incurring debt and losing income-earning potential and has contributed to land redistribution and decreased wealth among some of the poorest, despite indicators of increased housing standards and the potential for electricity to be installed reflecting a reduction in poverty.

Although the villagisation policy has had profound effects for some, the greatest negative impact was deemed to be incurred as a result of agricultural development policies. The new Rwandan Land Policy and Crop Intensification Program (ROR 2004, MINAGRI 2011) was put forward as a set of interventions aiming to improve the lives of smallholder farmers and to improve food security for the Rwandan population (i.e. policies with development objectives). These two policies seek to promote increased utilisation of agricultural land and production of marketable crops to enable smallholders to make a transition towards modern, market-oriented farmers. The effects of those policies on autonomy, expectations, tenure security over homes and land, and on social and cultural practices formed some of the most important trends and shocks occurring in the lives of rural inhabitants.

In rural Rwanda, the traditional mode of food production, growing multiple crops, is a method to reduce vulnerability to crop failure, high prices for alternative foods at market and to reduce times without harvestable staples in the face of highly sloping land and extreme and unpredictable dry and rainy seasons. Farming methods often develop over centuries as a response to environmental uncertainty and those methods represent a cultural and human resource (Zoomers 1999). In this polyculture system, 60 different crop types were recorded in western Rwanda, often with overlapping crop cycles (NISR 2010). Rwandan agricultural policies are directly opposed to this system, deeming it 'a mediocre agriculture that has no future' (ROR 2004). In the three areas within this study, maize, beans and tea were being gradually imposed upon farmers, despite their own perception that maize was unsuitable for their land and that cash crops were unsuitable for their subsistence needs. The approved crops require use of chemical fertilisers, which are provided on credit. However, as only 38% of households traded crops (the rest relying on subsistence), only a third of households actually used them, leaving the rest more vulnerable to crop failure, lack of food

and an increased need to sell their land simply to meet costs and feed their family. Among the 165 interview respondents only six households stated support for the policy whereas 68 described that it had affected them negatively, a large number despite the obvious reluctance of respondents to voice strong opinions about government policy.

Despite the potential impact on land tenure, land distribution and the ability of millions of rural smallholders to subsist, the policy may still be deemed a success, based on increased production yields of approved crop types. Furthermore, in terms of consistent measures of poverty, the loss of land and reduction in time spent producing food for a household may lead to increased reliance on labouring for income meaning that the purchasing power of households could actually be deemed to increase. The enforced change in land practices has affected not only material wellbeing but also non-material elements, having strong effects on social relations and cultural practices. It is the pervasive impact of such policies on a household's expectations, agency and ability to utilise their human, social and cultural resources which are unrepresented through objective, normative, conventional poverty indicators. This difference in the focus of different approaches to the study of poverty and vulnerability is a key factor explaining the difference between the two polar viewpoints of Rwandan development policy impacts described throughout this chapter.

Reflections on the application of integrated mixed methods

The previous sections highlighted some of the shortcomings of poverty research in Rwanda, both in quantitative and mixed methods studies, and illustrated why divergent conclusions about poverty trends and policies may emerge. The application of mixed methods in this study reveals how some of those limitations may be overcome and how consistent poverty measures and context-specific poverty studies may be reconciled. This was achieved through the integration of subjective and relational dimensions of wellbeing with research into the more objective, material aspects of people's lives. I focus particularly on the issue of positionality and integration of data collection methods.

Attention to positionality is a crucial factor enabling context specificity to emerge and is therefore also essential for integrated mixed methods study. But while this represents an opportunity for mixed methods study, it also highlights some tensions relating to the practical, logistical application of research methods. The requisite standards for qualitative

and quantitative data differ and to meet these multiple requirements means that mixed methods research is time-consuming, necessitates a broad range of skills for research teams spanning several academic disciplines and it can be costly to develop sufficient understanding and relations to perform well (Brannen 2005). Time was required in this study to build relationships and for continued reflection and flexibility, aspects which commonly receive less attention in quantitative household surveys. This was a mixed methods study, not ethnography, but the attention paid to the researcher-participant relationship went well beyond that of a rapid rural appraisal. Steps were taken as part of the research to spend sufficient time in the research sites to interact with people prior to interviews in order to allow participants to understand the motivations for this project, scope of the work, likely timescale, outcomes and, crucially, to enable them to make an informed decision on whether to take part or not. Numerous respondents confided about conflicts within the family, health issues and candid opinions about government policies or local level corruption and this provided me with confidence that my position was well understood and trusted as an independent, albeit white, British student. Up to three months were spent at each of the three study sites and, overall, more than one day in the field for each of the 165 interviews and eight focus groups conducted. This serves to illustrate well the trade-off between sample size for quantitative analyses and the levels of trust and depth of understanding required from qualitative analyses, when methods are conducted concurrently by the same researcher. A focus on researcher-participant relationships allows reflection on the way in which researchers and methods are perceived, questions are comprehended and any reasons for the provision of certain verbal or non-verbal responses (Tourangeau et al. 2000). For example, development discourse is prevalent in many developing countries and responses to surveys often strongly reflect this discourse if attention to positionality is not given sufficient consideration (Van Dijk 1990). Such an approach, although ultimately limiting potential sample sizes, enabled sampling and research methods to be tailored to the specific context, and to explore the nature of social differentiation and subjective factors which define wellbeing in the specific context. Relational factors can play a particularly important role in local experiences of poverty and vulnerability in sensitive political environments and ongoing reflexivity regarding the relationship with participants. Attention to research ethics and adaptability to changing circumstances were important factors in developing a suitable sampling strategy to enable the importance of power and culture to emerge.

This study integrated data collection for quantitative and qualitative analyses simultaneously into one household interview. Their combination in data collection also enhanced their compatibility in subsequent analyses and interpretation of results, for example, through identifying changes in holdings of material resources and matching these to perceptions of drivers of change. Qualitative and quantitative data collected in this study formed parts of one holistic (and multi-disciplinary) wellbeing framework rather than separate (and disciplinarily distinct) methodologies and this contributed to the different findings supporting, rather than contradicting, one another. This presents a possible case for mixed methods to be integrated, conceptually, through fieldwork and in the analytical approach, rather than each being designed, conducted and inevitably evolving, in isolation. The perceived cost of this integrated and detailed approach may be the limited sample size, geographic coverage and, therefore, relevance of results to other locations. However, for this study the validity of the findings and their generalisability may be relatively high, because many of the drivers of change were national scale policies.

Conclusions

The way in which poverty is defined, measured and researched strongly influence judgements about the suitability and effectiveness of development policies in alleviating poverty. In particular, poverty measures based on consistent, universal objective indicators can produce seemingly conflicting results with studies emphasising context-specific definitions, experiences and drivers of poverty. Mainstream measurements of poverty focus entirely on objectively measurable components of material wellbeing to maintain international consistency, while overlooking the subjective and relational dimensions which are often fundamental to people's wellbeing in that specific context (Ferguson 1990, Gledhill 2000). It is for this reason that questions about the effectiveness of development policies in Rwanda polarise opinion. The rapid reduction in the proportion of the population falling below poverty indicators has been praised as an example of successful development. However, at the same time, the Rwandan government is decried for implementing a suite of coercive policies which disrupt traditional modes of production and social systems and favour wealthier citizens while increasing the vulnerability of the poorest in society. This chapter has described how this apparent contradiction arises from different types of poverty research, those which seek to consistently measure poverty through objective indicators and those focused on more context-specific and

subjective factors. Yet the chapter has also illustrated that both represent valid perspectives. Furthermore these apparently contradictory perspectives can be reconciled using a mixed methods approach incorporating a more holistic analysis of wellbeing.

Access to material resources including water, shelter and medicine is universally important for people's wellbeing. In this sense great strides have been made in alleviating poverty in Rwanda. Yet such gains do not represent a complete picture of what is required to live free from poverty or enable conclusions to be drawn about overall poverty trends. Though the MPI adds further material and human resources to monetary-based poverty measurement, it also suffers the same limitations as consumption-based poverty indicators (Ravallion 2011). Research that pays attention not only to the incidence of absolute, measurable poverty but also to the dynamic social processes which effectively reproduce poverty outcomes may contribute greatly to the implementation of effective long-term solutions (Hulme and Shepherd 2003, Bevan 2004). In the application to Rwanda, for example, the extent of poverty differed considerably between socio-ethnic groups, suggesting that disaggregation on this basis may provide valuable insights into the social processes which reproduce poverty.

By utilising local definitions of wellbeing, poverty and inequality were shown to be increasing in the villages that were part of this study, even though mainstream, consistent poverty indicators may reveal an opposite trend. The exclusion of land and livestock from those indicators, such critical resources for the wellbeing of many rural people, was a key factor in this discrepancy and, more generally, represents a great anomaly surrounding research into rural poverty. Additionally, some of the resources which enable people to live free from poverty are not valued for their material contribution alone. Instead, their cultural and social meaning may form an important component, as was illustrated through the value conferred on land in rural Rwanda. To reveal some of these further elements of wellbeing, exploratory research approaches may be much more suitable than the 'confirmatory' approaches which may result from rigidly designed studies which only focus on consistent, objectively measurable poverty indicators (Copestake 2013). It is the study of the further subjective and relational dimensions of wellbeing and the combination of both qualitative and quantitative methods in doing so that support the counternarrative regarding Rwanda's developmental trajectory. Ultimately, by using multiple methods to define and analyse trends in wellbeing, negative poverty trends were shown to be exacerbated by development policies themselves.

The case supporting the wider adoption of mixed methods has come through the realisation that long-term development success has lagged far behind the potential and that the limited attention paid to context, culture and relational factors has been a contributory factor. But although mixed methods approaches can overcome many of the described weaknesses in poverty research, their application is far from common. There are a number of inter-related reasons why the credibility of this type of research is questioned by current academic and institutional norms. Mixed methods poverty research spans different disciplines (Kanbur and Shaffer 2007) and though this chapter has demonstrated this is not insurmountable, the clash between the skills required to perform 'thick description' of social complexity (Geertz 1973) and the strict academic and professional demands of development practice may form a considerable barrier. Attention to positionality and the use of research methods which explore subjective and relational aspects of wellbeing in addition to the material and human resources people possess or lack can allow more holistic interpretations of poverty and its causes to develop.

This study has shown that mixed methods research has a crucial potential role to play in the progression of development practices towards achieving poverty alleviation goals and avoiding negative policy impacts. Ethically, more holistic research approaches may help to break a cycle of development practice which, although in some circumstances may produce tangible gains, in others it may prove to be ineffective, and critically in others may result in an unregistered harm to some of the poorest people on earth (Mosse 2004). The challenge for those practising mixed methods research lies in continuing to utilise and refine approaches, in different fields and contexts, so that their use becomes common practice and the insights they bring forward become what is considered to be good or sufficient evidence. Acknowledgement of the credibility of mixed methods as a research method is growing. Acknowledgement of its role in providing credible evidence of development or policy impacts may also follow.

Note

- 1 This study was made possible through funding from the Social Sciences Faculty of the University of East Anglia and with support from the Ecosystem Services for Poverty Alleviation (ESPA) programme. Permitting and logistics were aided by staff at the Rwanda Development Board, Wildlife Conservation Society and Forest of Hope in addition to village chiefs and local executive secretaries. Above all we are thankful to those who participated in and contributed to the study.

References

- Alkire, S. and M. E. Santos. 2010. Acute Multidimensional Poverty: A New Index for Developing Countries. *Oxford Poverty & Human Development Initiative Working Paper No. 38*. Oxford: OPHI.
- Anonymous. 2013. Oxford University Says Rwanda to 'Eradicate' Poverty in 20 Years. *News of Rwanda*, 18 March. <http://www.newsforwanda.com/featured1/17405/oxford-university-says-rwanda-to-eradicate-poverty-in-20-years/>
- Ansoms, A. 2011. Rwanda's Post-genocide Economic Reconstruction: The Mismatch between Elite Ambitions and Rural Realities. In Straus, S. and Waldorf, L. (eds), *Reconstructing Rwanda: State Building & Human Rights after Mass Violence*. Madison, WI: University of Wisconsin Press, 240–251.
- Ansoms, A. and A. McKay. 2010. A Quantitative Analysis of Poverty and Livelihood Profiles: The Case of Rural Rwanda. *Food Policy* 35: 584–598.
- Bamberger, M., V. Rao, and M. Woolcock. 2010. Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development. *Policy Research Working Paper No. 5245*. Washington, DC: World Bank.
- Beswick, D. 2010. Managing Dissent in a Post-genocide Environment: The Challenge of Political Space in Rwanda. *Development and Change* 41: 225–251.
- Bevan, P. 2004. Exploring the Structured Dynamics of Chronic Poverty: A Sociological Approach. *Wellbeing in Developing Countries (WeD) Working Paper*. Bath: University of Bath.
- Booth, D. and F. Golooba-Mutebi. 2012. Developmental Patrimonialism? The Case of Rwanda. *African Affairs* 111: 379–403.
- Brannen, J. 2005. Mixed Methods Research: A Discussion Paper. *ESRC National Centre for Research Methods*. Southampton: ESRC National Centre for Research Methods Review Paper.
- Bryman, A. 2004. *Social Research Methods*. 2nd Edition. Oxford: Oxford University Press.
- Camfield, L. and D. Ruta. 2007. 'Translation is not Enough': Using the Global Person Generated Index (GPGI) to Assess Individual Quality of Life in Bangladesh, Thailand, and Ethiopia. *Quality of Life Research* 16: 1039–1051.
- Camfield, L., G. Crivello, and M. Woodhead. 2009. Wellbeing Research in Developing Countries: Reviewing the Role of Qualitative Methods. *Social Indicators Research* 90: 5–31.
- Cleaver, F. 2005. The Inequality of Social Capital and the Reproduction of Chronic Poverty. *World Development* 33: 893–906.
- Copestake, J. 2013. Credible Accounts of Causation in Complex Rural Contexts: Exploring Confirmatory and Exploratory Qualitative Impact Research Methodologies. *Plenary address at University of East Anglia/Institute of Development Studies Workshop on Mixed Methods Research in Poverty and Vulnerability*, 2 July 2013, London.
- Copestake, J. and L. Camfield. 2009. Measuring Subjective Wellbeing in Bangladesh, Ethiopia, Peru and Thailand using a Personal Life Goal Satisfaction Approach. *University of Bath/Wellbeing in Developing Countries Research Group (Wellbeing in Developing Countries (WeD) Working Papers; WeD Working Paper 09)*. Bath: University of Bath.
- Dawson, N. 2013. *An Assessment of Multidimensional Wellbeing in Rural Rwanda: Impacts of and Implications for Rural Development and Natural Resource Conservation*.

- PhD thesis, School of International Development, University of East Anglia, Norwich, UK.
- Des Forges, A. 2005. *Land in Rwanda: Winnowing out the Chaff*. L'Afrique des Grands Lacs: Annuaire 2006.
- Doyal, L. and I. Gough. 1991. *A Theory of Human Need*. Basingstoke: Macmillan.
- Du Toit, A. 2005. Chronic and Structural Poverty in South Africa: Challenges for Action and Research. *Chronic Poverty Research Centre Working Paper*. Cape Town: Centre for Social Science Research, University of Cape Town.
- Easterlin, R. A. 2003. Explaining Happiness. *Proceedings of the National Academy of Sciences of the United States of America* 100: 11176–11183.
- Ferguson, J. 1990. *The Anti-Politics Machine. 'Development,' Depoliticization, and Democratic Power in Lesotho*. Cambridge: Cambridge University Press.
- Geertz, C. 1973. *The Interpretation of Cultures: Selected Essays by Clifford Geertz*. New York: Basic Books.
- Gledhill, J. 2000. *Power and Its Disguises: Anthropological Perspectives on Politics*. London: Pluto Press.
- Gough, I. and J. A. McGregor. 2007. *Wellbeing in Developing Countries: From Theory to Research*. New York: Cambridge University Press.
- Gready, P. 2010. 'You're either with us or against us': Civil Society and Policy Making in Post-genocide Rwanda. *African Affairs* 109: 637–657.
- Green, M. and D. Hulme. 2005. From Correlates and Characteristics to Causes: Thinking About Poverty from a Chronic Poverty Perspective. *World Development* 33: 867–879.
- Holvoet, N. and H. Rombouts. 2008. The Challenge of Monitoring and Evaluation under the New Aid Modalities: Experiences from Rwanda. *The Journal of Modern African Studies* 46: 577–602.
- Howe, G. and A. McKay. 2007. Combining Quantitative and Qualitative Methods in Assessing Chronic Poverty: The Case of Rwanda. *World Development* 35: 197–211.
- Hulme, D. and A. McKay. 2005. Identifying and Measuring Chronic Poverty: Beyond Monetary Measures. *Chronic Poverty Research Centre Working Paper*. Cape Town: Centre for Social Science Research, University of Cape Town.
- Hulme, D. and A. Shepherd. 2003. Conceptualizing Chronic Poverty. *World Development* 31: 403–423.
- IFAD. 2011. Smallholder Cash and Export Crops Development Project, Republic of Rwanda. *Interim Evaluation Report No. 2302*. Rome: International Fund for Agricultural Development.
- IMF. 2011. Rwanda: Poverty Reduction Strategy Paper – Progress Report. *International Monetary Fund Country Report No. 11/154*, Washington, DC.
- Kanbur, R. and P. Shaffer. 2007. Epistemology, Normative Theory and Poverty Analysis: Implications for Q-Squared in Practice. *World Development* 35: 183–196.
- Kuhn, T. S. 1970. The Structure of Scientific Revolutions. In *International Encyclopedia of Unified Science 2*. University of Chicago Press.
- McGee, R. 2004. Constructing Poverty Trends in Uganda: A Multidisciplinary Perspective. *Development and Change* 35: 499–523.
- McKay, A. and D. Lawson. 2003. Assessing the Extent and Nature of Chronic Poverty in Low Income Countries: Issues and Evidence. *World Development* 31: 425–439.

- MINAGRI. 2011. Strategies for Sustainable Crop Intensification in Rwanda: Shifting Focus from Producing Enough to Producing Surplus. *Rwandan Ministry of Agriculture Report*. Kigali: Rwanda Ministry of Agriculture, p. 59.
- Mosse, D. 2004. Is Good Policy Unimplementable? Reflections on the Ethnography of Aid Policy and Practice. *Development and Change* 35: 639–671.
- Mosse, D. 2010. A Relational Approach to Durable Poverty, Inequality and Power. *The Journal of Development Studies* 46: 1156–1178.
- Newbury, C. (ed.) 2011. High Modernism at the Ground Level: The *imidugudu* Policy in Rwanda. In Straus, S. and Waldorf, L. (eds), *Reconstructing Rwanda: State Building & Human Rights after Mass Violence*. Madison, WI: University of Wisconsin Press, 223–239.
- NISR. 2010. National Agricultural Survey 2008. *National Institute of Statistics of Rwanda Report*. Kigali: Republic of Rwanda, p. 249.
- NISR. 2012. The Evolution of Poverty in Rwanda from 2000 to 2011: Results from the Household Surveys. *National Institute of Statistics of Rwanda*. Report, Government of Rwanda, Kigali.
- OPHI. 2013. Rwanda Country Briefing, Multidimensional Poverty Index Data Bank. *Oxford Poverty and Human Development Initiative*. Oxford: University of Oxford. Available at: www.ophi.org.uk/multidimensional-poverty-index/mpi-country-briefings/
- Pogge, T. and S. G. Reddy. 2005. *How Not to Count the Poor*. Available at: SSRN 893159.
- Pritchard, M. F. 2013. Land, Power and Peace: Tenure Formalization, Agricultural Reform, and Livelihood Insecurity in Rural Rwanda. *Land Use Policy* 30: 186–196.
- Purdekova, A. 2008. Building a Nation in Rwanda? De-ethnicisation and Its Discontents. *Studies in Ethnicity and Nationalism* 8: 502–523.
- Ravallion, M. 1996. *Issues in Measuring and Modeling Poverty*. Washington, DC: World Bank Publications.
- Ravallion, M. 2011. On Multidimensional Indices of Poverty. *The Journal of Economic Inequality* 9: 235–248.
- REMA. 2009. *Rwanda State of Environment and Outlook Report*. Rwanda: Rwanda Environment Management Authority.
- Reyntjens, F. 2011. Constructing the Truth, Dealing with Dissent, Domesticating the World: Governance in Post-genocide Rwanda. *African Affairs* 110: 1–34.
- ROR. 2004. National Land Policy. *Republic of Rwanda Report*, Kigali, Rwanda.
- ROR. 2007. *Economic Development and Poverty Reduction Strategy, 2008–2012*. Ministry of Finance and Economic Planning. Republic of Rwanda (ROR), Kigali.
- Sanchez, P., C. Palm, J. Sachs, G. Denning, R. Flor, R. Harawa, B. Jama, T. Kiflemariam, B. Konecky, and R. Kozar. 2007. The African Millennium Villages. *Proceedings of the National Academy of Sciences* 104: 16775–16780.
- Shaffer, P. 2013. Ten Years of ‘Q-Squared’: Implications for Understanding and Explaining Poverty. *World Development* 45: 269–285.
- Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies, and B. Befani. 2012. Broadening the Range of Designs and Methods for Impact Evaluations. *DFID Working Paper* 38. London, UK: DFID.
- Straus, S. and L. Waldorf. 2011. *Remaking Rwanda: State Building and Human Rights after Mass Violence*. Madison, WI: University of Wisconsin Press.

- Tendler, J. 1997. *Good Government in the Tropics*. Baltimore: Johns Hopkins University Press.
- Tourangeau, R., L. J. Rips, and K. Rasinski. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.
- UN. 2013. Assessing Progress in Africa toward the Millennium Development Goals. *MDG Report*. United Nations Report.
- UNDP. 2007. Turning Vision 2020 into Reality. From Recovery to Sustainable Human Development: National Human Development Report. *United Nations Development Programme, Rwanda 2007*. New York: UNDP.
- Van Dijk, T. A. 1990. Social Cognition and Discourse. In H. Giles and W. P. Robinson. *Handbook of Language and Social Psychology*. Chichester: John Wiley and Sons, 163–183.
- WB. 2013. Data downloaded from <http://info.worldbank.org/governance/wgi/> on 18/06/2013.
- WFP. 2009. Rwanda: Comprehensive Food Security and Vulnerability Analysis, July 2009. *World Food Programme Report*, Rome, Italy.
- White, S. C. 2009. Analyzing Wellbeing: A Framework for Development Practice. *University of Bath/Wellbeing in Developing Countries Research Group (Wellbeing in Developing Countries (WeD) Working Papers; WeD Working Paper 09)*. Bath: University of Bath.
- Wilson, J. 2013. Model Villages in the Neoliberal Era: The Millennium Development Goals and the Colonization of Everyday Life. *Journal of Peasant Studies* 41: 1–19.
- Zoomers, A. 1999. *Linking Livelihood Strategies to Development: Experiences from the Bolivian Andes*. Amsterdam: Royal Tropical Institute, KIT Press.

5

Measuring the Resilience of Rural Households in Burkina Faso

Lucrezia Tincani and Nigel Poole

Introduction

The concept of resilience is at the centre of current debates in development, climate change adaptation and humanitarian aid. Most recently, calls were made to integrate ‘resilient development’ and disaster risk management into each sector of the post-2015 Sustainable Development Goals (Hasan 2014). Conostas et al. (2014: 4) wrote, ‘In a world where conventional approaches to dealing with humanitarian aid and development assistance have been questioned, resilience has captured the attention of many audiences because it provides a new perspective on how to effectively plan for and analyse the effects of shocks and stressors that threaten the well-being of vulnerable populations.’ But what is meant by resilience, how can we measure it and what do we know about the factors that drive resilience?

The Sahel has been at the forefront of the debate on resilience due to the apparent failure to sustainably improve food security in the region, which has suffered from a recurring cycle of food shortages for decades (Jefferys 2013). Sahelian populations have been depicted as passive victims subject to exogenous features of the ‘complex, diverse and risk-prone’ environments they live in (Chambers 1997). Others, however, portray these populations as highly skilled at adapting their livelihoods to a challenging environment, with ‘resilience’ seen as an essential pre-requisite to reducing poverty and food insecurity (Gubbels 2013). Burkina Faso, a Sahelian landlocked country, provides a pertinent case study of these polarised views. However, it is not yet fully understood how such ‘resilience’ can be achieved and what indicators can be used to measure resilience at the scale of a livelihood system. With over 2.5 billion people living in drylands worldwide, it is important to

understand how sustainable agricultural livelihoods are constructed and maintained in risk-prone environments.

This chapter presents a mixed methods approach to unpicking livelihood resilience, documenting how rural populations in two provinces of Burkina Faso manage risk and adapt their livelihoods. The rationale for choosing a mixed methods approach is laid out below. The approach aims to identify factors enabling resilience, explore how strategies can vary within and between households and seasons, and understand how and why these drive household resilience and impact on a household's welfare – specifically their level of food security. Through the use of mixed methods this chapter provides a more nuanced understanding of how sustainable agricultural livelihoods are constructed and maintained in risk-prone environments. The benefits and challenges regarding the application of mixed methods are also discussed.

Conceptual background

Resilience: More than 'bouncing back'

The term 'resilience' is widely used in a range of disciplines from psychology to engineering, and is increasingly employed by humanitarian and development actors in the context of climate change adaptation and disaster risk reduction. However, the term 'resilience' is used in different ways within different disciplines because the concepts behind it have separate intellectual histories (Janssen and Ostrom 2006).

In engineering and psychology, the term 'resilience' is equated to the capacity of a physical structure or person to 'bounce back' or return to equilibrium following a perturbation or psychological stress (Gunderson and Holling 2002). In contrast, ecological resilience refers not to a simple return to status quo after a shock but refers to repeating cycles of adaptation and change which can result in a variety of different outcomes. This chapter uses the definition from ecology and applies it to the research on poverty, understanding resilience as the property which helps people and communities to anticipate, prepare for, cope with, and recover from shocks. It not only allows them to 'bounce back to where they were before the shocks occurred, but [also to] become even better-off' (Fan et al. 2014).

The poverty literature has had a long tradition of examining the links between shocks, poverty or food insecurity and coping, examining the processes leading to the outcome of poverty¹ (Chaudhuri 2003, Davies 1996). Widely documented processes include coping strategies, before as well as after a shock (Dercon 2005). These strategies are usually

examined at the scale of an individual or a household, with the term ‘resilience’ often employed as a loose antonym of vulnerability (Adger 2000). Examining the same links between shocks, food insecurity and coping through a ‘resilience’ lens would similarly concentrate on understanding the underlying processes, but focus the analysis on a different level: Resilience refers to a process occurring at the scale of a whole system – be it an ecosystem or a livelihood system – instead of a process at the level of a household or individual. As such it moves beyond understanding the individual determinants of vulnerability, and focuses on understanding the system-wide factors which help groups of people – extended families and communities – to anticipate, prepare for, cope with, and recover from shocks. For this reason, this study focused on polygamous families in Burkina Faso as they are large enough to each manage their own livelihood ‘system’.

Links between ‘resilience’ and ‘food security’

While resilience has been used effectively within ecology, psychology and engineering, and more recently in the humanitarian sector, resilience within the context of food and nutrition security is still a nascent concept (Fan et al. 2014). ‘Resilience’ and ‘food security’ relate to each other in two different ways.

Firstly, food and agricultural systems themselves need to be resilient to shocks, both large and small, to help preserve food availability and access even when shocks occur. Otherwise, economic shocks such as volatile food prices and financial crises, for example, can cause fluctuations in food prices which can reduce food consumption among net food buyers in developing countries (Zselezky and Yosef 2014). In other words, resilience of a livelihood system can ensure the food security of individuals within that system.

Secondly, being food secure facilitates the ability of an individual to anticipate, prepare for, cope with and recover from shocks simply because that individual is healthy and able to work and operate within their social network, all other things being equal. The food security of individuals can also enhance the resilience of whole economies by enhancing the long-term health and productivity of individuals (Hoddinott 2014). This concept is particularly relevant to the food security and nutrition sector because it emphasises the adverse long-term consequences of short-term shocks (Constas et al. 2014). In the context of addressing chronic undernutrition, there is a compelling body of evidence showing that not only do shocks and stressors such as civil war and drought have immediate effects on pre-school children’s nutritional

status but that these effects persist into adulthood. Conversely, children in households with greater resilience are likely to be better nourished and better schooled; and, as adults, these children will likely be more resilient to the shocks and stressors they face.

In summary, individual food security can enable individual resilience and therefore resilience of a livelihood system more widely – but a resilient livelihood system can also ensure the food security of individuals within that system. In other words being food secure can be both an enabling factor and an outcome of resilience.

Climate change literature is increasingly examining these inter-linkages between resilience and food security in the context of adaptation to climate change, focusing particularly on the inter-linkages between resilience and poverty alleviation. For example, the African Climate Change Resilience Alliance (ACCRA) has been investigating climate change adaptation in Ethiopia, Uganda and Mozambique for the past five years (2009–2014). They have developed a local adaptive capacity (LAC) framework which helps to understand what drives resilience at the level of local government institutions (Jones et al. 2010). Their framework was based on five elements which were thought to reflect a high capacity for resilience and adaptive capacity: the asset base, institutions and entitlements, knowledge and information, innovation and flexible forward-looking decision-making and governance.

This chapter adopts a similar approach as the LAC framework but focuses on a lower level: it examines these five elements in the context of the food security of individuals and their individual strategies – instead of focusing on local government institutions.

Application of mixed methods to ‘resilience’

Although resilience is a relatively intuitive concept, it remains difficult to measure because it can be measured at different levels (individual, household, community or country-wide) and because resilience is shock-specific. If a household is resilient to one type of shock, it does not follow that the household is resilient to all shocks (Carpenter et al. 2001: 767).

Two types of measures of resilience can be distinguished; those which capture the capacity for resilience before a shock (ex-ante) and those which track the outcome of being resilient after the shock has occurred (ex-post). Most studies have focused on developing an ex-ante indicator to capture the capacity for resilience, because ex-post indicators can simply rely on existing approaches to tracking the recovery of income or undernutrition after a shock. For example, in the case of undernutrition,

one would track weight for height and height for age assessing the timing and severity of their fall and the length of time it takes to recover from the shock (Hoddinott 2014).

Ex-ante indicators generally take the form of a single resilience index which aggregates a variety of characteristics that are thought to enable resilience. The 2013 report of an Expert Consultation organised by the FAO and World Food Programme (WFP) reviewed 11 methodologies currently being employed by a variety of research groups, NGOs and UN agencies to measure the ex-ante capacity for resilience – including the ACCRA study mentioned above (Frankenberger and Nelson 2013).² All computed an index which aggregated a variety of characteristics such as asset holdings, income levels, food security levels, coping strategies and exposure to shocks – with only three of the 11 indices also including ‘softer’ dimensions such as social capital, wellbeing and psychological stress, knowledge and awareness of adaptation options, and the flexibility of governance structures.

Only one of the 11 methodologies explicitly adopted a mixed methods approach whereby the index which was computed using secondary data was validated and further explored through key informant interviews and focus group discussions in Haiti.³ A subsequent review conducted by the authors found only two other resilience studies which adopted mixed methods approaches based on research in Somalia⁴ and Honduras.⁵ Out of the three, only the Somali study described a clear rationale for the sequencing of qualitative and quantitative elements.

In summary, it appears that the measurement of resilience is currently dominated by quantitative approaches. While easier to track over time, the over-reliance on quantitative aggregate indicators can make trends difficult to interpret, especially the ‘softer’ dimensions listed previously. This chapter adopts a mixed methods approach precisely to unpick some of these ‘softer’ dimensions.

Methodological approach

Overview

This chapter is part of a wider study (Tincani 2012) which aimed to better understand the different causal links between ‘resilience’ and ‘food security’ at the level of the individual and the household. It took the premise that resilience and food security are self-reinforcing drivers, meaning that a resilient household is expected to be more food secure (better able to meet the food needs of their family all year round); and a

food-secure household is expected to be better able to prepare for, cope with and recover from shocks, making it more resilient.

The wider study took a three-pronged mixed methods approach, which is summarised below. It includes both ex-ante elements for the capacity for resilience (the enabling factors) and ex-post elements regarding the result of possessing or not possessing resilience (outcome indicator).

- A variety of qualitative and quantitative *enabling factors* were explored over the course of the study to understand the causal links between household resilience and household food security levels. These factors were not aggregated into a single index in order to maintain the richness of the data.
- Quantitative *process indicators* of resilience were developed, and measured over the course of the study for each household, in order to capture variations in the determining factors driving household food security. The application of these process-level indicators is novel in the field of food security and poverty research, and draws on definitions of quantitative resilience indicators from the ecological sciences.
- The household's food security level was estimated at the beginning of the study, in order to capture variation in food security between households (for simplicity this is referred to as the *outcome indicator* even though food security can be both an enabling factor and an outcome of resilience – as previously explained).

All three elements were tackled through a mixed methods approach entailing quantitative surveys, qualitative semi-structured interviews, focus groups discussions (FGD), key informant interviews and participant observations over 14 months. Table 5.1 provides an overview of the different data collection tools used. While the wider study covered all three elements, this specific chapter focuses on a discussion of the enabling factors and briefly presents the variations found in the 'process indicators' of resilience. We present the whole research design below because the data on enabling factors analysed in this chapter was collected throughout the study and benefited from the careful sequencing of different elements of the mixed methods approach to inform the design of tools and enable continuous triangulation of findings. Before data collection began, a qualitative participatory wealth ranking was undertaken to better understand the local understanding of 'food security' and to define four food security levels. Thereafter, for each of the six survey rounds during the 14 months study, each round began with

a quantitative data collection phase and was followed three–four weeks later by a qualitative data collection phase. This timing meant that during each phase the findings from quantitative data could be verified and deepened through the qualitative tools applied straight after. However, these qualitative findings also helped to continuously refine the understanding of the concepts studied, and allowed the following round of quantitative questionnaires to be refined to better capture the indicators it intended to measure. The mixed methods design enabled a detailed understanding of which strategies household members were using and why they were choosing these.

Table 5.1 Overview of quantitative and qualitative tools used to address the research questions

Research questions	Qualitative data collection and analysis tools	Quantitative data collection and analysis tools
Which livelihood strategies were members of the household engaged in?	Individual interviews used to understand which strategies are undertaken by whom, in each season.	Quantitative cooking survey used to compare how much each strategy contributed to household food security.
How and why did different strategies contribute to the household's resilience level?	Individual interviews and FGD used to understand the decisions behind why different strategies were chosen.	Pairwise ranking exercise used to identify which criteria was most important in making decisions that resulted in strategies being chosen that improved the household's resilience level.
What was the household's resilience level?	FGD used to assess community perception of historic levels of resilience.	Compound resilience indicators constructed to measure household resilience levels.
What was the household's food security level at the beginning of the study and how did it relate to the household's resilience level?	Participatory wealth ranking used to define four food security levels. FGDs used to estimate which of these four levels applied to each household (self-reported).	Correlation analysis used to assess the relationship between levels of food security and levels of resilience.

Scope of the study sample

In order to examine the link between food security and resilience, households were chosen from two provinces within the Mossi ethnic group (see Figure 5.1). The two study provinces were deliberately chosen because of their varying level of food security. While at national level, Burkina Faso is just about self-sufficient in food production, covering 105% of its food needs over the last decade, the drier Northern provinces frequently experience a shortfall in staple crops, only meeting 73% of their food needs on average (INSD 2009). As the majority of agriculture is rain-fed, differences in cereal production are significantly affected by variations in rainfall patterns, with Yatenga province receiving average rainfall of 560 mm per year and Zoundwéogo province receiving average rainfall of 1050 mm per year. Extreme variability in rainfall onset, intensity and duration with consequent droughts and floods – a typical feature of the Sahelian climate – adds a significant element of uncertainty to the cropping season.

Within each province, two villages were selected. Each of the four study villages was of a similar size, containing 150–250 polygamous

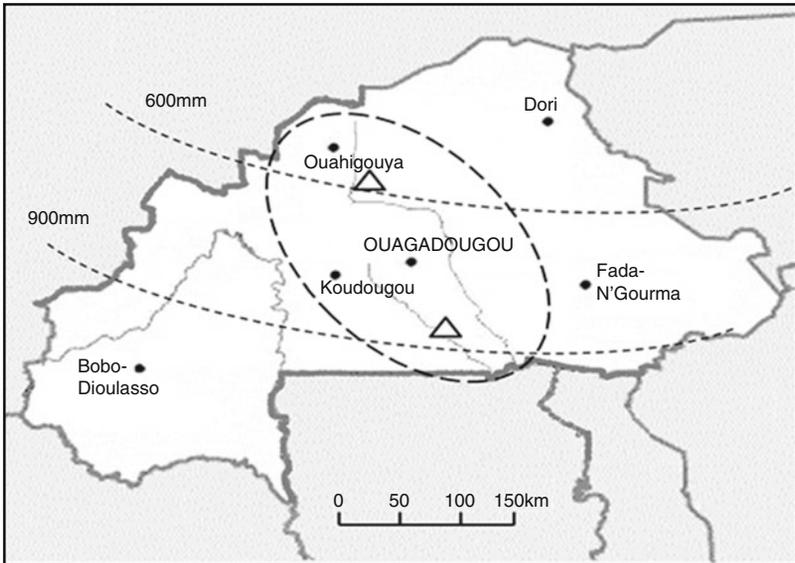


Figure 5.1 Map of Burkina Faso indicating major rivers and towns, the two field sites (triangles), rainfall gradients (dotted lines) and the area inhabited by the Mossi ethnic group (dashed circle)

families (roughly 2000–3000 people). Within each village, the study sample was selected based on varying levels of household food security; the outcome indicator for the study. First, different levels of food security were defined using a participatory wealth ranking exercise. Next, a participatory listing of polygamous families was undertaken with key informants in each study village. Following this, all households in study villages were grouped into four food security categories that had been defined in the participatory wealth ranking exercise. Finally, two polygamous families were purposely chosen per village which contained households of varying levels of food security, in order to specifically explore intra-compound variation. All polygamous families of the study were purposely chosen as those which could not assure their entire food supply from home-grown crops.

Within each household, every adult⁶ member of the polygamous family was interviewed repeatedly during the study period, resulting in a total study sample of 94 adults (see Table 5.2). Every adult member of the polygamous family was interviewed because all contribute in different ways to the family's livelihood system. In the study areas, polygamous families consist of related households communally living together in large family compounds that often include an ageing patriarch, his married sons and their wives and families. These households communally farm the large family field and share its harvest, while smaller individual fields are devolved to married sons and to each of their wives. Responsibilities relating to major expenses and to the cooking of main meals are shared by the whole compound, with minor expenses covered by households or individuals. Communal living also provides a safety net against other risks apart from agricultural variability, including disease morbidity and mortality. However, it is unclear if communal living is the best form of risk-pooling, and the delicate trade-off between individual and collective interests is one of the aspects explored in this chapter. Decision-making within polygamous families is particularly difficult to pick up through quantitative surveys and therefore allows more useful insights to be gained through the mixed methods approach adopted here.

Finally, the livelihood strategies of selected families were observed over a whole agriculture cycle, spanning a total of 14 months. This time frame was chosen due to the highly seasonal nature of livelihood activities in Burkina Faso. The dry season (February to May) is dedicated to maintenance work on the house and farm, as well as local or migratory trading activities. First rains usually begin in May and last until October. The 'rainy', 'lean' and 'harvest' seasons (May to January) are

Table 5.2 Overview of the household composition of the eight polygamous families of the study

Study province	Study village	Number of households per polygamous family	Number of wife sub-units per polygamous family	Number of active adult family members*	Number of dependents (children and elderly)**	Number of women who cooked***
Yátanga province (North)	Sima	2	5	4 men, 5 women	14	3
	Kouka-banko	4	8	7 men, 9 women	14	7
Zoundwéogo province (South)	Donsin	2	8	7 men, 8 women	10	6
		7	12	8 men, 13 women	36	11
	Kougris-sincé	1	3	2 men, 3 women	6	2
2		7	3 men, 7 women	18	5	
Total	4 villages	2	3	2 men, 3 women	6	2
		3	8	5 men, 8 women	22	5
		23	54	38 men, 56 women	126	41

(*) These comprised the study sample; (**) these were not interviewed; (***) these comprised the sample size for the quantitative cooking surveys.

dominated by agricultural tasks ranging from clearing the land, sowing and weeding through to harvesting. The lean season, known as the *période de soudure* (transition period), is thought to be the period of highest food insecurity, with food stocks running low because the previous year's dwindling staple crop reserves have not yet been replenished by the new harvest. As most rural farmers are not self-sufficient in food production, families engage in a wide range of livelihood activities such as livestock-keeping, agroforestry, trading activities and other off-farm activities in addition to farming.

Estimating household's food security levels (outcome indicator)

In order to better understand the quantitative element of the mixed methods approach, we first briefly describe how we estimated household food security levels and developed quantitative process indicators before moving on to enabling factors, the focus of this chapter.

Before data collection began, different levels of food security were defined together with a selection of key informants from the study sites using a participatory wealth ranking exercise based on local criteria, including the possession of physical assets, disposable income and subjective perception of food security (Grandin 1988). Interviewees defined four categories: 'very low', 'low', 'acceptable' and 'high' food security. This categorisation of the food security level of each household was only undertaken once – at the beginning of the study.

Developing and measuring quantitative process indicators of resilience

The process indicators of resilience were based solely on quantitative data and analysis, measured at the scale of one livelihood system (the sum of livelihood activities carried out by one polygamous family). The three indicators – diversity, co-variance and standard deviation of food sources – were based on resilience indicators that have been repeatedly tested and validated through the long-term study of resilience in ecosystems (Gunderson and Holling 2002, Holling 1973) but have not been empirically tested for livelihood systems. A full methodological discussion of these indicators can be found in Tincani (2012).

Due to the goal of the study – aiming to understand the causal link of how and why 'resilience' might increase food security – only those livelihood strategies contributing to the family's food security were used to calculate the three quantitative process indicators of resilience in accordance with Sen's Entitlement Theory⁷ (Sen 1981). In short, only

the following four strategies which were physically bringing food into the household were documented:

- *Home-grown staple foods* – principally sorghum, millet and maize;
- *Food purchased* from nearby towns with cash from agricultural sales or off-farm employment;
- *Food received* in the form of gifts from relatives, most commonly sacks of rice or millet; or
- *Food gathered* in the form of edible tree products, for example, leaves and fruits of a range of trees such as baobab.

The proportion of food stemming from each of the four pre-determined food sources – or livelihood activities – was quantified through cooking surveys, which were administered every two months from November 2009 to December 2010, resulting in six survey rounds (S1–S6). These surveys quantified in detail what each woman had cooked over the last six days, through two three-day recall questionnaires.

Contextual information was also gathered on livelihood strategies which generated income for other purposes than food purchase, in order to ensure data quality through triangulation. Contextual information on other livelihood activities such as child minding or maintenance of social relations was not recorded.

Exploring enabling factors

In order to understand the causal links between household resilience and household food security levels, a variety of enabling factors were explored using both quantitative asset surveys and qualitative tools, as well as contextual secondary data on rainfall patterns, seasonal livestock and staple crop prices. The enabling factors were grouped following the LAC framework developed by ACCRA (Jones et al. 2010). The only modification for this study was the level of analysis, which focused on the polygamous household as the ‘institution’ or ‘system’ under examination, whereas ACCRA focused on the higher level of local government institutions.

- The ownership of assets are well documented as playing an important role in determining the options households have in responding to shocks and thus conveying resilience (Chaudhuri 2003, Davies 1996). The tangible *asset base* of every household member was quantified, including land owned; monetary income and expenditure flows;

livestock assets; and food reserves remaining from the harvest, from food purchases and from collections made in forested areas (edible tree products). These asset flows – entailing both assets accumulated and assets consumed – were quantified every two months to capture the strong seasonality of wealth in Burkina Faso. They were quantified for every adult household member individually using a quantitative questionnaire to disentangle the distribution of assets within polygamous families⁸ common in the study area.

- Understanding not only the asset base but also access to assets through entitlements is essential to untangling how these contribute to conveying resilience. This research specifically focused on the household/extended family as an ‘institution’ through which access to resources is mediated. As such, *asset entitlements* were examined using the extent of asset sharing within polygamous families, through one-off qualitative interviews with every adult member of the household. These focused in particular on the ability of household members to negotiate access to resources held by other members of the polygamous family, as well as accessing common property resources held by the whole family – namely edible tree products gathered from land owned by the whole family clan.
- Understanding risks and how to adapt to them affects the ability of households to maintain their food security in the face of these risks. This ability is generally understood to be enhanced by an understanding of which future changes are likely, knowledge about adaptation options, the ability to assess options and the capacity to implement suitable interventions (Fankhauser and Tol 1997). This *knowledge base* was examined through one-off qualitative interviews, particularly focusing on the channels through which household members gained knowledge about adaptation options.
- The ability to foster innovation and support new practices is essential to allow existing practices to be altered and allow new ones to be adopted, in response to changing conditions (adaptive capacity). This study focused on *micro-level innovation*, namely initiatives taken by individual household members. These were captured through qualitative interviews held every two months, thus tracking changes in behaviour occurring against the backdrop of the seasonality of asset cycles.
- *Flexible forward-looking decision-making* is thought to be a key enabling factor of resilience, as it determines how individuals and institutions make decisions over asset allocation, in response to changing circumstances (Jones et al. 2010). This aspect was investigated through

qualitative interviews held every two months, which examined how and why household members took certain decisions. The trade-offs between decisions were examined through participatory pairwise ranking exercises.⁹ Analysis included understanding the ability of household members to influence decision-making over the use of shared resources, and how these power relations changed across seasons. Changes in attitudes to risk-taking were also mapped across seasons. Finally, the flexibility of the 'institution' of polygamous family as a whole was examined, exploring how social norms governing decision-making varied between families and between study locations.

Data analysis

The comparison of case studies was done within a factorial design¹⁰ which compared all combinations of location (northern field site and southern field site) and food security levels (three levels). Qualitative and quantitative data collection was identical across all sites. However, for statistical analysis, each case study formed a 'group', allowing the statistical comparison of the averages of key indicators within and between 'groups'.

A mixed methods approach to data analysis was adopted, entailing an iterative sequencing of quantitative and qualitative data analysis. First, the quantitative data emerging from the asset inventories and cooking surveys was analysed and tabulated to identify trends between villages, between seasons, and within households. Next, qualitative analysis was undertaken on the transcripts from key informant interviews and focus group discussions to further explore these trends, uncovering factors which could explain the differences identified between villages, between seasons, and within households. Lastly, quantitative resilience indicators were computed and analysed to assess if trends identified in the previous two rounds of analysis also held for the resilience indicators.

Limitations of the study

This study took a deliberate approach of investigating decision-making within the polygamous family, each of which contained several households. As such, the emphasis was placed on repeated and detailed data collection within the household, as opposed to covering a large number of households. The resulting study sample was small (23 households; 94 adults) and made no claim at achieving any level of local or regional representativeness. Instead, the study was intended as a case study,

providing a detailed snapshot of the decision-making processes of eight polygamous families.

Particular emphasis was placed on the processes affecting decision-making within the household, focusing on economic, cultural and behavioural drivers. While the socio-political context of the study villages and of the country as a whole undoubtedly affected behaviour, this aspect was not examined in detail by this study. The two study provinces were deliberately chosen within the same ethnic group to control for different social structures found in different ethnicities, but it is possible that local variations in the socio-political context may still have affected behaviour in different villages.

Results

This section reports the underlying factors enabling both diversification and specialisation, contextualising these by looking first at the households' resilience levels and how and why different strategies contributed to these.¹¹ The section is structured to first present the trends in resilience levels revealed through the quantitative data, and then present the qualitative findings which were used to unpick and understand the enabling factors that may have caused these trends. Note that this structure was chosen to highlight the added value of the mixed methods approach but during data collection quantitative and qualitative data was not collected in a sequence, instead it occurred in parallel over six iterative survey cycles.

Variations in resilience

This section briefly presents the findings from the quantitative element of the study, which measured variations in resilience through various 'process indicators' of resilience. To explore the robustness of the resilience indicators, we discuss whether they exhibited the pattern expected under varying levels of food security. For the level of diversification of food sources – one of the 'process indicators' of resilience measured in this study – one would expect a declining relationship between diversification and food security (Holling 2001). This idea is based on the premise that assets – in this case, food – can be successfully accumulated through fewer food sources, namely those which 'work best', thus increasing food security.

However, the data displayed no clear relationship between diversification and food security (non-significant Spearman's rank correlation). For some households, diversification appeared to be driven by need. Average

diversification levels were higher in the northern than the southern field site¹² (see Table 5.3), with qualitative analysis showing that households diversified in reaction to staple crop shortfall, particularly through purchase. Taken on a seasonal basis, households both in the northern and southern field site exhibited the highest diversification levels during the lean season and the lowest levels after the harvest¹³ (see Table 5.3), when staple crop stocks were at their highest.

It also appears other factors influenced diversification levels, as evidenced by the fact that some households diversified outside of the lean season. This effect became particularly clear when one controlled for the amount of labour available to the household. For polygamous households with more than one wife, a high level of diversification was maintained year-round, whereas monogamous households diversified only during the lean season – presumably to complement staple crop shortfalls (see Table 5.4).

Based on quantitative data alone it is neither possible to confirm or reject the claim that livelihood diversification levels affect households' resilience and food security, nor to examine the causal mechanisms for how this might be the case. As a result, the following sections discuss qualitative findings which shed light on the enabling factors that may have been causing the trends emerging from the quantitative data.

Table 5.3 Seasonal variation livelihood diversification for polygamous families in the northern and southern field sites

	Nov-Dec	Jan-Feb	Mar-Apr	Jun-Jul	Aug-Sep	Oct-Nov	Yearly average
Northern fieldsite	0.40 (0.15)	0.50 (0.18)	0.42 (0.12)	0.65 (0.15)	0.50 (0.12)	0.32 (0.14)	0.46 (0.16)
Southern fieldsite	0.40 (0.16)	0.33 (0.11)	0.47 (0.11)	0.50 (0.09)	0.30 (0.18)	0.33 (0.07)	0.40 (0.13)

Note: The standard deviation is indicated in brackets

Table 5.4 Yearly average of livelihood diversification for monogamous households and households with more than one wife

	Average yearly diversification level
Households with more than one wife	0.39 (0.16)
Monogamous households	0.45 (0.15)

Note: The standard deviation is indicated in brackets

Underlying factors enabling both diversification and specialisation

This section is structured around three of the five enabling factors introduced in the methodological section, namely asset entitlements, the knowledge base and the flexibility of decision-making. The remaining enabling factors, the asset base itself and micro-level innovation, were not found to play a significant role in supporting resilience (Tincani 2012).

Asset entitlements: Resilience is actively constructed

Table 5.5 summarises qualitative findings which suggest that resilience was not an emergent property, in other words it was not an automatic result of existing pre-conditions. This highlights the difference between resilience and vulnerability. Numerous studies have highlighted that vulnerability is higher for asset-poor households, but findings of this study suggest that the livelihood systems of asset-rich households were not necessarily more resilient; namely these systems did not necessarily result in a higher level of food security. Table 5.5 outlines examples of the decision-making of asset-rich and asset-poor households; one of the five types of enabling factors set out in the methodology. It is not only the asset levels themselves that mattered, but also the way they were managed. Asset-poor households with low food security levels did not manage to put savings away whereas those with high food security levels succeeded in saving, even if only at a very slow rate. In contrast, asset-rich households did not necessarily have high food security levels, as some invested their savings unwisely, thus depreciating their capital. In other words, resilience was not an emergent property but was actively constructed based on key decisions taken by household members.

The knowledge base: Ability to make decisions is based on the ability to anticipate future changes

As mentioned in the background section, the ability to anticipate future changes is generally understood to be enhanced by an understanding of which future changes are likely, knowledge about adaptation options, the ability to assess options and the capacity to implement suitable interventions (Fankhauser and Tol 1997).

Qualitative findings showed that this capacity was driven by a variety of factors. Regular exchanges with village elders, as well as personal experience, explained how household members had accumulated a host of potential causal drivers of food insecurity. For example, it was known that overgrazing of fallow fields degraded soils and limited their ability to produce substantial harvests in coming years. Exchanges with elders and

Table 5.5 Examples^a of decision-making observed by members of different household typologies

Typology of households	Quotes showing decisions made by household members
Asset-poor households with very low or low reported food security levels (<i>n</i> = 10)	<p>When times get really hard I use the running capital from my trading activities (90 GBP) to buy food but this makes it hard to start trading again next year. (Young male household head, April 2010)</p> <p>I would normally sell some animals during this time as prices are good, but my sheep were stolen and all my goats are pregnant females. I prefer to wait and see. If I have money maybe next month I will buy a chicken and wait until it has little ones. (Elderly male head of a polygamous family, February 2010)</p> <p>I am old. I need help to farm my field. But what is the point of paying someone to help clear a larger area to farm if you then have no money to pay someone to help you sow and harvest that land? I prefer to do it alone. (Elderly woman, December 2010)</p> <p>This year was hard because I was pregnant during the field work season, so I was not able to sow as much. I harvested little. Some of my field was flooded also. Next year will not be easy. (First wife, December 2010)</p>
Asset-rich households with very low or low reported food security levels (<i>n</i> = 2)	<p>The families who do well in this village are the ones who have relatives in town. Those can help them get access to good traders. Even if we have surplus harvest we don't know where to sell and get very poor prices. (Male focus group, January 2011)</p> <p>Last year I used all the income I gained from my surplus harvest to buy a motorcycle. It helps me get to the market to sell our produce. But now it is broken and I have little savings left to buy extra food for the lean season. (Elderly man, April 2011)</p> <p>My son got married this spring so I sold one cow to pay for his wedding, but now we have little savings left to buy extra food [for the lean season]. (Male household head, January 2011)</p>
Asset-poor households with medium reported food security levels (<i>n</i> = 6)	<p>Women just do what they can. I sold some cakes, when that did not work, I started boiling yams and selling those to the gold diggers. That worked better. You just try one at a time. If it doesn't work, you try a new way. (Elderly woman, January, 2011)</p> <p>In a good year I use the money I make to buy only one goat or sheep. That way slowly my flock will grow. (Second wife, April 2010)</p> <p>I would prefer to trade but I have no start-up capital to start doing that [activity]. For gold you need no start-up capital. You just need your physical strength. So I make my money that way. (Male household head, November 2009)</p>
Asset-rich households with medium reported food security levels (<i>n</i> = 5)	<p>I do better than my brother because I am organised. Every year I try to buy up a cow when it is small and resell it later. With the surplus I can buy cereals. (Male head of a polygamous family, April 2010)</p> <p>We do well because we help each other. I often lend my brothers money when they need some, at other times they send me some. Together we have done very well. (Elderly male head of a polygamous family, January 2010)</p>

^aAll quotes are labelled with the gender of the respondent and their position (head of the polygamous family, household head or wife). Where appropriate, it is indicated if the woman is in a monogamous marriage or is the first, second or third wife in a polygamous marriage. All men unless otherwise indicated are married. Guidance on the age of respondents is indicated through 'elderly' – referring to individuals above 50 years – and 'young', referring to individuals under 30 years of age. All other respondents are 30–50 years old.

other village inhabitants also contributed to gaining an understanding of which factors can be expected to occur more often, and which early-warning signs can be used to prompt adaptation before a crisis occurs.

Some of this 'common wisdom' was documented in proverbs – such as the saying that if rain had not fallen by mid-June, women should start collecting edible forest products to complement dwindling staple crop reserves (Bonnet 1982). Other behaviours were born out of experience, which differed between the northern and southern field sites. In the northern field site, all households skipped breakfast in the lean season, with four out of 15 households closing the main granary store all together.¹⁴ None of the southern households displayed either of these behaviours. In the northern field site, such behaviour was undertaken to save food following the commonly cited rule that 'if the granary is already half empty by March it is best to skip breakfast'. This behaviour was born out of experience, which showed that reducing food consumption during the hot dry season months (February to April) was more feasible than during the lean season (June to September) when physical strength was required for farming. In fact, meal sizes notably increased during the lean season period to provide strength to those farming, as a necessary investment to ensure that the harvest turned out well. The observed behaviour was therefore proactive, reducing food consumption to conserve grain reserves, as opposed to reactive, with meal sizes being smallest once granary reserves had run out in the lean season, just before the next harvest.

Flexible decision-making: Based on access to labour and asset pools

Remaining flexible and adaptable was cited repeatedly by interviewed households as a means to meeting the family's food needs throughout the year. Interviews revealed that three main factors enabled flexibility: (i) having access to labour and assets (land, draught animals), (ii) having sufficient savings to experiment with and (iii) having sufficient autonomy within village society to not stifle behaviour. These aspects are dealt with in turn, below. It is important to note that all three aspects varied considerably within the polygamous family, reinforcing the importance of a study design which interviewed every adult family member separately.

(i) Having access to labour and assets on demand

Economic theory has long highlighted that ownership of assets (control over a labour pool, ownership of land, draught animals and so on)

is a significant factor in determining the choices people make, based on their options. However, interviews showed that not only ownership over assets mattered. Household members such as wives and unmarried sons who owned few assets highlighted that they were still able to negotiate access to them when needed:

Life was not easy 100 years ago. Men refused to give [their wives] land because they were afraid she'd become too independent . . . or they'd just give her a bad plot of land . . . and as soon as she had worked it hard and restored soil fertility, he'd take it away out of jealousy and farm it himself. Thankfully such 'stories' are finished now. But a similar thing still happens to immigrants¹⁵ today – when a tenant asks for land, the owner sometimes takes [the plot] back just when it is starting to be productive again (elderly woman, January 2011).

I don't have money to buy a plough but I borrowed one from my brother, and helped him farm his field in return (young male household head, May 2011).

I asked the head of another clan for access to his trees, and he allowed me to gather enough to cook dinner for my children. But if anyone comes with sacks [with the intention of collecting large quantities for selling], they would be chased (monogamous wife, January 2011).

In closely knit polygamous societies assets are often shared between wives of the same husband, between households of the same family and within family clans. A sophisticated system of negotiation is in place whereby those requiring a plot of land, for example, can borrow it from a relative for a fixed period of time. The ability to negotiate access through social ties was just as important as the physical ability to access a market, such as the presence of a road which did not flood in the rainy season. As a result, even household members in remote villages in northern Burkina Faso were found to still regularly trade with major urban centres in the south of the country, thanks to a network of relatives.

The only asset for which negotiation was less effective was access to labour. As the quantitative data showed, households with more than one wife had additional labour available to procure food through additional sources. Qualitative analysis showed that the ability to access additional food sources depended on a wide variety of factors, including purchasing power, proximity to markets, desire to diversify the diet,

quality of relationships with relatives and power dynamics within the household. Additional food sources could be pursued for a host of reasons, including to accumulate food in the form of savings which can be later sold and to attenuate power imbalances within the household (Tincani 2012). In contrast, qualitative analysis confirmed that labour shortage due to few adult household members or due to pregnancy or illness was a significant factor in limiting the ability to access additional food sources.

(ii) Having sufficient savings to experiment with

A certain pool of savings was a pre-requisite to be able to experiment with different strategies and adapt to circumstances. However, interviews showed that it was not the absolute but the relative amount of savings that mattered. Women reported that if they were able to save only 15,000 CFA (approximately 18 GBP) per season, they could start a new business such as baking, allowing them to sell cupcakes in the market and gain additional income. Bi-monthly asset inventories clearly showed that savings levels were highly seasonal, with household members typically exhibiting the highest levels of savings at the start of the dry season (February–April) and least during the lean season (June–September), see Table 5.6. As a result, the diversity of income-generating strategies carried out peaked during the dry season, thanks to having sufficient savings to invest, and diminished in the rainy season.¹⁶ If successful, income-generating activities created a feedback loop which generated more income that could be reinvested in other livelihood activities.

Table 5.6 Seasonal variation of cash income and cash expenditure streams for all four households of the Tao family (all currencies indicated in CFA, with £1 = 827 CFA)

	Nov–Dec	Jan–Feb	Mar–Apr	Jun–Jul	Aug–Sep	Oct–Nov	Yearly total
Combined income	226,800	265,200	274,875	147,200	192,250	309,350	1,415,675
Combined expenditure	255,750	174,000	230,600	292,450	223,050	212,675	1,388,525
Assumed savings	-28,950	91,200	44,275	-145,250	-30,800	96,675	27,150
% of income as savings	-13	34	16	-99	-16	31	2

- (iii) Having sufficient autonomy within village society to not stifle behaviour

Household members reported that even with sufficient savings and access to assets and labour, some were constrained by social norms which hindered them from carrying out certain activities – even though these activities would have helped them improve their family's food security level.

Particularly women and unmarried sons from traditional families reported that they did not have the freedom to pursue income-generating activities, due to lack of permission from the head of the household or due to other social obligations:

I would like to buy young cows and fatten them up [to sell them]. But I cannot take care of many. I do not have the time to cut much grass for them. I am the only son left. All my brothers left for town. I need to help my father [to feed the family] (male household head, December 2010).

I wanted to start selling cakes so I asked my husband for permission. I waited until it was my turn¹⁷ and asked him then. As the proverb goes 'it is best to sleep on it' (second wife, November 2009).¹⁸

My trading is supposed to be for me. The reason it is not improving is because I keep using the profit to support the whole family [instead of re-investing it back into my trade]. As the proverb goes, 'where there is one rich man and nine poor, there will soon be ten poor' because income is invariably redistributed. I only take care of the most urgent problems [otherwise I get too many requests] (male household head, May 2010).

The more money I make, the more people ask for help so maybe it is better to stay poor. I tried making cakes to sell in the market but my grandchildren ate most of them . . . I can hardly refuse; can I accept that they go hungry? (elderly woman, December 2010).

In response, some pursued their activities in secret or deliberately disregarded certain social norms:

I would like to buy some goats with my savings so that I can later sell the little ones [and accumulate savings]. But my husband would never accept it; we are poor – he will not accept that I have more

animals than him. So I bought two [goats] anyway but kept them at my uncle's house, in a neighbouring village where I grew up.¹⁹ Now I can keep my goats [and any income from their sale] without upsetting my husband (elderly woman, December 2010).

My married son and I share the same household. I am old and I rely on him to help me. I gave him his own field so that he could grow his own crops and help us all eat better. He planted red millet, but sold all of it by December, instead of stocking and drying it for the whole family to use later (elderly male head of a polygamous family, December 2009).

I prefer to put the fertiliser I bought on my aubergines rather than keep it for the family's staple crops, that way I harvest more [aubergines] and have more income [for myself] (young unmarried man, December 2009).

Other interviewed household members chose to pursue a different strategy, preferring to take advantage of strong social norms instead of going against them. Such behaviour was very common, with many household members reporting that they would help their neighbours harvest in the expectation that they would help them in return if they ever experienced a shortage. The Mossi proverb stating that 'the best place to store your food is in someone else's stomach' was widely cited. It summarises the notion that investing in social networks, by deliberately redistributing your assets (e.g. by inviting your neighbours for dinner), allows you to secure favours which can be used as contingency in times of need. Such exchanges of favours were particularly common within complex households, with the data gathered from cooking surveys showing that at least 5–15% of staple crops consumed stemmed from gifts from neighbours.

Discussion on the application of mixed methods

The findings of the study showcase several elements of added value in adopting a mixed methods approach. They are discussed in turn below.

More holistic understanding of resilience

Adopting a purely quantitative approach would have captured the variation in resilience indicator levels between households and over time, and correlation with certain factors such as asset levels and labour

availability. It would not, however, have allowed a deeper examination of why certain households appeared to have higher levels of resilience than others. Adopting a purely qualitative approach would have revealed a multitude of different objectives driving the choice of livelihood strategies – such as the need to procure sufficient quantities of food and diversity of food types, to fulfil income-generation needs, to fulfil social obligations and to maintain social standing. It would not, however, have permitted a classification of how these objectives facilitated or hindered resilience. By combining both approaches, the quantitative data was able to reveal certain patterns – such as the relationship between labour availability and livelihood diversification – which could then be further examined to understand the differing objectives driving behaviour through semi-structured interviews with household members. This more holistic understanding is the main benefit offered over conventional resilience approaches, which has focused on purely quantitative approaches. Having a more nuanced understanding of ‘resilience’ makes findings that adopt a mixed methods approach more useable, as they better reflect the reality at hand and can therefore better inform policymaking.

Improved insight into polygamous households

The mixed methods approach also allowed dynamics within polygamous households to be better understood, which are challenging to pick up on in quantitative surveys usually addressed to the household head. For example, the approach allowed the identification of different and sometimes conflicting decision-making processes within the household. It showed that the interests of the household head are not always aligned with the interests of other household members. It also highlighted that different household members have access to different level of savings and/or control over different pools of labour. In addition, different household members have different levels of social autonomy. This suggests that each household member contributes in different ways to the resilience of the livelihood system of the polygamous family. Some may strengthen its resilience by building up assets and strong social ties while others might make it more vulnerable. While the exact causal relationships could not be unpicked, the choice of a mixed methods approach helped to reveal important themes which appear to drive the relationship between food security and resilience. As with the point above, having a better insight into the behaviour and incentive structures within polygamous households allows

a mixed methods approach to produce more useable findings, which better reflect the reality at hand.

Improved credibility of the data through triangulation of findings

The repeated sequencing of quantitative and qualitative data collection phases over six survey rounds enabled findings to be triangulated over several survey rounds and using different tools. The sequencing also allowed the tools themselves to be refined, for example, by ensuring that all household assets were included in asset inventories. Initially, the gender of livestock had not been recorded in asset inventories. It was only after the qualitative data showed the importance the gender of livestock played in asset management (namely that males were usually sold) that this indicator was included in the asset inventory. Both the triangulation of findings and the iterative refining of tools allowed the mixed methods approach to produce more credible findings.

Challenges in combining qualitative and quantitative datasets

The benefits of triangulation mentioned above were only possible when qualitative and quantitative approaches were combined in such a way that the same concept could be explored through two different approaches. This was achieved for the enabling factors discussed above, but was more challenging for the concept of 'food security'. For example, 'food security' was only captured using self-reported ranking methods, and not triangulated with other sources. As a result, the food security levels obtained were less credible than if they could have been compared against other food security indicators.

Conclusion

This study set out to understand the concept of 'resilience' in the context of rural livelihoods in Burkina Faso. It adopted a mixed methods approach, using qualitative indicators as well as quantitative indicators adapted from ecology, to examine how resilience can best be measured.

The quantitative data displayed no clear relationship between diversification – one of the quantitative process indicators of resilience – and food security. Trends did appear when the data was disaggregated by seasonality (diversification was higher in the lean season) and by household size (diversification was higher when the household head had more than one wife).

These trends could be further explained through the qualitative data which revealed that three of the enabling factors introduced in the methodological section played an important role:

- *The asset base*: By comparing the strategies used by asset-poor and asset-rich households, it became clear that not only the asset levels themselves mattered in determining food security and resilience levels, but also the way they were managed (prudently or not prudently).
- *The knowledge base* was important in driving decision-making, for example, with regard to meal sizes – where a notable difference was observed between the two field sites, and between seasons. Such behaviour was found to be the result of personal experience as well as common wisdom.
- *The flexibility of decision-making* was cited repeatedly by interviewed households as a means to meeting the family's food needs throughout the year. Interviews revealed that three main factors enabled flexibility: having access to labour and assets (land, draught animals) on demand, having sufficient savings to experiment with and having sufficient autonomy within village society to not stifle behaviour.

In summary, using the qualitative data to understand the trends in the quantitative data revealed that the high heterogeneity found in the quantitative data was due to the different ways household members made decisions. This level of detail within the household would have been impossible to pick up using a conventional survey approach addressed at the household head. As a result, adopting a mixed methods approach resulted in more usable findings, as they better reflect the reality at hand and can therefore better inform policymaking. Adopting a mixed methods approach also made findings more credible, as results could be triangulated across several data sources thus making them more accurate. The findings of this study therefore provide a more nuanced understanding of the processes underlying resilience. They stress the need for future resilience studies to move beyond only constructing resilience indexes which aggregates resilience characteristics, but should also include qualitative approaches which unpick the causal mechanisms that enable these characteristics.

Nonetheless, challenges remained in the application of combining methods in such a way that would allow a fuller triangulation of all findings. For example, the concept of 'food security' was only captured on one occasion using self-reported ranking methods, and not triangulated with other sources. This provides an important lesson in mixed methods

design: each research question must be explored through both qualitative and quantitative methods to allow findings from both approaches to complement each other. Future studies should carefully explain in their methodology how and why different quantitative and qualitative approaches were combined to answer a given research question, and what added value their combination can bring. This would encourage researchers to critically assess the design of their mixed methods approach and result in a clearer justification of how approaches were combined.

Notes

- 1 In the poverty literature, the distinction between process and outcome is also reflected in the debate on 'capability' versus 'poverty' (Dreze and Sen 1991).
- 2 Thirteen methodologies are presented in the report but two of these are ex-post resilience indicators.
- 3 This study analysed the effects of humanitarian assistance on resilience outcomes in the aftermath of the 2010 earthquake in Haiti (DRLA 2012).
- 4 This joint FAO-UNICEF-WFP initiative has not yet been published but findings were presented at the Resilience and Development Conference in Montpellier in 2014 (see <http://resilience2014.sciencesconf.org/25572>).
- 5 This was a study on the impact of climate change on social and ecological systems (USAID 2014).
- 6 An 'adult' was not defined by a particular age, but by whether the person considered themselves as an active member of the household with regard to contributing to the household's food security. In this study sample, the age of 'active adults' varied from 16 to around 80 years of age.
- 7 Sen stipulated that people access food in three 'entitlements': through their own food production; through exchange of goods for food via the market; or through food transfers from relatives. A fourth entitlement channel is also included, namely 'environmental entitlements', where informal community-level institutions grant access to communally owned resources such as edible tree products (Leach et al. 1999), a coping strategy against famine that has been amply documented in the literature (Waal 1990).
- 8 In the study area, polygamous families usually consisted of 10–25 adults. They contained several different households, each headed by one married man – most commonly, several brothers. Each married man had one or several wives, with each wife and her children forming a 'wife sub-unit'.
- 9 Pairwise ranking is a common participatory technique used to rank preferences or rank lists of factors influencing decisions, such as lack of market access or lack of credit facilities (Russell 1997).
- 10 The 23 households are unequally distributed within the factorial study design because, historically, compounds in Yatenga province are larger than in Zoundwéogo (West 2009), resulting in a larger sample of households in the northern field site.

- 11 The other research questions in Table 5.1 are covered by the wider study (Tincani 2012).
- 12 The difference was statistically significant (two-sample Wilcoxon rank-sum test, d.f. = 1, $p < 0.01$).
- 13 The difference was statistically significant in the northern field site (two-sample Wilcoxon rank-sum test, d.f. = 1, $p < 0.01$) and in the southern field site (two-sample Wilcoxon rank-sum test, d.f. = 1, $p < 0.05$).
- 14 The main granary holds the staple crop harvest of the collective household field, owned by the household head. These staple crops are usually used for communal meals. When the household head decides to 'close' this granary, forbidding its use, food is instead provided by the women – based on what they had harvested from their own small fields – and whatever additional food can be purchased with available savings.
- 15 'Immigrants' refers to people from another ethnic group that are not originally from the area.
- 16 The diversity of livelihood strategies was not only driven by the ability to experiment with different strategies, but also by the availability of labour. During the rainy season, most people were busy farming their fields and had little time to pursue other livelihood activities.
- 17 In polygamous marriages, women took turns cooking for the whole household. When it was their turn to cook, it was also their turn to sleep in their husband's house that evening. On other nights she slept in her own house. Note that every adult man and woman had their own houses.
- 18 Other studies have confirmed the double meaning of the proverb (Bonnet 1982).
- 19 In Mossi society, it is common for women to move to the village of her husband's family once she is married.

References

- Adger, W. N. (2000). Social and Ecological Resilience: Are They Related? *Progress in Human Geography*, 24(3): 347–364.
- Bonnet, D. (1982). *Le proverbe chez les Mossi du Yatenga, Haute-Volta*. Paris: SELAF.
- Carpenter, S., Walker, B., Anderies, J. M., and Abel, N. (2001). From Metaphor to Measurement: Resilience of What to What? *Ecosystems*, 4(8): 765–781.
- Chambers, R. (1997). *Whose Reality Counts? Putting the First Last*. London: Intermediate Technology Publications.
- Chaudhuri, S. (2003). *Assessing Vulnerability to Poverty: Concepts, Empirical Methods and Illustrative Examples*. New York: Department of Economics, Columbia University (draft).
- Constas, M. A., Frankenberger, T., and Hoddinott, J. (2014). Resilience Measurement Principles. *Food Security Information Network Technical Series*. Rome: Food and Agriculture Organization and World Food Programme.
- Davies, S. (1996). *Adaptable Livelihoods: Coping with Food Insecurity in the Malian Sahel*. Basingstoke: Macmillan.
- Dercon, S. (2005). Risk, Insurance, and Poverty: A Review. In S. Dercon (ed.), *Insurance Against Poverty*, Chapter 1: pp. 9–37. New York: Oxford University Press.

- Dreze, J. and Sen, A. K. (1991). *Hunger and Public Action*. Oxford: Clarendon Press.
- DRLA. (2012). *Haiti Humanitarian Assistance Evaluation from a Resilience Perspective*. New Orleans: Tulane University's Disaster Resilience Leadership Academy (DRLA).
- EIU. (2009). *Burkina Faso Country Report (August 2009)*. London: Economist Intelligence Unit (EIU).
- Fan, S., Pandya-Lorch, R., and Yosef, S. (eds) (2014). Chapter 1: Introduction. In *Resilience for Food and Nutrition Security*. IFPRI.
- Fankhauser, S. and Tol, R. S. J. (1997) The Social Costs of Climate Change: The IPCC Second Assessment Report and Beyond. *Mitigation and Adaptation Strategies for Global Change*, 1: 385–403.
- Folke, C. (2006). Resilience: The Emergence of a Perspective for Social-ecological Systems Analyses. *Global Environmental Change*, 16(3): 253–267.
- Frankenberger, T. and Nelson, S. (2013). *Background Paper for the Expert Consultation on Resilience Measurement for Food Security*. Rome: FAO.
- Grandin, B. E. (1988). *Wealth Ranking in Smallholder Communities: A Field Manual*. London: Intermediate Technology Publications.
- Gubbels, P. (2013). Escaping the Hunger Cycle: Pathways to Resilience in the Sahel. *Report Jointly Commissioned by Christian Aid, CARE International UK, Concern Worldwide, Oxfam GB, CAFOD, Plan UK, Save the Children UK, Tearfund, and World Vision UK*. Washington, DC: Groundswell International.
- Gunderson, L. H. and Holling, C. S. (2002). Resilience and Adaptive Cycles. Chapter 2. In Gunderson, L. H. and Holling, C. S. (ed.), *Panarchy: Understanding Transformations in Human and Natural Systems*, 25–62. Washington, DC: Island Press.
- Hasan, M. (2014). SDGs Open Way for 'Resilient Development'. Blogpost on post2015.org, posted 16/01/2014. <<http://post2015.org/2014/01/16/sdgs-open-way-for-resilient-development/>>
- Hoddinott, J. (2014). Looking at Development through a Resilience Lens. Chapter 3. In Fan, S. et al. (eds), *Resilience for Food and Nutrition Security*. Washington, DC: IFPRI.
- Holling, C. S. (1973). Resilience and Stability of Ecological Systems. *Annual Review of Ecology and Systematics*, 4: 1–23.
- Holling, C. S. (2001). Understanding the Complexity of Economic, Ecological, and Social Systems. *Ecosystems*, 4(5): 390–405.
- INSD. (2009). *Annuaire statistique 2009*. Ouagadougou: L'Institut National de la Statistique et de la Démographie.
- Janssen, M. A. and Ostrom, E. (2006). Resilience, Vulnerability, and Adaptation: A Cross-cutting Theme of the International Human Dimensions Programme on Global Environmental Change. *Global Environmental Change*, 16(3): 237–239.
- Jefferys, A. (2013). *The R-word – Rhetoric versus Reality in the Sahel*. IRIN news website, 4 March 2013.
- Jones, L., Ludi, E., and Levine, S. (2010). *Towards a Characterisation of Adaptive Capacity: A Framework for Analysing Adaptive Capacity at the Local Level*. London: ODI.
- Leach, M., Mearns, R., and Scoones, I. (1999). Environmental Entitlements: Dynamics and Institutions in Community-based Natural Resource Management. *World Development*, 27(2): 225–247.

- Russell, T. (1997). Pair wise Ranking Made Easy. Chapter 6. In IIED (ed.), *PLA Notes No. 28: Methodological Complementarity*, 25–27. London: IIED.
- Sen, A. (1981). *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford: Oxford University Press.
- Tincani, L. (2012). *Resilient Livelihoods: Adaptation, Food Security and Wild Foods in Rural Burkina Faso*. Unpublished Ph.D. thesis, School of Oriental and African Studies, London.
- USAID. (2014). Vulnerability and Resilience to Climate Change in Western Honduras. *African and Latin American Resilience to Climate Change Project*. Burlington, VT: United States Agency for International Development (USAID).
- Waal, A. (1990). A Re-assessment of Entitlement Theory in the Light of the Recent Famines in Africa. *Development and Change*, 21(3): 469–490.
- West, C. T. (2009). Domestic Transitions, Desiccation, Agricultural Intensification, and Livelihood Diversification among Rural Households on the Central Plateau, Burkina Faso. *American Anthropologist*, 111(3): 275–288.
- Zselezcky, L. and Yosef, S. (2014). Are Shocks Becoming More Frequent or Intense? Chapter 2. In Fan, S. et al. (eds.), *Resilience for Food and Nutrition Security*. Washington DC: IFPRI.

Section II

Evaluation Research

6

Assessing Rural Transformations: Piloting a Qualitative Impact Protocol in Malawi and Ethiopia

James Copestake and Fiona Remnant

Introduction

This chapter reports on pilot testing of a qualitative impact protocol – referred to as the QUIP – that aims to provide credible, timely and cost-effective evidence of impact based on the testimonies of intended beneficiaries of rural livelihood interventions without the need for a control group. The QUIP aims to address the perennial question of how international development agencies evaluate the impact of their work, with particular reference to the challenges faced by NGOs seeking to assist smallholder farmers with often complex agricultural and rural livelihood transformations associated with market integration and adaptation to climate change. Evidence of programme impact is potentially useful both for organisational learning and for building legitimacy through improved external accountability. Its importance has been reinforced by the seemingly inexorable rise of results-based and performance management culture in development practice (Gulrajani, 2010; Ramalingam, 2013) notwithstanding concern that this approach is undemocratic (Eyben, 2013) and can encourage what Natsios (2010) refers to as ‘obsessive measurement disorder’. While often framed in technical terms, the issue of how the impact of development interventions can realistically and credibly be evaluated has been one battleground for these debates (Camfield and Duvendack, 2014).

A central methodological issue is attribution: or how particular outcomes can reliably be causally linked to specific projects, interventions or mechanisms in different contexts. The dominant approach defines impact as the difference in the value of an outcome indicator (Y_1) for a given population after a particular intervention or ‘treatment’ (X) compared to what the value would have been for the same population if

the treatment had not occurred (Y_0) (White, 2010: 154). Putting aside the problem of consistent measurement of X and Y, a central issue is how to establish a plausible counterfactual. If the evaluator can make a large number of observations of X and Y then they can draw on well-known quantitative approaches to address this problem, including the use of randomised control designs. In contrast, the research summarised in this chapter addresses the scope for more qualitative and 'small n' approaches. Our motivation for this is that while there are a range of established qualitative impact evaluation methods to choose from (process tracing, for example) there has been insufficient empirical research into how best to employ and to adapt these to disparate kinds of development activities (Stern et al., 2012: 1; White and Phillips, 2012: 5).

Among various criticisms of quantitative approaches that rely on experimental or quasi-experimental designs, perhaps the most important concern is the feasibility of addressing the practical threats to internal validity.¹ In an immensely complex, diverse, fast changing, emergent and recursive social world, many researchers have argued that it is simply too slow and expensive to generate sufficient data using experimental or quasi-experimental designs. It may be possible to measure a large vector of variables Y for a given population and time period, and to demonstrate how they are affected by exposure to a vector of interventions or treatments X. But each set of results is specific in time and space to a vector of confounding or contextual variables (Z) that is too small to be measured reliably, or too quickly becomes out of date (Pawson and Tilley, 1994). Realist evaluation offers one counterpoint to this, emphasising the need for a cumulative process of broadening understanding of context-mechanism-outcome interactions or knowledge of 'what works for whom in what circumstances, in what respects, over which duration . . . and why' (Pawson and Manzano-Santaella, 2012: 177). This pursuit of realism can be viewed as being achieved at the expense of the precision gained from experimental methods which generate statistically significant results through artificially restricting variation in treatment and contextual vectors (Levins, 1966). In this sense, the quest for alternatives to precise quantitative methods of impact evaluation entails dealing with 'organised complexity' on its own terms, rather than through a process of deliberate reduction into a closed model with a more manageable number of variables and/or statistical properties.² Employing more than one strategy for addressing attribution can also add to the credibility of impact claims by broadening the range of possible causal mechanisms explored and revealing consistency (or lack of it) in evidence between methods subject to different sets of assumptions and potential biases.

An alternative to estimating a counterfactual on the basis of statistical comparisons between respondents subject to different levels of exposure to a project/treatment is simply to ask intended beneficiaries what they think. If we are interested in finding out whether particular men, women or children are less hungry as a result of some action, it seems ethically important as well as common sense to just ask them (Anderson et al., 2012). But even putting aside problems of construct validity (over the definition of hunger, for example) it is not obvious how easily they will be able to attribute changes in their experience to specific activities; and there may also be reasons to doubt the reliability of their responses, including confirmation bias (Haidt, 2012: 93) or a tendency to anchor their responses to what is familiar or expected (Kahneman, 2011). In this chapter we also use the term pro-project bias to refer to the possibility that someone consciously or otherwise conceals or distorts what they think they know about an activity in the hope that doing so will reinforce the case to keep it going. The instrumental value of asking people directly about attribution is practical and empirical. To what extent is it possible to find ways to benefit from their direct experience of the impact of a project in a way that is not undermined by potential pro-project bias?

The structure of the chapter is as follows. The remainder of the first section elaborates on what is meant by a credible evaluation. The second section provides a short factual description of the methodology underpinning the QUIP as designed and tested on two NGO projects in Malawi and two in Ethiopia. The third section presents selected findings from these pilot studies to illuminate the methodological discussion. The fourth section discusses three key methodological issues – attribution, confirmation bias and generalisability – and the fifth section concludes.

Defining credible impact evaluation

White (2010: 154) notes that the term impact evaluation is widely used to refer both to any discussion of outcome and impact indicators, and more narrowly to studies that explicitly seek to attribute outcomes to a specified intervention. This chapter adopts the second definition. It also allows for the possibility that specific impact assessment methods (including those within a positivist tradition) can be nested within broader (including interpretive) evaluation approaches. Attributing impact is only one issue that evaluation addresses – others including how an intervention works and whether it constitutes value for money (Stern et al., 2012: 36).

As a servant of action in a changing context, the scientific rigour of impact evaluation also has to be weighed alongside cost, timeliness and fitness, to purpose. Without rejecting the quest for consensus about

what constitutes quality in qualitative research, Hammersley (2013: 83) also favours use of the term credibility rather than scientific rigour as a criterion for assessing impact evaluation, echoing the more general distinction between reasonableness and rationality (McGilchrist, 2010).³

By credibility we refer to one party being able to offer a sufficient combination of evidence and explanation to convince another party that a proposition is reasonable in the sense of being sufficiently plausible to act upon – not rational in a logical sense, perhaps, but neither irrational. While this emphasises the importance of context and trust, the rigour with which conclusions about impact are logically derived from stated evidence and assumptions is also clearly important.⁴ A more specific approach to defining credibility with respect to impact evaluation is to agree on what constitutes reasonable evidence of causation. For example, an evaluator's claim to establishing impact (i.e. X causing Y in particular contexts) might be regarded as being credible if: (a) there is strong evidence that X and Y happened in such contexts, (b) X is described by a diverse range of stakeholders as having been a necessary component of a package of actions that are sufficient to cause Y in those contexts, (c) their explanations of the mechanism by which X caused Y in those contexts are independently arrived at and mutually consistent and (d) the counter hypothesis that they have other reasons for making the statement can reasonably be refuted. The point is not to secure universal agreement, but to be as clear and precise as possible about what can reasonably be expected in a given context. For example, our emphasis here being on qualitative methods, the definition excludes the requirement for (e) evidence of how much Y varies according to exposure to X.⁵

The idea of credible causation, based on reasonableness, can be further elaborated by specifying minimum conditions for mitigating the risks of systematic bias. The definition above, for example, proposes structures and processes of evaluation that reduce the plausibility of complicity among different stakeholders. This falls short of scientific certainty but in complex contexts it is often as much as we can hope for, particularly given the possibility that efforts to aim higher may be counterproductive in terms of cost, timeliness and policy relevance. In other words, we are not suggesting that this definition is universal or even widely accepted, rather that it is a realistic one in contexts where overcoming the attribution problem is particularly difficult.

Methodology

This section reports on action research comprising the design and testing of QUIP.⁶ Initial piloting was conducted with four projects sponsored

by international NGOs: two in Malawi and two in Ethiopia. Details of them are set out in Table 6.1. Projects 1 and 3 concentrated their activities (X) on specific crops, while Projects 2 and 4 incorporated a broader spectrum of activities intended to promote livelihood diversification. However, all of them aimed to strengthen the livelihoods and food security of selected rural households, enabling the QUIP to be designed around a common set of impact indicators (Y) listed in the second column of the table. The context of all the projects can be described as one of organised complexity arising from the presence of inter-connected, uncertain and hard-to-measure confounding factors (Z) affecting the causal links between X and Y. In both Malawi and Ethiopia these include climate change, commercialisation (Collier and Dercon, 2009; Future Agricultures, 2014), the activities of other NGOs working in the same area and the evolution of public policy (e.g. Abro et al., 2014; Chirwa and Dorward, 2013) and social protection (Wedegebriel and Prowse, 2013). In contrast to quantitative impact assessment methods, the QUIP sets out to generate differentiated empirical evidence of impact based on narrative causal statements of intended project beneficiaries without the requirement to interview a control group. Evidence of attribution is sought through respondents' own account of causal mechanisms linking X to Y alongside Z, rather than by relying on statistical inference based on variable exposure to X.

Drafts of written guidelines for the QUIP were prepared for a methodology workshop held in June 2013 and attended by staff from the University of Bath, the University of Malawi, Self Help Africa, Farm

Table 6.1 Summary of pilot projects, impact indicators and confounding factors

Interventions (X)	Impact indicators (Y)	Confounding factors (Z)
Project 1. Groundnut production and marketing (Central Malawi)	Food production	Weather
	Cash income	Climate change
	Food consumption	Crop pests and diseases
	Cash spending	Livestock mortality
Project 2. Livelihood diversification (Northern Malawi)	Quality of relationships	Activities of other external organisations
	Net asset accumulation	Market conditions
Project 3. Malt barley production and marketing (Southern Ethiopia)	Overall wellbeing	Demographic changes
		Health shocks
Project 4. Livelihood diversification (Northern Ethiopia)		

Africa, Evidence for Development, Oxfam, UK, and Irish Aid.⁷ Each section was subject to detailed discussion at the workshop, and further refined through field testing of the protocol with two NGO projects in Malawi in November 2013, and two in Ethiopia in May 2014. The guidelines cover commissioning of impact assessment, its relationship to other impact evaluation activities, sample selection, data collection methods, briefing and debriefing the field researchers, facilitating interviews, data analysis, quality assurance and use of findings.⁸

Data collection by two field researchers for each pilot study was intended to last 10 days, comprising four days of household level interviews, one day of focus group discussions and five days of data transcription. For the initial pilot studies in Malawi eight households were interviewed, and four focus groups were carried out; in Ethiopia the number of households was increased to 16, while the focus groups remained the same (sufficient to cover groups of older and younger men and women). Sample sizes were dictated primarily by constraints on time and funding, with all data collection restricted to one or two villages only, selected purposively as reasonably typical of the project area.

The field researchers were independently contracted by the University of Bath, acting as lead evaluator. They set up interviews and focus group discussions without any contact with the selected NGO or project staff, or indeed without knowledge of the project being analysed. In the absence of this information the research team entered the field with an introductory letter to relevant local officials and a list of individuals in selected villages from which to randomly draw the interview sample. They introduced themselves to respondents as independent researchers conducting a study of general changes in the rural livelihoods and food security of farmers in the selected area. The purpose of this 'blinding' procedure was primarily to reduce potential for pro-project bias on the part of respondents, and is discussed in the section on discussion. It also minimised diversion of the NGO's time and effort into impact evaluation.

The household interview schedule started by asking respondents about changes in household composition. It then worked through a series of discrete sections covering different impact domains to explore how changes in food production and other sources of real and cash income relate to changes in spending, food consumption, asset accumulation, relationships and overall wellbeing. Each domain section starts with an open-ended generative question and finishes with one or more closed questions, as summarised in the Appendix. Optional probing questions (also shown) were also available to help the interviewers sustain and

deepen the conversation. A final section asked respondents to list organisations they interact with from outside their village, and to rank their importance, thereby providing them with an additional opportunity to volunteer information about the NGO being evaluated.⁹

The researchers recorded narrative data in the field on a paper pro forma, subsequently copying it into an Excel spreadsheet with an identical layout. They then passed the data to staff at the University of Bath for analysis. Their task – having also been briefed about details of the project – was to identify and code cause-and-effect statements embedded in the data according to whether they (a) explicitly attributed impact to project activities, (b) made statements that were implicitly consistent with the project's theory of change, and (c) referred to drivers of change that were incidental to project activities. These statements were also classified according to impact domains and whether respondents described effects as positive or negative.¹⁰ A similar process was followed for analysis of the focus group data.

Findings were fed back to the NGO in the form of a ten-page report for each project (to a standard format), accompanied with an annex setting out the coded cause-and-effect statements in full. The body of these reports comprised a series of tables with frequency counts of different kinds of narrative statements. Simple quantification of responses in this way was not intended to support any kind of statistical claim. Rather it provided an initial indication of the extent of congruence in responses across the sample. At the same time the project reports encouraged readers to draw on the coded narrative statements, which were provided as an appendix. These statements were organised thematically making them easier to read, whilst retaining the richness of the original data.

Findings

Although asked only after an open-ended discussion, we start with answers to closed questions from household interviews as these reflect respondents' own overall assessment of the direction of change in selected impact indicators.¹¹ The data from Malawi shown in Table 6.2 refers to perceived changes over the previous two years.

For Project 1 (groundnut) the data indicates positive change in food production, cash income, cash spending and food consumption for all but two respondents. For Project 2 (climate adaptation) the picture is more mixed with six out of eight respondents reporting falling food production and three of them also reporting negative changes with respect to the other indicators. The final column refers to net asset

Table 6.2 Responses to closed questions: Malawi projects

Gen	Age	FP	CY	CS	FC	AA	Gen	Age	FP	CY	CS	FC	AA
Project 1 (n = 8)							Project 2 (n = 8)						
F	61	=	+	-	-	+	F	58	-	+	+	+	+
F	31	+	+	+	+	+	F	39	-	-	-	-	+
M	49	+	+	+	+	+	M	33	-	+	+	+	+
F	22	+	+	+	+	+	M	23	-	-	-	-	+
F	31	-	-	-	=	-	M	54	-	-	-	-	-
F	22	+	+	+	+	-	F	43	-	+	+	+	=
M	26	+	+	+	+	+	M	32	+	+	+	+	+
M	43	+	+	+	+	+	M	42	+	+	+	+	+

Notes: FP – Food Production; CY – Cash income; CS – Cash Spending; FC – Food consumption; AA – net Asset accumulation. See last column of Appendix table for details of questions.

accumulation, and in the majority of the cases this follows the pattern of responses to the other questions: positive changes being associated with asset accumulation (seven cases) and negative changes being associated with asset sales (two cases), possibly as a coping strategy. But the number of mixed responses is also noteworthy, including three cases where the direction of net asset accumulation bucked the trend of changes in the other indicators. Open-ended interviews offered various explanations, including negative health shocks and positive remittance flows, illustrating the complexity of household livelihood systems and the environment within which they operate.

Table 6.3 presents similar data on perceived changes (this time over the previous two years) for the Ethiopia projects. Project 3 (malt barley) reveals a consistent pattern of increasing or stable food production, cash income and food consumption, with most respondents also reporting improvements in cash spending (i.e. overall purchasing power), net asset accumulation and rising overall wellbeing. In contrast, Project 4 (livelihood diversification) reveals a more mixed picture of change. Positive responses outnumber negative for perception of food production, food consumption and overall wellbeing, but it is the opposite way round for changes in cash income and purchasing power. This illustrates a recurring theme in narrative interviews of retail price inflation eroding hard-won improvements in real income.

In the case of Projects 1, 3 and 4, the selected NGO was picked out by respondents as the main organisation working with them from outside their village, although its precise identity was often confused by reference

Table 6.3 Responses to closed questions: Ethiopia projects

G	A	FP	CY	CS	FC	AA	WB	G	A	FP	CY	CS	FC	AA	WB
Project 3 (n = 16)								Project 4 (n = 16)							
M	28	+	+	-	+	+	+	F	33	+	+	+	+	+	+
M	45	+	+	+	+	+	+	M	38	-	-	-	+	+	+
M	58	+	+	+	+	+	+	M	37	+	+	+	+	+	+
M	28	+	+	-	+	+	+	F	52	+	-	-	=	-	+
M	40	+	+	+	+	-	-	F	52	-	-	-	=	-	-
M	38	+	+	+	+	+	+	F	40	-	=	+	+	+	+
M	67	+	+	+	+	=	+	F	47	+	+	+	+	+	+
M	30	+	+	+	=	+	+	F	27	-	-	-	-	+	=
M	40	+	+	-	=	=	=	F	51	=	=	+	+	=	=
M	31	+	+	+	+	+	+	M	50	+	=	-	=	-	-
M	26	+	+	+	+	+	+	F	40	=	=	=	=	+	=
M	50	+	+	+	+	+	+	F	45	+	+	-	+	+	=
M	60	+	=	=	=	?	=	M	43	=	-	-	=	=	=
M	55	+	+	+	+	+	+	F	46	=	-	-	=	=	=
M	49	+	+	+	+	+	+	F	38	+	-	=	=	=	?
M	65	=	=	=	+	-	=	F	50	=	=	+	=	+	=

Notes: G – Gender; A – Age; FP – Food Production; CY – Cash income; CS – Cash Spending; FC – Food consumption; AA – net Asset Accumulation; WB – wellbeing. See last column of Appendix table for details of questions.

to the name of the project and/or local partners including local government extension workers. The institutional landscape was particularly confused in the case of Project 2, where the selected NGO was coordinating a project that also involved several other local agencies. Resolving these identity issues and precisely establishing who was doing what in which localities emerged as an important preliminary task to coding of the narrative data. Exactly how the selected NGOs are labelled by ‘their’ intended beneficiaries within the institutional landscape is itself potentially insightful; for example, some government and NGO projects were confused.

Table 6.4 shows the number and type of cause-and-effect statements extracted from the narrative data, juxtaposing it with closed question data already discussed. The first number indicates the number of household respondents making a statement of this kind, and the second the number of focus groups. In so doing we move from evidence of perceived change to evidence of attribution classified according to whether respondents volunteered statements that explicitly mentioned the project as causal drivers, made statements that were implicitly consistent with the project’s theory of change or were incidental to it (the note to

Table 6.4 Frequency of causal statements and responses to closed questions compared from semi-structure interviews (first number) and focus groups (second number)

	Positive responses				Negative responses					
	Expl	Impl	Inci	Unat	Closed	Expl	Impl	Inci	Unat	Closed
Project 1 (n = 8)										
Food production	5,2	2,1	1,0	0,0	6	0,0	0,3	1,0	0,0	1
Cash income	4,4	4,0	2,0	0,0	7	0,0	1,2	2,0	0,0	1
Cash spending	4,4	1,0	1,0	0,0	6	0,0	0,1	2,0	0,0	2
Food consumption	3,1	1,0	1,0	0,0	6	0,0	0,2	1,0	0,0	2
Relationships	4,1	1,0	2,0	0,0	Na	0,0	2,2	2,0	0,0	Na
Asset accumulation	2,2	0,1	2,0	0,0	6	0,0	1,1	2,0	0,0	2
Project 2 (n = 8)										
Food production	1,0	2,0	0,0	0,0	2	0,0	1,0	4,4	0,0	6
Cash income	1,1	5,3	0,0	0,0	5	0,0	1,0	3,4	0,0	3
Cash spending	0,0	2,1	0,0	1,0	5	0,0	0,0	5,3	0,0	3
Food consumption	0,0	4,0	1,0	1,0	5	0,0	0,0	2,4	0,0	3
Relationships	0,0	0,3	3,3	1,0	Na	0,0	0,0	1,2	0,0	Na
Asset accumulation	0,0	3,0	0,2	3,0	6	0,0	0,0	0,3	0,0	1

Notes: The first number in each cell refers to how many household interviews yielded such statements, and the second to how many focus groups did so. The four 'types of statement' were defined as: Expl = change explicitly attributed to the project or explicitly named project activities; Impl = change confirming or refuting the specific mechanism or theory of change by which the project aims to achieve impact, but with no explicit reference to the project or named project activities; Inci = change attributed to other forces incidental to (not related to) the activities included in the project's theory of change; and Unat = change not attributed to any specific cause. Domains refer to sections of the interviewing and focus group schedules. Na = not applicable. Analysts classified statements as positive or negative according to the impact on respondents' wellbeing as expressed by respondents themselves; an option to classify responses as 'neutral' or unclear in its impact on the stated domain was also available, and used in the coded transcript to highlight where it was unclear, but not used in the analysis tables.

Table 6.4 provides a more precise definition). The first point to note here is the frequency of explicit positive statements about impact relative to the complete absence of negative statements. The fewer explicit positive statements about Project 2 reflect at least in part the fact that project activities were less advanced in selected villages, and there was some confusion over the withdrawal of another NGO from the area. Reference to incidental negative drivers of change (many weather related) were also higher for Projects 2 and 4. In contrast, many respondents of the study of Project 3 volunteered statements about positive incidental drivers of change. Those relating to increased food production referred either to livestock rearing, vegetable (including potato) production or both, often linked to the work of government Development Agents (DAs). There were also numerous references to the benefits of government training in nutrition and gender relations adding up to a consistent story of a community of farmers that were highly tuned into and responsive to progressive government outreach. Having avoided linking the field researchers to the NGO in order to reduce the risk of pro-project bias towards the project, it is likely that these responses collectively reflect a tendency towards positive confirmation bias towards government activities.

The frequency counts presented in Table 6.4 do not convey the detail and diversity of information about causal processes in the narrative data. For example, an interesting finding about Project 1 was the mental accounting through which farmers linked income from groundnut production to the cost of fertiliser purchases for their main maize crop: a rise in income from groundnuts being discounted as of little importance if offset by the rising price of fertiliser, even though this probably would have happened anyway. This also illustrates how discrete drivers were often interestingly linked – for example, positive endorsement of help with purchasing livestock, but hedged by reference to disease and mortality problems. To give another example, there was explicit and implicit support for the NGO project activities in Project 4 (particularly irrigated vegetable production) but such statements were often combined with reference to the magnitude of the incidental negative drivers, particularly lack of rainfall. For example, one of the focus groups of older men was reported as saying the following:

As the agricultural land is so small and not suitable for crop production, many Development Agents have been advising farmers and providing training on how they can use their land for alternative sources of income. Because of the drought our income has recently been reduced. But still many farmers are struggling to make use of the

limited water in the check dams and hand dug wells to produce crops and vegetables to earn some money.

It is very easy to pull out narrative quotations such as this to support specific points, but thereby also to present the evidence in a biased or cosmetic way to support prior views. To counter this danger the data analysis for each project included a process of inductively grouping and then systematically tabulating drivers of change mentioned by at least two respondents (cf. Benini et al., 2013). The main drivers identified in this way are summarised in Tables 6.5 and 6.6. Data in brackets again indicates the frequency with which the driver was mentioned in both household interviews and focus groups. Asterisks indicate those drivers that explicitly or implicitly support or negate project theory. One unsurprising finding here is that the same drivers were mentioned repeatedly in relation to different impact indicators: the importance of advice from DAs in the case of Project 3, for example. This repetition is nevertheless important. For example, it is no surprise in the case of Project 1 that groundnut production was widely cited as improving food production, cash income and spending, and the same for new varieties of barley in Project 3. But it is significant that as crops grown primarily for sale, these factors are also mentioned as positive drivers of food consumption.

Discussion

This section critically reflects on methodological issues encountered in designing and piloting the QUIP. These are grouped into three. The first section reviews the potential of the QUIP to generate internally valid evidence of project impact, subject to the premise that confirmation bias and related problems can be addressed. The second reflects on the strategy for mitigating confirmation bias, and the third reflects on questions of sampling bias, timing and external validity. The chapter concludes with a preliminary assessment of the overall credibility and cost-effectiveness of the approach taking into account all these considerations.

Attribution

One motivation behind the action research presented here was to explore scope for addressing the problem of impact attribution not only through statistical inference based on variation in exposure of a population to project interventions but also through self-reported attribution, in the form of narrative statements from intended beneficiaries, explaining what happened to them over a period of time compared to

Table 6.5 Most widely cited positive and negative drivers of change, Malawi projects

Domain	Positive	Negative
<i>Project 1: Groundnut seed, Malawi (n = 8,4)</i>		
Food production	NGO support for groundnut crop (4,0)* NGO advice on making manure (2,2)* NGO advice on small-scale irrigation (2,1)*	Low sale price for crops (1,3)*
Cash income	NGO support for groundnut crop (5,3)* NGO pass-on livestock programme (3,2)* NGO support for farming as a business (3,0)*	Low sale price for crops (2,3)*
Cash spending	NGO support for groundnut crop (5,3)* NGO support for farming as a business (3,0)* Village savings and loan groups (3,0)	Increased prices, including food (0,3)
Food consumption	NGO support for groundnut crop (2,1)*	Increased prices, including food (0,2)
Quality of relationships	NGO support for farming as a business (1,1)*	Economic hardship (0,2)
Net asset accumulation	NGO support for groundnut crop (2,0)*	
<i>Project 2: Climate change adaptation, Malawi (n = 8,4)</i>		
Food production	NGO livestock rotation programmes (2,3)* Training in conservation farming (2,0)*	Poor weather conditions (4,4) Livestock diseases (2,0)*
Cash income	NGO village savings and loan groups (2,3)* NGO small-scale irrigation projects (2,1)*	Low sale prices for crops (2,1)
Cash spending		Poor weather conditions (4,3) Low sale prices for crops (0,2)
Food consumption	Training in nutrition (0,2)	Poor weather conditions (1,4)
Quality of relationships	NGO training in financial management (0,3) NGO village savings and loan groups (0,2)	
Net asset accumulation		Poor weather conditions (0,3)

Table 6.6 Most widely cited positive and negative drivers of change, Ethiopia projects

Domain	Positive	Negative
<i>Project 3: Malt barley seed, Ethiopia (n = 16,4)</i>		
Food production	Agricultural advice from DAs (16,4)* New varieties of barley from NGO (13,4)* Advice from DAs on livestock rearing (8,4)	
Cash income	Agricultural advice from DAs (13,4)* New varieties of barley from NGO (11,3)* Advice from DAs on livestock rearing (9,4)	
Cash spending	Agricultural advice from DAs (5,4)* New varieties of barley from NGO (2,2)* DA training in financial management (2,0)	Increase in market prices of food, fertiliser and clothes (8,0) Increase in contributions to govt. bodies (2,0)
Food consumption	Agricultural advice from DAs (8,2)* Advice on diet and nutrition from HEAs (5,4) New varieties of barley from NGO (3,2)*	
Quality of relationships	Kabele training in gender equality (10,4) DA training in working together (10,2) New varieties of barley from NGO (1,1)*	Increased work demands and competition between households (0,2)
Net asset accumulation	New varieties of barley from NGO (0,3)*	
<i>Project 4: Livelihood diversification, Ethiopia (n = 16,4)</i>		
Food production	Increased fruit and veg production (4,4)* Goat rearing (1,3)* Beekeeping (1,1)* Purchase of ox, camel or cow (2,0)	Snow in August (& shortage of rain) (6,1) Lack of water/drought (3,3) Problems maintaining livestock (2,0)* Decreased labour (2,0)
Cash income	Increased fruit and veg production (4,3)* Goat rearing (3,1)*	Lack of water/drought (5,3) Snow in August (and shortage of rain) (4,0) Decreased labour (2,0)

(continued)

Table 6.6 (continued)

Domain	Positive	Negative
Cash spending	Increased fruit and veg production (3,1)* Employment abroad (2,0)	Increased prices (5,0) Lack of water/drought (1,3) Price of fertiliser (2,0)
Food consumption	Increased fruit and veg production (5,4)* Cheaper vegetables (2,0)*	Lack of water/drought (3,1)
Quality of relationships	Sharing ideas and resources in new farming practices (4,2)*	Lack of voluntary community support (3,1)
Net asset accumulation	Goat rearing (6,4)* Increased fruit and veg production (1,1)* Purchase of ox, camel or cow (2,0)	Problems maintaining livestock (4,0)*
Overall wellbeing	Increased fruit and veg production (3,3)* Beekeeping (2,0)* Improved health services (1,1)	Lack of voluntary community support (1,1)

what *would have* happened to them in the absence of the activities being evaluated. To put it another way, the attribution strategy being explored relies on respondents being able and willing to imagine and communicate statements about change relative to a hypothetical counterfactual of zero exposure to particular activities. It is certainly not rare for us to communicate contingent statements of this kind to each other: 'if I hadn't been at the meeting I would not have got the job', for example. The tougher questions concern how much information such statements can reliably carry in different contexts, and how explicitly the contingent nature of the statement needs to be spelt out. For example, to say 'I got the job because I went to the meeting' implies causation, but is rather more relaxed, for I might still have got the job if I had gone to some other meeting instead.

The four pilot studies certainly generated lots of cause-and-effect statements of the kind 'X caused Y'. But even if accepted as unbiased and truthful, their interpretation is not easy. One observation that can be made is that relatively few statements attempted to assess the magnitude of observed impact. The most precise statements referred to the effect of new varieties of barley seed on yields (Project 3), while others downplayed the impact of project activities relative to larger forces like climate (Project 4). In line with discussion of sampling issues below,

the frequency with which certain statements about impact were made constitutes evidence of their credibility rather than magnitude or importance. Hence in most cases the magnitude of the impact per household remains unknown, and so in isolation the QUIP should therefore primarily be viewed as a method for contribution analysis rather than impact assessment.

One strategy for addressing this limitation is to use the QUIP in conjunction with more precise quantitative monitoring of changes in key variables – in other words to use mixed methods.¹² In a second round of pilot studies, ongoing monitoring surveys will be used to estimate the magnitude of changes in food security, with the QUIP providing complementary qualitative evidence from respondents of the main causes behind these changes. This can at the very least help establish limits to the magnitude of change that might conceivably be attributed to an intervention. For example, if monitoring reveals at some future date that an indicator, Y1, of household disposable income on average rose by 2% between baseline and a repeat survey, it will still be possible for the intervention to have had an average impact of more than 2% because it might have offset the negative impact of a change in some confounding variable, Z1, such as rainfall. However, claims of impact in excess of observed changes would also need to be substantiated by evidence that these confounding causal effects were indeed present. If sufficiently detailed then QUIP data on causal mechanisms can be combined with monitoring data on the relative magnitudes of key variables to construct models with which to simulate the impact on Y of different combinations of X and Z. Armed with such estimates it would then be possible to make cost-benefit calculations in order to compare the cost-effectiveness of selected projects relative to alternatives.¹³ This illustrates one example of the potential for synergy between qualitative and quantitative methods in impact evaluation that is quite different from mixed method combinations where one is used to frame the other sequentially, or they are used in parallel to obtain more robust results through triangulation.

Confirmation bias

One criticism of impact evaluation based on self-reported attribution is that it generates weak evidence on the magnitude of change. Another even more potentially damning argument is that it is particularly vulnerable to confirmation bias whether based on a respondent's effective willingness to please or a more strategic calculation that exaggerating

impact can contribute to continued or further project support. Nor is the risk of bias confined to respondents. Researchers can also accentuate the importance of project interventions by downplaying or remaining ignorant of other influences on respondents' lives, particularly given the dominance of performance management culture in development practice, prompting evaluations to focus narrowly on assessing progress towards stated project goals (Picciotto, 2014: 35).¹⁴ In contrast the QUIP approach aims to be even-handed in eliciting evidence on the impact of treatment and of potentially confounding variables.¹⁵ It thereby also seeks to explore the tension between 'exploratory' and 'confirmatory' approaches to impact evaluation (Copestake, 2014).

The QUIP pilots attempted a robust response to potential confirmation bias problems by recruiting independent field researchers in a way that meant they were unaware of the identity of the project being evaluated and the NGO implementing it.¹⁶ This emphasis on avoiding pro-project bias appears to be in tension with the argument for placing project theories of change at the heart of impact evaluation to facilitate formulation of clear and testable impact hypotheses (cf. Ton, 2012). However, the piloting of the QUIP demonstrated that this apparent tension can at least partly be resolved by separating the role of data collection from that of analysis. In other words, an exploratory data collection stage of the QUIP was nested within, but contractually separated from a confirmatory analysis stage. One feature of this strategy was the involvement of another agency to serve as lead evaluator: recruiting and briefing the lead researchers, providing them with lists of potential respondents from project staff and then carrying out the data analysis by cross-analysing the narrative data against information on the goals, activities and intended outcomes of the project. The good news from the pilots is that it demonstrated that the process of 'blinding' was indeed feasible. Lead researchers remained unclear which projects they were specifically helping to evaluate, yet the protocol nevertheless succeeded in generating a substantial amount of useful data about their impact.

At the same time, the piloting experience revealed at least four limitations of this approach to dealing with confirmation bias. First, removing the association between field workers and the implementing NGO left a vacuum in the minds of respondents that they presumably filled with other possibilities.¹⁷ In all cases the field researchers identified themselves as being affiliated with national universities; and while this may not have eliminated pro-authority bias entirely, it perhaps encouraged respondents to be more honest and hopeful. But in at least one case

(Project 3) there seems the possibility that pro-NGO project bias was replaced by a generalised pro-government bias.

A second problem is the replicability of the model used for these pilot studies. The pool of suitably qualified researchers (combining knowledge of local languages with social research skills) is limited, and being part of a UK university-sponsored research project helped to recruit some of the best, which may be more difficult for NGOs to replicate over the longer term. Although our collaborators readily understood and entered into the spirit of conducting the work blind, it could easily be misconstrued in other contexts as distrustful and is in any case hard to guarantee or sustain. Ultimately, blinding is perhaps less important than building up the pool of qualified social researchers with professional commitment to high research standards of independent evaluation and research ethics.

Third, while field researchers were left in the dark about the project this was not the case for the role of data analysts for whom knowledge of project theory was necessary in order to code whether it was consistent or not with the empirical evidence collected. This raises the question of how far they too might have been prone to bias in coding and interpretation of the data. Distinguishing between explicit and implicit attribution, deciding how far multiple positive and negative cause-and-effect statements can be unbundled and aggregation of these into groups were three of the analytical tasks that proved difficult to do in a completely mechanical and objective way. However, this point should not be overstated: the subjective space for using the written transcripts is much smaller than that available to respondents and researchers in constructing those narratives, and in principle the analytical role is also more easily audited, particularly since the coded transcripts are attached in full to the report (enabling readers to take issue with coding if they so wish).

Fourth, not being fully transparent with respondents about the purpose of interviews raises deeper ethical issues. In the case of the QUIP this did not involve an outright lie: the field researchers did indeed come from national universities and the research was indeed motivated by a broad interest in the lives and livelihoods of farmers in the selected areas. Having explained this it was made clear to respondents that their participation was entirely voluntary, and that their anonymity would be protected. It is also unlikely that concealing the identity of the NGO caused any harm. However, farmers were nevertheless deprived of information that might have prompted them to withdraw or to give voice to stronger views about the NGO, whether positive or negative.

Thus there is an unavoidable ethical choice to be made between adherence to categorical principles (such as being fully transparent as possible) and pragmatism about means (being economical with the truth) in pursuit of hopefully sufficiently important ends (more reliable and useful evaluation). While it may accurately reflect human psychology, a more contentious issue for some may nevertheless be the decision to base research methods on implicit distrust in what other people will say when presented with a fuller explanation of why the data is being collected.

These ethical issues cannot be fully posed in isolation from the wider political economy of any impact evaluation as a mechanism for accounting for the use of scarce resources, and in relation to the cost and ethics of methodological alternatives. For example, one motivation for the QUIP research was to investigate methods of impact evaluation that (a) give voice to respondents' own explanations of change rather than inferring this indirectly from often rather simple comparisons of their behaviour and (b) avoid assigning some people or villages, randomly or otherwise, into a control group that entails questioning them even when they are not benefiting directly and immediately from the project being evaluated.¹⁸ More fundamentally still there is the issue of how to balance evaluation practices with different development ends, with QUIP falling somewhere between more extractive survey approaches and more participatory and democratic approaches.

Overall confirmation bias may significantly undermine the credibility of qualitative impact evaluation, and the QUIP pilots suggest ways of addressing this. But doing so does not come without having to make compromises, and since the extent of such bias is itself very hard to evaluate or quantify it is not easy to assess how much importance should be paid to this problem in methodological design.¹⁹ While the piloting exercise illustrates that there is scope for reducing some bias through innovations such as partial blinding, these are unlikely to be able to offer the same reassurance as triangulation of findings between studies that are prone to different sources of bias.

Generalisability

The reflections above have focused on credibility of what QUIP findings reveal about the impact of each project on selected respondents, but not on how generalisable these findings are beyond the relatively small sample of project participants actually contacted and the time period – of two years or less – covered by the questions they were asked. This section first considers selection over project space, within communities

and over time. It then reviews scope for generalisation beyond project boundaries and time horizons.

For monitoring surveys that aim at precise estimation of the typical (hence overall) value of selected indicators subject to acceptable levels of statistical significance there is a relatively well understood science for sample selection. In contrast, qualitative research is designed primarily to identify not only the main causal mechanisms affecting key indicators but also unexpected outcomes; thus, criteria and processes for sample selection are unavoidably less precise. In the case of the QUIP, the ideal scenario would have been to randomly select a sub-sample of all households covered by systematic monitoring surveys, and keep open the option to augment the size of an initially small sample until it becomes apparent that additional interviews are not generating sufficient additional evidence to justify the effort. A relatively higher level of duplication of responses can be observed, for example, across the sample of 16 household interviews conducted for Project 3, than for Project 4.

The pilot studies were not able to draw samples this simply, not least because randomly selecting respondents across large and scattered project areas would have massively increased the cost of finding and reaching respondents. Consequently, selection proceeded in two stages, with an initial purposive selection of one or two villages followed by random selection of households from within them. The issue of how representative the selected villages were of the wider project area is not one that can be addressed by the procedure described above (augmenting a random sample opportunistically) because of the relatively small numbers involved. In practice, purposeful selection relied on secondary data, and the number of villages selected was limited by the constraint to limit data collection to five days for two researchers. Best practice combined two steps: documenting key sources of variation between sub-areas within the project area (e.g. agro-climatic, including altitude, and proximity to markets); and inviting knowledgeable local stakeholders to sort villages into like groups on the basis of what they anticipate being the most important sources of variation in project performance. This at least can clarify how far villages selected for qualitative studies compare with others across the project, as well as the extent of within project contextual homogeneity. It quickly became apparent, for example, that farming systems across Project 2 were hugely diversified (with maize, rice, sorghum and cassava competing as staples). In contrast, the farming system in Project 3 was relatively homogenous, with barley dominant at intermediate altitudes and giving way to wheat and oats at the

lower and upper margins, respectively, of the project area. An additional and underestimated source of factors that affected the QUIP pilot studies was variation in the nature and timing of project activities between villages. For example, in the case of Project 4, households were earmarked for one of five distinct livelihood diversification packages, and data collection was restricted to one of the two villages where they had all been introduced.

The challenge of minimising or at least clarifying the extent of geographical bias is complicated by the need to ensure adequate coverage of variation in project effects within villages and indeed within households. For example, projects may accentuate differences in access to resources between households, and feed intra-household tensions over gender and age-specific allocation of labour, cash and other resources. One way to address the second problem is through multiple interviews within each household to provide greater detail of information and gender sensitivity, but at the extra cost of doubling up on interviewers, and having to invest time in reconciling potentially inconsistent data. Separate second interviews within each household can also be difficult to arrange (due to absences for work, for example), and resolving differences in answers risks creating or accentuating tensions within the household. For these reasons QUIP interviews during the pilot stage were limited to one per household, starting with the primary respondent identified from project lists (e.g. almost entirely men in the case of Project 3), but without ruling out participation of other household members. At the same time the QUIP pilots augmented household data with exploratory gender- and age-specific focus groups to explore whether replicating discussions within small peer groups rather than a household setting might elicit different data.²⁰ For example, we hypothesised that respondents might be more likely to complain about gendered effects arising from a shift to cash cropping outside their own household and without having to refer to it specifically. Focus groups did throw up some interesting contrasts: younger people often being more positive about change than the elderly, for example. But Table 6.4 does not reveal a consistent difference across the four studies in the ratio of positive to negative statements collected through household interviews and focus groups.

In addition to respondent recruitment at the extensive and intensive margin, complex issues arise with respect to timing and frequency of interviewing (Camfield and Roelen, 2012; Devereux et al., 2012; Woolcock, 2009). With many project interventions linked to the farming cycle, the minimum period for assessing change is a year while

at the other extreme it is optimistic to expect farmers to provide a detailed account of how different drivers of change interacted over more than a two-year period. However, data for over more than two years is clearly necessary to address the sustainability of post-project impacts, implying that repeat studies are essential – particularly for projects such as the ones considered here that are profoundly influenced by long-term fluctuations and trends in market activity, climate, demography and even culture.²¹ A potential strength of qualitative assessment is that findings are separable and additive – that is, each additional interview can independently add to understanding. Additional studies can also be organised relatively quickly over time and across space – for example, in response to findings generated by routine monitoring of key indicators. They are also potentially valuable early in project design and implementation to challenge project assumptions (Lensink, 2014).²²

Overall, there are practical constraints as to how far scope for generalisation can be increased through better sampling methods without also taking into account the budget available for impact assessment in relation to the heterogeneity of activities and contexts within and between projects, and over time. The four projects reviewed here illustrate how bespoke design of projects around time and space bound technological and market opportunities are critical to supporting livelihood diversification and adaptation. Hence, while building concurrent impact evaluation into large-scale development programming can help, expanding the mix of assessment methods that can be used flexibly and iteratively is also important. The goal of the action research reported here is to develop a QUIP with a unit cost of less than £5,000 (the budget used in these pilots), that can be conducted from start to finish in a few weeks and can be scaled up and adapted to reflect changing project activities and conditions; but this assumes the protocol can build upon comprehensive and reliable quantitative monitoring of changes in key variables. A second round of pilot studies is planned for 2015, and there is clearly scope for further work both over time and in other contexts.

Conclusion

This chapter has presented results from a first round of pilot testing of a qualitative impact assessment protocol tailored to provide independent feedback on how rural livelihood and climate adaptation projects are affecting household level production, income and food

security. First, it has suggested that it is possible to address problems of attribution and contribution using qualitative as well as quantitative methods by relying on narrative accounts of drivers of change collected directly from intended beneficiaries, particularly if this can be part of a mixed methods strategy by being combined with quantitative monitoring and estimation of changes in key indicators through model-based simulation (not described here, but planned as part of a second round of piloting). Second, it has identified some scope for addressing pro-project or confirmation bias through the use of independent evaluators distanced from project implementation. Third, it has pointed towards the importance of strengthening scalable methods of research that can be used adaptively, particularly in conjunction with routine monitoring of key indicators. Despite many years of effort to improve monitoring and evaluation of rural development, considerable scope remains for improvement. While the focus of the action research reported here has been on rural livelihood transformations and their effect on relatively familiar and uncontroversial indicators of economic security, there is potential also to explore how the ideas and methods of qualitative assessment being tested relate to methods being utilised in other areas of intervention and with other indicators of wellbeing.

At a more general epistemological level this chapter is unapologetic in promoting improvement in impact evaluation through systematic research and testing, including through comparative use of multiple and mixed methods to perform the same role. At the same time it implicitly recognises that success hinges upon building trusted and sustained collaborative relationships that erode the frequently made but overdrawn distinction between research and practice. It also recognises the limitations of a positivist approach to improving development in the face of overwhelming contextual complexity and multiple stakeholder interests that spawn diverse and competing interpretations of what constitutes credible and useful evidence. More specifically, responses to problems of attribution, confirmation bias and generalisability have to be assessed against standards of construct, internal and external validity and reliability, simultaneously. Likewise the messy details of design, data collection, analysis and use also have to be tackled together. Action research, such as that reported in this chapter need not be premised on rational production of universal best solutions. Rather its purpose is to spur progress towards a wider range of more reasonable better practices, recognising that they will still be contested.

Appendix. Summary of the QUIP Household interview schedule

Section/domain headings and key generative question	Optional probing questions	Closed questions
<p><i>A. Introduction (interview details)</i></p> <p>We are conducting a study into how the income and food security of people living in this area is changing and what can be done to improve this</p> <p><i>B. Household composition</i></p> <p>Please tell me the main things that have happened to your household during this period</p>	<p>How has the composition of the household changed?</p> <p>How has your health and those of other household members been?</p> <p>What are the main reasons for these changes?</p>	<p>Name of lead respondent, contact details, place and time of interview, name of interviewer.</p>
<p><i>C. Food production and cash income</i></p> <p>C1. Please tell me how your ability as a household to produce your own food has changed since then, if at all.</p>	<p>What are the reasons for these changes?</p> <p>What about livestock?</p> <p>Have you taken up any new activities to help you produce more food?</p> <p>Is there anything you have stopped doing?</p> <p>Are you doing anything differently compared to others?</p> <p>Have you taken up any new activities for earning cash?</p> <p>Have you stopped any activities?</p> <p>Are you doing anything differently compared to others?</p> <p>What are the reasons for these changes?</p>	<p>Please can you tell me who currently belongs to your household? [Relation to interviewee, gender, age, education, residency, chronic illness or disability]</p>
<p>C2. Please tell me how your ability as a household to earn money has changed since then, if at all.</p>	<p>What are the reasons for these changes?</p> <p>What about livestock?</p> <p>Have you taken up any new activities to help you produce more food?</p> <p>Is there anything you have stopped doing?</p> <p>Are you doing anything differently compared to others?</p> <p>Have you taken up any new activities for earning cash?</p> <p>Have you stopped any activities?</p> <p>Are you doing anything differently compared to others?</p> <p>What are the reasons for these changes?</p>	<p>C3. Overall, how has the ability of your household to produce enough food to meet its needs changed in this time? [Better, No change, Worse, Not Sure]</p> <p>C4. Overall how has your cash income as a household changed over this time? [Higher, No Change, Lower, Not sure]</p>

D. Cash spending and food consumption

- D1. Please tell me how what you spend money on as a household has changed since then, if at all.
- What are the reasons for this? Is there anything you are spending more on now? Why? Is there anything you are spending less on? Why? Do you think these changes are good or bad? Are there things people consume now, but didn't before? Or no longer consume? What is the reason for these changes?
- D2. What about food consumption: how has this changed, if at all?

E. Relationships

- E1. Please tell me how relationships within your household have changed since then, if at all.
- How about changes in how decisions are made over food? Or money? Or about how work is shared out? Or use of assets? What are the reasons for these changes? What are the reasons for these changes?
- E2. Please tell me how your relationships with others living in this village have changed since then, if at all.

F. Change in household assets

- F1. Please tell me whether you have acquired or disposed of any assets during this period?
- Why did you do this? Who from/to? Describe the asset
- F2. Overall, do you feel the combined total value of all your assets has gone up or down over the period? [Up, Down, Same, Not sure]

- D3. Overall, how has what you as a household can purchase with money changed over the period? [Better, Worse, Same, Not sure].
- D4. Overall how much are you eating as a household compared to this time two years ago? [Better, Worse, Same, Not sure].

G. Change in relationships

- G1. Please list the most important links you have with organisations from outside the village. What have you done with this organisation and what difference has this made to you? How have your links changed and why? Please rank the organisations you have listed, starting with the one you value most.
- G2. Overall, taking all things into account, how do you think the wellbeing of your household has changed during this period?
- F2. Overall, do you feel the combined total value of all your assets has gone up or down over the period? [Up, Down, Same, Not sure]

Notes

- 1 Randomisation is also no guarantee against pro-project bias (White, 2010: 156), particularly if Y is obtained from respondents (and/or by researchers) who are not blind to whether they belong to the treatment or control sample, and may therefore be prone to different degrees of response bias, including Hawthorne and John Henry effects (Duvendack et al., 2011). For further discussion see Camfield and Duvendack (2014).
- 2 Complexity is much discussed, but often rather loosely. For discussion of the term 'organised complexity' see Ramalingam (2013: 134). Here we take it to mean that the influence of X on Y is confounded by factors Z that are impossible to fully enumerate, of uncertain or highly variable value, difficult to separate and/or impossible to fully control. Additional complexity arises if the nature and value of X and/or Y is also uncertain.
- 3 McGilchrist (2010) suggest humans are all capable of thinking in two distinct and complementary ways. The first more rational, depersonalised and certainty seeking abstracts and simplifies, producing narrower, more precise and focused models of the world. The second aims to be reasonable, concrete, less certain, contextual, person rather than idea oriented, emphasising difference rather than sameness and quantification over meaning. It is associated with open forms of attention and vigilance, alongside broader, contextualising and holistic ways of thinking. Much of the time we employ both together, and this confers immense potential evolutionary advantages: to think narrowly (as forensic hunter-gatherer) and broadly (as agile evader of other hunters) at the same time, for example. But that does not rule out individuals having a stronger pre-disposition towards one way of thinking over the other. Rowson and McGilchrist (2013: 30) make clear that this 'horizontal' distinction is complementary but distinct from the 'vertical' one between 'fast' and 'slow' thinking made by Kahneman (2011).
- 4 A common way of further elaborating on the credibility of evidence is to distinguish between the validity of an approach, and the reliability of results arising from its application in a particular context. However, we agree with Lewis and Ritchie (2003: 270) that this distinction is harder to sustain and therefore less useful for qualitative impact evaluation given that no study can ever be replicated in precisely the same time and setting in order to identify how far results are sensitive to implementation rather than design.
- 5 Although scope for quantification will be explored through a second round of pilot studies making greater use of ongoing monitoring data using the Individual Household Method (IHM). This is further discussed in endnote 12.
- 6 It covers work carried out between November 2012 and May 2014 as part of the three-year 'ART Project' programme of research into 'assessing rural transformations'. This is in turn funded under a joint call of the UK Economic and Social Research Council (ESRC) and Department for International Development (DFID) for research into 'measuring development'.
- 7 This in turn drew upon a QUIP designed during the 1990s to meet the specific needs of microfinance organisations that also linked in-depth impact interviews with routine quantitative monitoring of borrower or 'client' level indicators (see Imp-Act, 2004).

- 8 A draft copy of the QUIP is available at <http://www.bath.ac.uk/cds/projects-activities/assessing-rural-transformations/index.html>
- 9 This echoes the more holistic area approach to assessing impact adopted by both the WIDE programme in Ethiopia (Bevan, 2013) and PADev in West Africa (Dietz et al., 2013).
- 10 The analysis of the first two Malawi pilots was conducted in parallel by two analysts, one using bespoke Excel software and the other the qualitative analysis package NVivo. This served a quality control function (e.g. leading to identification of spreadsheet errors), and also stimulating discussion and reflection on how to improve both coding and presentation of findings. The field research teams also provided feedback on the field work process and results.
- 11 It is worth noting at this point that only eight households were interviewed in each area in Malawi, as compared to 16 in Ethiopia. The pilot will interview 24 individual respondents in the next round in Malawi, giving us valuable information on the relative advantages of different sample sizes.
- 12 In the case of the selected projects, the NGOs are monitoring the food security of intended beneficiary households using the IHM developed by the NGO Evidence for Development (EFD). This approach is based on a combination of participatory rapid rural appraisal, structured household interviewing and simulation using bespoke software. Field data is used to generate estimates of how the production, exchange and transfer entitlements (in cash and kind) of a sample of households compare with estimates of their food consumption needs based on standardised nutritional requirements and food conversion ratios. Adult equivalent entitlements for a cross section of households are then compared with a benchmark, absolute poverty threshold, and can be used to simulate the heterogeneous impact of price, output, income and other shocks, as well as the impact of project interventions.
- 13 Mueller et al. (2014) propose an alternative approach that entails using more specific questions to encourage respondents to quantify hypothetical counterfactuals.
- 14 In the absence of scope for placebos and double blind interviewing, even quantitative impact evaluation methods that incorporate 'control' groups are prone to this problem – in the form of Hawthorne and John Henry effects, for example. However, these problems can to some extent be mitigated by ensuring interview questions focus on general changes experienced by respondents, thereby concealing project intentionality and minimising (though never eliminating) differences in the way interviews with 'treatment' and 'control' respondents are framed and structured.
- 15 The repeated mention of the significant impact of the work of government agricultural experts in Project 3 is a good example of this – whilst not part of the NGO's project, the positive effects of both were inextricably intertwined, and it was important to note this relationship.
- 16 This approach may be viewed as a form of 'goal free evaluation', as reviewed by Youker et al. (2014), who also point out the dearth of systematic and empirical studies into its use.
- 17 Anthropologist Thayer Scudder once recounted being told categorically by a villager in Zambia that he must be from the government. When asked why he thought this, the villager replied 'only three sorts of outsiders come here: government people, missionaries and traders. And if a missionary or a trader

then you're the worst of either I've ever met.' The world has of course moved on, but there is still something satisfyingly robust about the generalisation that outsiders in rural areas have either political, commercial or religious motives (see Levine, 1972: 56–58).

- 18 Such respondents can be compensated with money, lottery tickets or other token gifts, but this raises still more ethical dilemmas.
- 19 It would be possible to test the blinding approach by randomly informing some respondents but not others of the identity of the NGO evaluated. However, the problem would remain of how to assess the extent to which results could be generalised to other contexts, as discussed in the next section.
- 20 More specifically the QUIP guidelines were for four focus group discussions per study (for younger men, younger women, older men and older women), with a minimum of three people present in each and a maximum of eight. The guidelines suggest inviting participation from additional members of selected households (other than the lead respondent), augmented by encouraging them to bring along a friend (the idea being to encourage freer peer discussion of more sensitive topics). In practice, selection of participants across the four studies was more ad hoc, with only 38 out of 96 belonging to selected households.
- 21 With respect to culture, the studies of Projects 3 and 4 both raised the question of how projects were responding to (and perhaps influencing) a shift towards more individualistic and competitive relations between neighbouring households, including having less time to share coffee and being less likely to offer help to those in need.
- 22 The QUIP carried out for Project 2 in this study demonstrates this: it was too early in the lifecycle of the intervention to provide much information about the effectiveness of the project, but it did provide useful information on what respondents saw as the most significant positive and negative forces affecting their livelihoods.

References

- Abro, Z. A., Alemu, B. A., and Hanjra, M. A. (2014). Policies for Agricultural Productivity Growth and Poverty Reduction in Rural Ethiopia. *World Development*, 59: 461–474.
- Anderson, M. B., Brown, D., and Jean, I. (2012). *Time to Listen: Hearing People on the Receiving End of International Aid*. Cambridge, MA: CDA Collaborative Learning Project.
- Benini, A., Chowdhury, W. S., Khan, A. A., and Bhattacharjee, R. (2013). *Reflections on Research Processes in a Development NGO: FIVDB's Survey in 2013 of the Change in Household Conditions and of the Effect of Livelihood Trainings*. A Research Note. Dacca: Friends in Village Development Bangladesh (FIVDB).
- Bevan, P. (2013). *Researching Social Change and Continuity: A Complexity-Informed Study of Twenty Rural Community Cases in Ethiopia, 1994–2015*. Oxford: Mokoro Ltd.
- Camfield, L. and Duvendack, M. (2014). Impact Evaluation – Are We 'Off the Gold Standard'?. *European Journal of Development Research*, 26(1): 1–12.

- Camfield, L. and Roelen, K. (2012). Chronic Poverty in Rural Ethiopia through the Lens of Life Histories. *Working Paper 399*. Brighton: Institute of Development Studies.
- Chirwa, E. and Dorward, A. (2013). *Agricultural Input Subsidies: The Recent Malawi Experience*. Oxford: Oxford University Press.
- Collier, P. and Dercon, S. (2009). *African Agriculture in Fifty Years: Smallholders in a Rapidly Changing World? Expert Meeting on How to Feed the World in 2050*. Rome: FAO. <http://ftp.fao.org/docrep/fao/012/ak983e/ak983e00.pdf>
- Copestake, J. (2014). Credible Impact Evaluation in Complex Contexts: Confirmatory and Exploratory Approaches. *Evaluation*, 20(4): 412–427.
- Devereux, S., Sabates-Wheeler, R., and Longhurst, R. (eds) (2012). *Seasonality, Rural Livelihoods and Development*. London: Earthscan from Routledge.
- Dietz, T., Bymolt, R., Belemvire, A., van der Geest, K., de Groot, D., Millar, D., Zaal, F. (2013). *PaDev Guidebook: Participatory Assessment of Development*. Leiden: University of Amsterdam, Tamale University for Development Studies, Expertise pour le Developpement du Sahel, ICCO Alliance, Prisma, Woord en Daad.
- Duvendack, M., Palmer-Jones, R., Copestake, J., Hooper, L., Loke, Y., and Rao, N. (2011). *What is the Evidence of the Impact of Microfinance on the Well-Being of Poor People?*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Eyben, R. (2013). *Uncovering the Politics of 'Evidence' and 'Results'. A Framing Paper for Development Practitioners*. Brighton: Institute of Development Studies (IDS). www.bigpushforward.net
- Future Agricultures. (2014). Supporting Small Farmers to Commercialise. *CAADP Policy Brief 11*. Brighton: Institute of Development Studies. <http://www.future-agricultures.org/>
- Gulrajani, N. (2010). New Vistas for Development Management: Examining Radical-Reformist Possibilities and Potential. *Public Administration and Development*, 30(2): 136–148.
- Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. London: Penguin.
- Hammersley, M. (2013). *The Myth of Research-based Policy and Practice*. Los Angeles: Sage.
- Imp-Act. (2004). QUIP: Understanding Clients Through In-depth Interviews. *Practice Note 2*. Brighton: IDS. http://spmresourcecentre.net/iprc/assets/File/PN2_QUIP.pdf
- Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Allen Lane.
- Lensink, R. (2014). What Can We Learn from Impact Evaluations? *European Journal of Development Research*, 26(1): 12–17.
- Levine, D. (1972). *Wax and Gold: Tradition and Innovation in Ethiopian Culture*. Chicago: Chicago University Press.
- Levins, R. (1966). The Strategy of Model Building in Population Biology. *American Scientist*, 54(4): 11.
- Lewis, J. and Ritchie, J. (2003). Generalising from Qualitative Research: A Guide for Social Science Students and Researchers. In Ritchie, J. and Lewis, J. (eds), *Qualitative Research Practice* (pp. 263–286). London, Thousand Oaks, New Delhi: Sage.
- Mayne, J. (2012). Contribution Analysis: Coming of Age? *Evaluation*, 18(3): 270–280.
- McGilchrist, I. (2010). *The Master and His Emissary: The Divided Brain and the Making of the Western World*. New Haven: Yale University Press.

- Mueller, C. E., Gaus, H., and Rech, J. (2014). The Counterfactual Self-Estimation of Programme Participants: Impact Assessment without Control Groups or Pretests. *American Journal of Evaluation*, 35(1): 8–25.
- Natsios, A. (2010). The Clash of the Counter-Bureaucracy and Development. *Working Paper*. Washington, DC: Center for Global Development.
- Pawson, R. and Manzano-Santaella, A. (2012). A Realist Diagnostic Workshop. *Evaluation*, 18(2): 176–191.
- Pawson, R. and Tilley, N. (1994). What Works in Evaluation Research?. *British Journal of Criminology*, 34(3): 15.
- Picciotto, R. (2014). Is Impact Evaluation Evaluation? *European Journal of Development Research*, 26(1): 31–38.
- Ramalingam, B. (2013). *Aid on the Edge of Chaos: Rethinking International Cooperation in a Complex World*. Oxford: Oxford University Press.
- Rowson, J. and McGilchrist, I. (2013). *Divided Brain, Divided World: Why the Best Part of Us Struggles to Be Heard*. London: RSA Action and Research Centre. www.thersa.org
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., and Befani, B. (2012). *Broadening the Range of Designs and Methods for Impact Evaluations*. London: DFID.
- Ton, G. (2012). The Mixing of Methods: A Three-Step Process for Improving the Rigour in Impact Evaluations. *Evaluation*, 18(1): 20.
- Wedegabriel, Z. B. and Prowse, M. (2013). Climate Change Adaptation in Ethiopia: To What Extent Does Social Protection Influence Livelihood Diversification?. *Development Policy Review*, 31(S2): 35–56.
- White, H. (2010). A Contribution to Current Debates in Impact Evaluation. *Evaluation*, 16(2): 11.
- White, H. and Phillips, D. (2012). *Addressing Attribution of Cause and Effect in 'Small n' Impact Evaluations: Towards an Integrated Framework*. London: International Initiative for Impact Evaluation.
- Woolcock, M. (2009). Towards a Plurality of Methods in Project Evaluation: A Contextualised Approach to Understanding Impact Trajectories and Efficacy. *Working Papers*. University of Manchester, Brooks World Poverty Institute.
- Youker, B. R., Ingraham, A., and Bayer, N. (2014). An Assessment of Goal-free Evaluation: Case Studies of Four Goal-free Evaluations. *Evaluation and Program Planning*, 46: 10–16.

7

Evaluating the Impacts That Impact Evaluations Don't Evaluate

Stephen Devereux and Keetie Roelen

Introduction

Development interventions aim to achieve a pre-determined set of positive improvements in the lives of people who receive them. They embody an implicit or explicit theory of change: people's wellbeing is impaired by certain deficits or challenges they face, so providing specific benefits to them will alleviate these constraints and improve their wellbeing. Impact evaluations test whether the theory of change holds in practice: did the intervention actually improve wellbeing as predicted?

The simplest way to test this is to identify relevant indicators of wellbeing and to measure them before and after the intervention. For example, if an intervention aims to reduce child hunger then children's nutrition status should be assessed pre- and post-intervention. If children's nutrition status improves over this period, the intervention has succeeded and this appears to validate the theory of change. However, it is possible that nutrition indicators improved over this period for unrelated reasons. Hence, a 'control' group of children with similar characteristics, except that they do not receive the intervention, must also be assessed at the same two points in time, and the net difference between changes in the 'treatment' group and changes in the 'control' group is declared the impact of the intervention. This is the logic underpinning 'difference-in-difference' estimations, and rigorous randomised control trials (RCTs) that follow these protocols are often considered to be the 'gold standard' for impact evaluations.

As Camfield and Duvendack (2014) acknowledge, RCTs have their limitations. They do not explain why any observed changes occurred, or why an intervention failed. They are limited to assessing pre-determined indicators, ignoring a range of other potential impacts.

They favour quantitative indicators, so are biased towards material outcomes as opposed to more subjective or social changes. Experimental research designs assume away or ‘control for’ the messy complex realities that explain actual outcomes. These limitations explain why some critics have argued instead for mixed methods as the ‘platinum standard’ (Khagram and Thomas 2010) for impact evaluations: quantitative surveys can answer the ‘how much’ questions, but complementary qualitative research is needed to answer the ‘why’ or ‘how’ questions.

This chapter builds on arguments for mixed methods or ‘Q-squared’ (Shaffer 2013) approaches to impact evaluation by focusing on two aspects of development interventions that are not adequately considered in impact evaluations: ‘programme processes’ and ‘social dynamics’. The following section examines programme processes using school feeding schemes as a case study and argues that intended impacts might not be achieved because of the way the intervention is designed and delivered, or because the theory of change is fundamentally flawed. The next section, on ‘social dynamics’, draws on several case studies to reveal how the intended impacts of a development intervention can be compromised in practice because the intervention establishes or modifies a series of key social relationships.

The chapter concludes by arguing that mixed methods are necessary to incorporate the analysis of programme processes and social dynamics into comprehensive, more accurate impact evaluations. This conclusion reinforces the growing recognition that evaluations must become more responsive to the complex, non-linear nature of poverty reduction pathways (Befani, Barnett and Stern 2014), and the increasing momentum behind mixed methods approaches for more insightful analyses of poverty and vulnerability (Kanbur and Shaffer 2007). Moreover, we add a call to critically interrogate the development intervention itself – its theory of change, design and implementation modalities – and to give due attention to unintended outcomes as well as to intended impacts.

Programme processes

Development interventions deliver inputs or resources through a process or modality that generates outcomes leading to short- and long-term impacts, some intended, others unforeseen. In theory, the intention is to influence selected aspects of wellbeing by transferring relevant resources to people whose wellbeing is believed to be impaired because of deficits in these resources. This linear theory of change is illustrated in Figure 7.1.

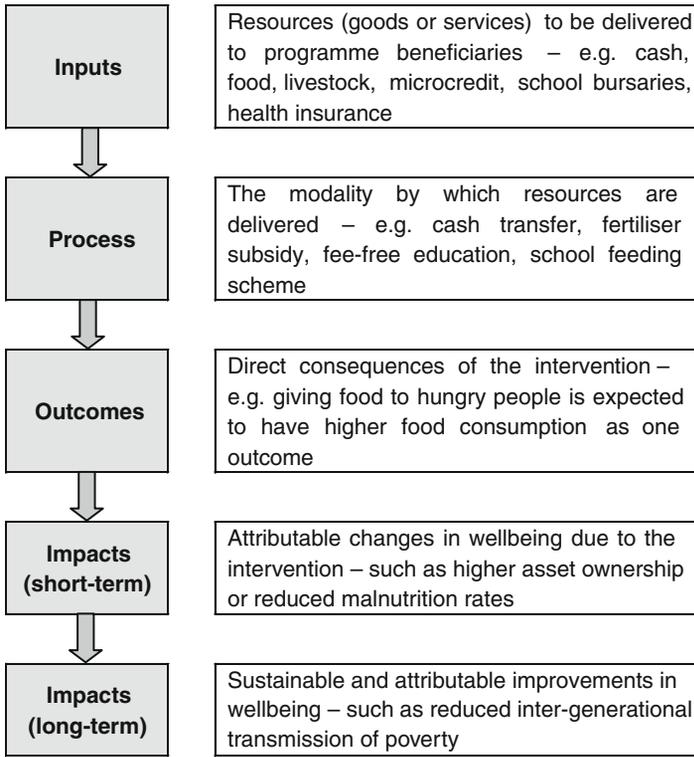


Figure 7.1 Theory of change: development interventions

Source: authors

Impact evaluations reflect but rarely challenge this simple narrative. Current dominant evaluation approaches fail to address ‘the multiplicity of contributions to development outcomes, their interrelationships, or their complex trajectories over a long period of time’ (Befani, Barnett and Stern 2014: 3). For instance, Roelen and Devereux (2014) argue that ‘programme processes’ – design choices, implementation modalities and the assumptions implicit in the theory of change – substantively affect the outcomes and impacts of development interventions, but programme processes are not subjected to scrutiny in standard evaluation methodologies.

Consider a school feeding scheme that provides a free breakfast or lunch to learners every school day. What motivates governments and donor agencies to introduce such a scheme? What is their theory of change?

Interrogating the programme processes surrounding school feeding schemes can reveal why their desired impacts are not always achieved. Giving children from poor families free meals at school is expected to generate improvements in two areas: food consumption and education. Specifically, for children who benefit from the scheme, food consumption and school attendance rates are both expected to increase. These positive outcomes are, in turn, expected to translate into measurable impacts in both areas. Higher food consumption should lead to improved nutrition status, and increased school attendance should lead to improved learner performance. Children will progress further through the primary, secondary and possibly tertiary education system, they will achieve better examination results and they are less likely to fail or drop out.

There is also a hypothesised synergy between the consumption and education outcomes that is predicted to reinforce these positive impacts: better nourished children can concentrate better in class and should outperform hungry children in tests and exams. So a single instrument – school feeding – is predicted to improve both the nutrition and education of poor children, enhancing their human capital and ultimately improving their chances of earning higher incomes and breaking the inter-generational transmission of poverty. Figure 7.2 illustrates this hypothesised causal pathway.

This is a plausible narrative that posits a linear series of effects of an intervention on beneficiaries. It may generate the expected outcomes and impacts in many cases. But in reality school feeding schemes rarely achieve such clear positive impacts, because of a range of confounding factors and methodological challenges that intervene between inputs and impacts, at each step in the theory of change. Some of these factors and challenges are discussed next.

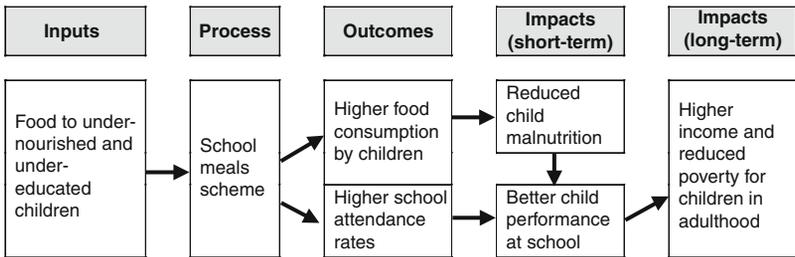


Figure 7.2 Theory of change: school feeding schemes

Source: authors

Between inputs and process

- *Were undernourished and under-educated children actually reached by the school feeding scheme?*

The assumption that school meals will be delivered to hungry under-educated children needs to be tested in evaluations. In fact the evidence suggests that school feeding schemes typically have high ‘exclusion errors’ in low-income countries, because the poorest children are least likely to be attending school so they might not be reached by this intervention. Exclusion errors are often also gendered, because in many cultures girls are much less likely to be sent to school than boys. School feeding schemes have the objective – and the effect – of attracting poorer children to school, so exclusion errors can be reduced to some extent (Bundy et al. 2009: 16). But free school meals are rarely a sufficient inducement to overcome the sociocultural barriers that keep girls out of school, unless gender targeting is introduced specifically to narrow gender gaps in access to education.

Methodological implications: Understanding the reasons why certain children do not attend school requires mixed methods investigation. Are the barriers to access for children from poor families financial (unaffordable fees and non-fee costs), logistical (no physical access to schools in remote communities), education-related (quality of teaching is perceived to be poor) or labour market-related (returns to education are perceived to be low)? Are barriers to access for girls security-related (it is unsafe for girls to walk to school, or girls are sexually harassed by teachers), policy-related (pregnant girls are excluded from school) or cultural (girls are expected to become home-based wives and mothers so educating them is considered redundant)?

School feeding schemes also usually register high ‘inclusion errors’, because targeting tends to be either universal (all children in all schools) or geographic (entire districts or entire schools rather than individual children within a school), so all poor and non-poor children attending schools running a feeding scheme will benefit equally. It is theoretically possible to restrict free school meals to poor children only, but this raises challenges of targeting (how to identify and separate out poor children) and possible jealousy from excluded classmates. As pointed out by Bundy et al. (2009): ‘While targeting individual children on the basis of need can have considerable benefits in cost-effectiveness, it has potential social costs from stigmatization’ (54). Policymakers therefore need to trade-off the financial costs of universal schemes (‘leakages’ of free school

meals to non-needy children) against the social costs of targeted schemes (marginalisation or stigmatisation of school feeding beneficiaries).

Methodological implications: Quantitative research should determine the extent of leakages to non-poor children from targeting entire schools rather than poor children within each school. This should be complemented with qualitative research to establish whether targeting poor children for school meals has negative social consequences for them – will they be stigmatised by other children?¹ This will provide policymakers with essential information they need to decide whether to introduce a universal scheme or geographic targeting, or targeting of individual poor children, to maximise cost-effectiveness and positive intended impacts while minimising any negative unintended impacts.

Between process and outcomes

- *Are school meals additional to food consumed at home?*

The assumption that food provided at school is 100% additional to food consumed at home has been tested in several countries, with highly variable results. Many studies reviewed in Kristjansson et al. (2006) found that total consumption increased by approximately the size of the transfer, but other studies have recorded a significant degree of 'substitution'. For example, 77% of children and 82% of caregivers surveyed in Malawi reported that school feeding beneficiaries receive less food at home. Studies in India and Peru that measured the net increase in calories consumed found that it amounted to less than half the calories transferred through school meals. One study in Kenya even found a net decrease in excess of 100 calories – these children would have eaten more at home if they were not fed at school. Substitution effects are generally higher in poorer and larger households. These findings have adverse implications for the impacts of school feeding schemes on children's nutrition and cognitive performance (Kristjansson et al. 2006: 10, 39).

School meals are also not provided every day – they stop over weekends and school holidays – and supplies might be interrupted or inadequate, even during school terms. 'Home-grown school feeding' (HGSF) is an innovative modality that sources food from local farmers rather than using commercially imported food or food aid, in order to stimulate the local economy and match school meals to local tastes and consumption patterns. However, local procurement is vulnerable to disruptions in supplies (e.g. due to agricultural seasonality or weather shocks) and

to challenges in terms of food safety and quality (e.g. food fortification might be needed). As indicated in a recent World Food Programme (WFP) report: 'HGSF must guarantee an uninterrupted supply of appropriate quality food to targeted children' (WFP 2009: 23).

Methodological implications: It is erroneous to assume that a child's food consumption will increase by an amount equivalent to the nutritional content of food they receive at school. Additional mixed methods research is needed into consumption patterns at the child's home and what determines intra-household food allocation rules, including the reasons for and extent of any 'substitution' effects by household size and poverty status, as well as by child age and gender, also bearing in mind the sensitivity of this issue and potential biases in survey reporting.

Between outcomes and short-term impacts

- *Did children who received school meals improve their nutrition status?*

Anthropometric evidence from several studies confirms that school feeding schemes that increase caloric intake can increase children's body size and muscle mass, while micronutrients in fortified foods, especially iron and zinc, have positive impacts on children's height (Adelman, Gilligan and Lehrer 2008: 45). One study in Kenya found that children who received a mid-morning snack of meat at school for 21 months performed better than children who received milk, others who were given an energy snack, and a control group who received nothing, on cognitive tests for abstract and perceptual abilities (Whaley et al. 2003). These results suggest that the quantity and quality of food provided through school feeding schemes have differential impacts on children's nutrition status and cognitive performance. On the other hand, it is well known that nutritional deficits in the first 1,000 days of life have permanent consequences for physical and cognitive development. From this perspective, school feeding schemes intervene too late to reverse the damage done by maternal and early childhood malnutrition, except to a marginal extent (Bundy et al. 2009).

Methodological implications: If school feeding schemes fail to significantly reduce the incidence of malnutrition among learners, this could be because the food provided at school is inadequate in quality or quantity. Evaluations from other contexts can provide lessons about the appropriate amount and kinds of food to deliver in school meals. However, if the theory of change is flawed because school feeding schemes, being a late intervention, cannot be expected to reverse

early childhood malnutrition, then nutrition indicators should not be assessed in impact evaluations.

- *Did children who received school meals perform better in examinations?*

While providing free school meals offers a positive incentive for children to attend and stay in school, this is purely a demand-side effect. If there are insufficient teachers or teachers are poorly trained, then increased attendance rates (a positive *outcome*) might not result in improved learner achievement in exams (no positive *impact*). Increasing the demand for education without investing in the supply side could even create a negative synergy. Specifically, introducing a free school meal scheme often precipitates a surge in enrolment, which can undermine the quality of education if learner-to-teacher ratios rise above optimal levels, or if teachers and administrators are diverted away from their core functions of teaching learners and managing schools. One study in India found that school feeding responsibilities took teachers away from teaching for two to three hours each day (WFP 2009). More children might get access to education, but if the quality of education falls they might perform worse in exams and failure rates could rise (a negative impact).

Methodological implications: If children who receive free meals at school do not perform better in learner assessments than they did before and compared to their peers, this could be attributed to deficiencies in the school feeding scheme, or to substitution of meals provided at home, or to problems with the quality of education provided, which might be exacerbated by the demands imposed by the school feeding scheme itself. This potential negative synergy between access to education (a positive demand-side outcome) and the quality of education (a negative supply-side outcome) requires expanding the theory of change. Understanding this complex chain of causality requires a combination of quantitative and qualitative research methods to establish the net impacts of a school feeding scheme.

Between short-term and long-term impacts

- *Do children who received free meals in school earn higher incomes and escape poverty in adult life?*

No panel studies that we are aware of have traced children over decades and isolated the effects of school feeding schemes on their incomes in adulthood. The methodological challenges of setting up such a panel study are perhaps insurmountable, or at least too complex and costly to

justify the expense. But there is credible evidence from other relevant sources suggesting that even well-nourished and well-educated children from poor backgrounds will not necessarily escape poverty in adult life. Breaking the inter-generational transmission of poverty is not achieved only by investing in the human capital (nutrition and education) of poor children. Some causes of poverty are structural rather than due to individual characteristics such as educational qualifications. For example, rigidities in the labour market or systematic discrimination against minority groups can present barriers to access to decent employment opportunities, even for well-educated school-leavers. Evidence comes from Mexico's conditional cash transfer programme, *Oportunidades*. An evaluation found that requiring children from indigenous families to attend school as a condition for the family receiving cash transfers is not sufficient to overcome the challenges they face from the poor quality of education they receive and the discrimination they face from employers when they leave school and apply for jobs (Ulrichs and Roelen 2012).

Methodological implications: Investing in children's nutrition and education effectively increases the supply and quality of labour. A quantitative assessment that finds little long-term impact of school feeding schemes on poverty reduction should not assume the programme has failed. Rather, it should be supplemented by qualitative research to understand why school-leavers failed to secure well-paid employment, and by a qualitative assessment of the labour market – is there adequate demand for labour in aggregate, and are there barriers such as discriminatory attitudes and practices by employers that effectively reduce the demand for certain types of labour, including job-seekers who were once recipients of school meals?

Social dynamics

Having elaborated the importance of considering and unpacking aspects of programme processes, we now move to examining the importance of social dynamics in evaluations of development interventions.

People live in complex networks of relationships with others, both within and beyond their family units. These relationships are characterised by inequalities of power that are typically overlooked when programmes are designed, implemented and evaluated. Yet the impacts of an intervention could be affected directly by these social relations, or the intervention could have effects on social relations which, in turn, impinge on impacts. It has long been known that development interventions affect and are affected by their sociocultural context. USAID (1995: 1) argued that development planners should consider

Table 7.1 Typology of social relations around development programmes

Level	Social relationships		
Intra-household	Female	↔	Male
	Older generation	↔	Younger generation
	Biological children	↔	Non-biological children
Intra-community	Beneficiary	↔	Recipient
	Beneficiaries	↔	Non-beneficiaries
Programme actors	Beneficiaries	↔	Programme staff
Economic actors	Beneficiaries	↔	Traders
Political actors	Beneficiaries	↔	Local politicians

both: ‘the compatibility of the project with the sociocultural environment in which it is to be introduced’, and ‘the social impact on distribution of benefits and burdens among different groups, both within the initial project and beyond’.

So impact evaluations should look in two directions: at how the socio-cultural context affects the intervention and at how the intervention affects the sociocultural context. Table 7.1 presents a partial typology of social relationships that are potentially affected by an intervention, or that can impinge on the intervention’s effectiveness. This section presents case studies of each of these relationships.

Intra-household dynamics

Within the household or family unit, power imbalances between genders (men and women, boys and girls) and generations (older persons, working-age adults, children), as well as between blood relatives and others (e.g. biological vs adopted children) can intervene to subvert the intended impacts of an intervention.

Female vis-à-vis Male

Patriarchal gender relations and assigned gender roles within families can substantively affect the outcomes of development interventions, both in terms of who benefits and in terms of impacts on gender relations.

An evaluation of a public works programme in Ethiopia found that, on aggregate, higher household incomes benefited children through reduced needs for child labour and higher school attendance rates. However, for some groups of households, girls significantly increased their hours of labour, mainly on domestic chores rather than income-earning activities, probably because they were substituting for mothers

who were engaged in public works projects (Hoddinott et al. 2011). This finding, that some girls were left worse off than before the intervention, qualifies the headline conclusion that child wellbeing was enhanced overall by the public works programme and illustrates the imperative for gender-disaggregated impact evaluations.

An evaluation in the 1990s of the income and empowerment effects of microcredit programmes in Bangladesh (pithily titled ‘Who takes the credit?’) found that women were increasingly targeted as the recipients and intended beneficiaries of loans made by microfinance institutions. However, many women were required to transfer the money they borrowed to their husbands, due to patriarchal norms dictating that men should control all cash coming into the household. Some women were actually left worse off than before, because their husbands used the loan money for their own enterprises or for personal consumption, leaving their wives with the responsibility of repaying the loan (Goetz and Sen Gupta 1996). This study highlights the need to look beyond ‘first order’ outcomes – did women receive the loans – and to assess ‘second order’ outcomes – did women control the use of these loans – in evaluating actual programme impacts.

Older generation vis-à-vis younger generation

Dynamics between generations, particularly in multi-generational households, can greatly affect how transfers or returns from development interventions are negotiated within households and who benefits from them.

A qualitative evaluation of South Africa’s Child Support Grant found that the regular receipt of monthly cash transfers instilled confidence in primary caregivers, many of them young women. Nevertheless, findings also suggested that decision-making processes regarding the use of the grant can lead to tensions in multi-generational households, with grandmothers receiving or making decisions about the use of the grant for grandchildren on behalf of their young mothers (DSD, SASSA and UNICEF 2011). With respect to a cash transfer pilot project in Lesotho, Slater and Mphale (2008) found that generational conflicts are of concern in households with elderly widows caring for orphans and vulnerable children, as they find it difficult to respond to demands on the sharing of household food and money or use of livestock, particularly with respect to older children. These findings suggest that evaluations should pay greater attention to generational dynamics within households and the ways in which they impact upon decision-making processes.

Biological children vis-à-vis non-biological children

Despite well-documented positive impacts of cash and in-kind transfers on child wellbeing and good intentions to target such impacts to the particularly vulnerable group of orphans, intra-household dynamics may undermine such impacts and potentially cause unintended negative consequences.

In terms of the caregiver–child relationship, non-biological children are widely known to be at a higher risk of exploitation, often being considered as a source of labour within the family and receiving lower quality care and attention. Targeting non-biological children with cash or in-kind transfers may reinforce these risks, as evidenced in research on linkages between social protection and children’s care in Rwanda and Ghana (Roelen and Karki Chettri 2014; Roelen and Shelmerdine 2014). Adults (including caregivers of biological and non-biological children) indicated that while the provision of transfers constitutes much-needed support for most caregivers of non-biological children, it may also play into bad intentions when caregivers are primarily motivated by economic gain. Rather than going to school like their biological siblings, such children may serve as domestic workers or be put to work on the family farm, for example. Interventions aiming to support orphans and vulnerable children must be cognisant of the potential for perverse incentives.

With respect to the relationship between biological and non-biological children, the explicit targeting of non-biological children can lead to tensions following stigmatisation or power dynamics. Social workers in Botswana, for example, indicated that orphans receiving in-kind transfers used their position to negotiate greater shares in the distribution of household food and money (Roelen et al. 2011). Thus, while programmes directly targeting non-biological children may improve material outcomes, they may also lead to intra-household tensions and marginalise individual children.

Beneficiary vis-à-vis recipient

There is sometimes a distinction between the recipient of a social grant and its intended beneficiary. For instance, cash transfer programmes that target children usually give the grant money to parents or carers rather than to the children themselves. As a consequence, the child does not necessarily receive the full benefit of the intervention that was designed to assist them. When asked how she uses the Child Support Grant (CSG) cash that she collects for her daughter each month, one South African mother replied: *‘Basically it is for the child, but because of poverty this*

money is used to assist in the household food requirements' (DSD, SASSA and UNICEF 2011: 46). The methodological implication is that evaluations should never assume that the full value of an intervention will accrue to the direct beneficiary. In this case, any increase in food consumption that is attributable to the CSG is shared among the beneficiary child's family, and the average CSG household in South Africa has just over six members. So the food consumption and potential nutritional impacts of the CSG are diluted among several 'secondary beneficiaries'.

A similar dynamic arises when the beneficiary is unable to collect his or her benefit due to illness or infirmity, and nominates a proxy or 'designated person' to collect the benefit on their behalf. Evaluation results from a cash transfer programme in Tigray, Ethiopia, suggest that dynamics between the main beneficiary and proxy or designated person allowed to collect the transfer on behalf of the main beneficiary can play an important role in the final impact of the transfer (Berhane et al. 2012). As the programme was targeted to the extreme poor and those unable to work, more than half of the beneficiaries in the rural area were unable to travel the relatively long distance to physical pay points and therefore relied upon a designated person to collect the transfer on their behalf. Survey data and interviews with key informants suggested that designated persons are not reliable to collect transfers in timely fashion and asked for money or favours in return. Programme impacts may be undermined following the dynamics between beneficiaries and designated persons as these led to beneficiaries receiving fewer funds than intended, and at more infrequent intervals.

Intra-community dynamics

Beyond the household, any intervention that delivers benefits (e.g. cash, food, assets, free services) to some members of a community and not to others immediately divides that community into two groups: 'beneficiaries' and 'non-beneficiaries', which can create tensions between the two groups.

Beneficiaries vis-à-vis non-beneficiaries

An evaluation of the Zimbabwe Emergency Cash Transfer (ZECT) pilot project in 2010 found that the choice of transfer modality led to differentiated social consequences. ZECT transferred food to some households and an equivalent value of cash to others, and compared the outcomes. Both food and cash transfers registered positive material impacts on intended indicators, notably consumption of staple food, with no inflationary impact from cash transfers on local food prices.

Cash transfers also enabled recipients to pay for education and health care, which made cash more flexible and more effective than food transfers in meeting multiple basic needs.

However, there were severe unintended social consequences. Recipients of cash transfers reported experiencing resentment, jealousy, even 'hatred' from non-beneficiary neighbours. Although there is a culture of sharing food in rural communities, sharing cash is much less common, and the switch from food aid (seen as a basic need) to cash transfers (seen as a sign of wealth) clearly had profound negative effects on social relations. The consequences of this deterioration in intra-community relations were felt in tangible ways. Some households that received cash reported that they could no longer rely on their neighbours if they needed to borrow agricultural implements, to share childcare responsibilities or to work together on communal activities (MacAuslan and Riemenschneider 2011). Evidently, the introduction of cash transfers to rural Zimbabwean communities disrupted traditional norms of reciprocity and initiated a process of commercialisation of informal social support systems.

Programme actors

Actor-oriented approaches recognise the significance of the relationships between the key actors involved in development interventions (Long 2001). Imported development models that are not well adapted to local context specificities, and that do not acknowledge the different expectations and interests of people who deliver the intervention and people who receive it, often fail for reasons that will not be revealed by impact evaluations that deploy standard evaluation methodologies to assess a simplistic theory of change.

Beneficiaries vis-à-vis programme staff

In the 1990s ActionAid established women's Savings and Credit groups in Ghana, for purposes of disbursing small micro-enterprise loans. An anthropologist undertook an 'interface analysis'² of this programme by attending many group meetings over one year and observing interactions between staff and participants.

One group was visited each week by a young male ActionAid fieldworker who arrived on his motorcycle and remained seated on it while each woman came to him to make her weekly loan repayment or savings deposit. Within 10–15 minutes this task was completed and he left. A second group was visited weekly by a young female fieldworker who arrived on her bicycle which she would dismount to sit on the ground in

a circle with the women. After talking with them for some time, advising them on their businesses and hearing about their families and other issues, she would collect their repayments and deposits and leave. While the atmosphere in the first group's meetings was formal and slightly tense, the second group's meetings were friendly and relaxed. Within one year the first group disbanded without fully repaying their loans, but the second group repaid in full and most of their ActionAid-supported micro-enterprises were thriving.

Solomon (2003) concluded that field staff should be trained not only in their core functions of disbursing and collecting loans and savings but also in how to interact with their clients, since programme outcomes were mediated by the different priorities and approaches of each individual fieldworker. Solomon also recommended that evaluations of development programmes should include an assessment of fieldworker-client relationships, given how profoundly this aspect affected programme performance in this case study.

Economic actors

The involvement of the private sector in development interventions, either directly or indirectly, sets up a relationship between programme participants and economic agents such as banks or traders. Some programmes are undermined by corruption or inefficient implementation. How participants are treated by these agents can also affect the effectiveness of the intervention. Here we consider the role of traders in programmes that transfer cash to poor people.

Beneficiaries vis-à-vis traders

Many cases exist of traders exploiting the dependence of cash transfer beneficiaries on them to profit from this dependence. This can happen directly, if private traders are sub-contracted to pay out cash transfers, or indirectly, if traders take advantage of pay-days to profiteer. Below we provide two examples. An example of private actors benefiting from being delivery agents comes from the social grants in South Africa, which can be collected from supermarkets as well as government offices, post offices and banks. An example of profiteering comes from a cash transfer project in Swaziland, which boosted the income of local retailers as well as that of project participants.

South Africa's social grants can be collected at designated retail outlets using a swipe card but many retailers abuse this facility by imposing a compulsory spending rule, arguing that cashiers cannot open the till unless the beneficiary buys something. In some cases the amount spent

can be nominal, but other retail chains insist that the beneficiary spends 10% (R24) or even up to 20% (R50) of their grant money in the store. (*'When I collect the money at Rhino the shop-owner requires me to spend at least R24 of the money that I am coming to collect.'*) Among grant recipients interviewed this practice was the main source of complaints made against retailers (DSD, SASSA and UNICEF 2011: 38).

A combination of a half ration of food and the equivalent value of cash was delivered to 40,000 drought-affected people in Swaziland for six months during 2007/2008. A pre-intervention feasibility study predicted that the cash transfer would push up local food prices by only 5–7%, but in project communities the retail price of maize actually increased by over 35% during the intervention period. Opportunistic behaviour by traders was observed, especially on pay-days, when traders established temporary markets near the queues of people waiting to collect their cash, selling a wide diversity of commodities for prices typically 10–20% or even up to 50% higher than normal. Local shopkeepers also reported surges in demand and higher turnover on pay-days, but 'cash recipients claimed that shop owners raised the prices of basic commodities on cash collection days, taking advantage of the "windfall income" that caused recipients to be less price conscious than usual' (Devereux and Jere 2008: 38).

These market responses to South Africa's social grants and a humanitarian intervention in Swaziland could be interpreted either positively, as a stimulus to local economic activity, or negatively, as exploitation of vulnerable poor people by unscrupulous traders. Either way, both cases offer clear evidence of a modified relationship between beneficiaries and local economic actors that should be captured and analysed in evaluations of cash transfer programmes, but is often overlooked.

Political actors

Political actors also play an important role in positively reinforcing or undermining programme impact. Issues of clientelism or elite capture, for example, can divert programme resources away from the intended beneficiaries.

Beneficiaries versus local politicians

A public works programme in the western hills of Nepal – the Karnali Employment Programme (KEP) – aims to reduce poverty and improve livelihoods by providing cash-for-work to poor households in local communities. Recent reports, however, suggest that a lack of functioning accountability mechanisms between stakeholders at district, village

and community levels and programme participants leads to inclusion of non-eligible programme participants and misappropriation of funds (NPC 2012). In one particular community, the owner of the land across which a road was being built as part of the programme demanded to be a participant, despite not being poor. In other communities, funds earmarked for KEP projects were used to pay private contractors to do the job, or allocated to other purposes altogether (Roelen and Karki Chettri 2015). This example illustrates the importance of strong accountability mechanisms as part of development interventions, particularly by including procedures for programme beneficiaries to hold politicians and officials to account at various levels of government.

Assessing social impacts

Harrison (2014: 39) points out that most evaluations limit themselves to considering observables only, in effect 'looking for keys where the light is better'. Evaluations of development interventions should account for their unintended 'difficult to measure' consequences as well as their intended quantifiable impacts. Although some interventions explicitly attempt to transform social relations (e.g. to empower women or to reverse discriminatory attitudes and practices), these are a minority and for simplicity we equate 'intended' with 'material' impacts (on income, nutrition, education, assets, etc.) and 'unintended' with 'social' impacts (on gender relations, stigma, jealousy, etc.).

Also for our purposes, 'intended material impacts' operate mainly at the level of targeted individuals or households, while 'unintended social impacts' operate between individuals or households and extend to relations with non-beneficiaries. ('Second order' material impacts, such as the effects of cash transfers on prices and local economy multipliers, are also sometimes evaluated [Davies and Davey 2008; Kagin et al. 2014].) From a methodological perspective, intended material impacts are typically observable and quantifiable while unintended social impacts are less easily observed and are more amenable to qualitative than quantitative research methods. This requires an extension of the evaluation toolkit beyond the use of randomised control trials or other quantitative techniques, to include more qualitative and participatory methods.

Figure 7.3 offers a holistic tool for assessing the intended and unintended impacts of a development intervention in combination, in a matrix format. Standard impact evaluations are oblivious to social impacts – in effect, they assume the intervention being evaluated had a neutral social impact, so they consider only the middle column in Figure 7.3. But even if the evaluation records improvements in the

		Unintended (social) impacts		
		Positive	Neutral	Negative
Intended (material) impacts	Positive	'Double success' [+/+] Intervention achieved its objectives and also had unplanned, beneficial social consequences	Success [+/=] Intervention achieved its material objectives and had no discernible social consequences	'Mixed result' [+/-] Intervention achieved its material objectives, but had unplanned negative social consequences
	Neutral	'Qualified success' [=/+] Intervention did not achieve its objectives, but recorded improved social indicators	No impact [=/=] Intervention effectively never happened; it had no discernible intended or unintended impacts	'Qualified failure' [=/-] Intervention did not achieve its objectives and left beneficiaries worse off on social indicators
	Negative	'Failure plus' [-/+] Intervention left its beneficiaries worse off in material terms but better off on social indicators	Failure [-/=] Intervention left its beneficiaries worse off in terms of its intended material indicators	'Double failure' [-/-] Intervention left its beneficiaries worse off in both material and social indicators

Figure 7.3 A matrix for assessing intended and unintended impacts

Source: authors

material wellbeing of beneficiaries [+/=], it is quite likely that the intervention had social impacts, either positive [+/+] or negative [+/-]. However, these more holistic results are rarely recorded because social impacts are not usually assessed.

Consider a cash transfer programme that targets the poorest members of a community. A standard evaluation would report on how the cash was used and what changes followed in key indicators of interest, such as the food security and nutrition status of beneficiaries. Effects on social dynamics within the community would be overlooked, but could go either way. A 'double success' [+/+] would occur if the community valued the delivery of cash transfers to its poorest members, because this alleviates the burden of care from their neighbours who become 'indirect beneficiaries' of the programme. But a 'mixed result' [+/-] is also possible, if neighbours resent being excluded from the programme, believing they are equally deserving of support. The social tension that this creates counterbalances the positive economic gains to 'direct beneficiaries'.

There is also a synergy to be evaluated here. If formal cash transfers substitute for informal support previously provided by the community (similar to the 'substitution effect' discussed in school feeding schemes

above), the direct beneficiaries could gain very little in material terms, and neighbours who gain might feel positively about the programme [=/+] while others who are poor but excluded might feel resentful [=/-].

This example highlights the complexity of unintended material impacts and social dynamics, which are never uniform but could diverge across the beneficiary group. Some beneficiaries could be significantly better off thanks to the intervention, while others might be left even worse off than before if their informal support systems are severed and they lose more than they gain. Drawing on one of the examples discussed above, while a child grant may improve household income and children's nutritional outcomes it may at the same time undermine a young mother's position if she lives in a multi-generational household with an older woman. Similarly, some social relationships might be enhanced while others are compromised by this intervention, as illustrated by intra-household dynamics between biological and non-biological children. Analysing these kinds of impacts requires less conventional methods, such as social network analysis (Camfield and Duvendack 2014).

Conclusion

Standard impact evaluations that deploy (quasi-) experimental design protocols are locked into testing a theory of change that might be inappropriate, and are at risk of overlooking unintended impacts that could compromise the wellbeing of beneficiaries in unforeseen ways and could also undermine the effectiveness of the intervention. Conversely, qualitative methods alone are unable to assess the scale or significance of material impacts, nor can they attribute any observed changes to a specific intervention. It is therefore axiomatic that mixed methods approaches are imperative to investigate the roles of programme processes and social dynamics in shaping outcomes and impacts.

Several methodological implications follow from the analysis and arguments made in this chapter. Most fundamentally, these reflections question the hegemony of randomised control trials as the 'gold standard' in evaluations of development interventions, and propose a more open-ended, holistic and exploratory approach to research design, data collection, data analysis and reporting.

Impact evaluations need to be more aware of possible divergences from the intervention's predicted trajectory of change. Specifically, the way the intervention is designed and delivered can in itself affect outcomes and impacts, either positively or negatively. One such pathway is

through unintended and unpredictable impacts on ‘social dynamics’ – the inter-personal relationships that are affected by the intervention. Understanding these potential social consequences could be enhanced by undertaking a pre-intervention ‘sociocultural analysis’ that would include a ‘gender audit’, especially if gender empowerment is one objective of the intervention, for example, by transferring cash or microcredit directly to women.

In addition to standard qualitative tools such as focus group discussions, key informant interviews and in-depth case studies, impact evaluations should draw on fieldwork methods from disciplines including social anthropology and rural sociology – such as participant observation and social network analysis. These methods remain under-utilised in impact evaluations, but they might be essential to capture complexity to better understand why programmes succeed or fail and to report more comprehensively on their outcomes and impacts.

Finally, if a development intervention fails to achieve its intended impacts, the theory of change should be interrogated. Often reality is more complex than is implicit in a simple linear ‘cause-effect’ hypothesis, and more nuanced research methods are needed not only to understand this complexity but also to adjust expectations of what a single development intervention can realistically achieve.

Notes

- 1 See Adato (2008) and Ellis (2012) for wider discussions about social tensions resulting from targeting procedures.
- 2 ‘Interface analysis problematises the notion that development interventions are implemented according to linear blueprints which culminate in projected outcomes, highlighting the agency of actors that transform, undermine and subvert policy and give rise to the host of unacknowledged and unplanned outcomes’ (Solomon 2003: i).

References

- Adato, M. (2008). Combining Survey and Ethnographic Methods to Improve Evaluation of Conditional Cash Transfer programmes. *International Journal of Multiple Research Approaches*, 2: 222–236.
- Adelman, S., Gilligan, D. and Lehrer, K. (2008). How Effective are Food for Education Programs? A Critical Assessment of the Evidence from Developing Countries. *Food Policy Review 9*. Washington, DC: International Food Policy Research Institute (IFPRI).
- Befani, B., Barnett, C. and Stern, E. (2014). Introduction: Rethinking Impact Evaluation for Development. *IDS Bulletin*, 45(6): 1–5. doi: 10.1111/1759-5436.12108.

- Berhane, G., Devereux, S., Hoddinott, J., Tegebu, F. N., Roelen, K. and Schwab, B. (2012). *Evaluation of the Social Cash Transfers Pilot Programme, Tigray Region, Ethiopia: Baseline Report*. Washington, DC: International Food Policy Research Institute (IFPRI); Brighton: Institute of Development Studies (IDS); Mekelle: Mekelle University.
- Bundy, D., Burbano, C., Grosh, M., Gelli, A., Jukes, M. and Drake, L. (2009). *Rethinking School Feeding: Social Safety Nets, Child Development, and the Education Sector*. Washington, DC: World Bank.
- Camfield, L. and Duvendack, M. (2014). Impact Evaluation: Are We 'Off the Gold Standard'? *European Journal of Development Research*, 26: 1–11.
- Davies, S. and Davey, J. (2008). A Regional Multiplier Approach to Estimating the Impact of Cash Transfers on the Market: The Case of Rural Malawi. *Development Policy Review*, 26(1): 91–111.
- Department of Social Development (DSD), South African Social Security Agency (SASSA) and UNICEF. (2011). *Child Support Grant Evaluation 2010: Qualitative Research Report*. Pretoria: UNICEF South Africa.
- Devereux, S. and Jere, P. (2008). 'Choice, Dignity and Empowerment'? *Cash and Food Transfers in Swaziland: An evaluation of Save the Children's Emergency Drought Response, 2007/08*. Mbabane: Save the Children Swaziland.
- Ellis, F. (2012). 'We Are All Poor Here': Economic Difference, Social Divisiveness and Targeting Cash Transfers in Sub-Saharan Africa. *Journal of Development Studies*, 48(2): 201–214.
- Goetz, A. and Sen Gupta, R. (1996). Who Takes the Credit? Gender, Power, and Control Over Loan Use in Rural Credit Programs in Bangladesh. *World Development*, 24(1): 45–63.
- Harrison, G. (2014). Impact Evaluation and Welfare Evaluation. *European Journal of Development Research*, 26(1): 39–45.
- Hoddinott, J., Gilligan, D. and Taffesse, A. (2011). The Impact of Ethiopia's Productive Safety Net Program on Schooling and Child Labour. Chapter 5. In Handa, S., Devereux, S. and Webb, D. (eds.), *Social Protection for Africa's Children*. London: Routledge, 71–95.
- Kagin, J., Taylor, J., Alfani, F. and Davis, B. (2014). Local Economy-wide Impact Evaluation (LEWIE) of Ethiopia's Social Cash Transfer Pilot Programme. *From Protection to Production Report*. Rome: FAO, <http://www.fao.org/3/a-i3815e.pdf>
- Kanbur, R. and Shaffer, P. (2007). Epistemology, Normative Theory and Poverty Analysis: Implications for Q-Squared in Practice. *World Development*, 35(2): 183–196. doi: <http://dx.doi.org/10.1016/j.worlddev.2005.10.016>
- Khagram, S. and Thomas, C. (2010). Toward a Platinum Standard for Evidence-Based Assessment by 2020. *Public Administration Review*, December: S100–S106.
- Kristjansson, E., Robinson, V., Petticrew, M., MacDonald, B., Krasevec, J., Janzen, L., Greenhalgh, T., Wells, G., MacGowan, J., Farmer, A., Shea, B., Mayhew, A. and Tugwell, P. (2006). School Feeding for Improving the Physical and Psychosocial Health of Disadvantaged Students. *Campbell Systematic Reviews*, 2006: 14. Campbell Collaboration. doi: 10.4073/csr.2006.14.
- Long, N. (2001). *Development Sociology: Actor Perspectives*. London: Routledge.
- MacAuslan, I. and Riemenschneider, N. (2011). Richer But Resented: What Do Cash Transfers Do to Social Relations?. *IDS Bulletin*, 42(6): 60–66.
- NPC. (2012). *Assessment of Karnali Employment Program*. Kathmandu: National Planning Commission (NPC).

- Patel, L., Hochfeld, T. and Moodley, J. (2013). Gender and Child Sensitive Social Protection in South Africa. *Development Southern Africa*, 30(1): 69–83. doi: 10.1080/0376835X.2012.755872
- Roelen, K. and Devereux, S. (2014). Evaluating Outside the Box: Mixing Methods in Analysing Social Protection Programmes. *CDI Practice Paper 06*. Brighton: Institute of Development Studies.
- Roelen, K. and Karki Chettri, H. (2014). *Researching the Linkages between Social Protection and Children's Care in Ghana*. London: Family for Every Child.
- Roelen, K. and Karki Chettri, H. (2015). *Child Sensitive Social Protection in Nepal*. Kathmandu: Save the Children Nepal (forthcoming).
- Roelen, K. and Shelmerdine, H. (2014). *Researching the Linkages between Social Protection and Children's Care in Rwanda*. London: Family for Every Child.
- Roelen, K., Edstrom, J., Sabates-Wheeler, R. and Davies, M. (2011). *Lessons from the Children and AIDS Regional Initiative (CARI): Child- and HIV-sensitive Social Protection in Eastern and Southern Africa*. Brighton: Centre for Social Protection, Institute of Development Studies.
- Shaffer, P. (2013). Ten Years of 'Q-Squared': Implications for Understanding and Explaining Poverty. *World Development*, 45: 269–285. doi: <http://dx.doi.org/10.1016/j.worlddev.2012.12.008>
- Slater, R. and Mphale, M. (2008). *Cash Transfers, Gender and Generational Relations: Evidence from a Pilot Project in Lesotho*. London: Humanitarian Policy Group, Overseas Development Institute. <http://www.odi.org.uk/resources/download/1922.pdf> (last accessed 12 January 2015)
- Solomon, C. (2003). Giving Women Choices? Development Interfaces: Women and Microcredit in Northern Ghana. *DPhil Thesis*. Brighton: Institute of Development Studies, University of Sussex.
- Ulrichs, M. and Roelen, K. (2012). Equal Opportunities for All? A Critical Analysis of Mexico's *Oportunidades*. *IDS Working Paper*, 413 and *CSP Working Paper*, 007. Brighton: Institute of Development Studies.
- USAID. (1995). *Guidelines for Social Soundness Analysis*. Washington, DC: United States Agency for International Development.
- Whaley, S., Sigman, M., Neumann, C., Bwibo, N., Guthrie, D., Weiss, R., Alber, S. and Murphy, S. (2003). The Impact of Dietary Intervention on the Cognitive Development of Kenyan School Children. *American Society for Nutritional Sciences*, 133: 3965S–3971S.
- World Food Programme (WFP). (2009). *Home-Grown School Feeding: A Framework for Action*. Rome: WFP.

Section III
From Research to Policy

8

An Inclusive Proposal for the Use of Mixed Methods in Studying Poverty: An Application to a Colombian Municipality

María Fernanda Torres Penagos and Edna Bautista Hernández

Introduction

Historically, poverty studies have focused on the quantitative analysis of information as a key for decision-making for policy makers. However, recent studies have shown that a mixed approach captures the richness of the information, and also allows us to understand the extension, the pattern and the nature of poverty (Bourguignon, 2001; Carvalho, 1997). This chapter aims to demonstrate the validity of the use of mixed methods in studying poverty and its additional value for informing policy processes. Through the integration of information, methodologies and experience in quantitative and qualitative approaches, we designed a mixed methods diagnosis of poverty in a municipality in Colombia. This study used mixed methods in order to update the information on, and explain and understand the phenomenon of, poverty holistically for this small municipality.

The technical advantages of using mixed methods are plentiful. Each approach has clear objectives that differ from each other and complement rather than compete: 'the limitations of one method can be offset by the strengths of the other method, and the combination of quantitative and qualitative data provide a more complete understanding of the research problem than either approach by itself' (Creswell and Plano, 2011, p. 45). The essential purpose of mixing methods in this context is to get the best measurement of multidimensional poverty for the study population using the strengths of each of the methods proposed. In the research presented here, the use of mixed methods was important for three main reasons: (i) allowing for a robust analysis following triangulation of data, sequencing of data collection and research methods

and methodological bilingualism, (ii) identification of priority areas for policy and (iii) promoting policy uptake at the local level.

For the study of poverty used in this chapter, mixed methods were particularly effective in gathering information that would not have been collected by either quantitative or qualitative methodologies in isolation. The integration of the methodologies was essential to develop a complete diagnosis of the situation of poverty in a municipality that did not have enough information about the living conditions of its inhabitants since it did not have up-to-date diagnoses of the poverty condition, neither quantitatively nor qualitatively. Mixed methods research allowed for identifying individual dimensions of poverty in the particular study site, as well as to obtain results generalisable to the entire population of the municipality while deepening the understanding of causes and consequences of poverty experienced by some of the inhabitants. The research also allowed for identifying the institutional arrangements of the municipality, the relationship between citizens and local authorities, and programme rationales about fighting poverty and the needs of the town's inhabitants to improve their quality of life.

This investigation contributes to the literature in public policy to overcome poverty in the local contexts in three aspects: (i) the application of a mixed methodology for making a diagnosis, and following from this (ii) the identification of priority areas for state interventions and (iii) the use of study findings for policy uptake at the local level.

This chapter proceeds as follows: First, we discuss the use of mixed methods research, its main objectives and its importance in the study of poverty. Second, we briefly present the research and its purpose, the methodology for undertaking a poverty diagnosis and main results. Third, we present a discussion of the contributions of our work to poverty studies using mixed methods research, as well as the advantages of using mixed methods for measuring and analysing poverty. Finally, we provide recommendations for policy and/or future research in mixed studies of poverty.

From individual approaches to mixed methods

Social science research integrates different tools to approach the study of poverty. However, at a methodological level 'much analysis of chronic poverty to date has focused on monetary measures of living standards, partly because these are among the measures that can fluctuate most over even quite short time periods (McKay and Lawson, 2003, p. 426).' The main feature of the quantitative methodology is quantifiable measurement of the observed phenomena. Statistics is an appropriate tool to

measure 'objective' and 'regular' phenomena, so as to study their variability and degree of generalisation (Castro, 1996). Thus, the quantitative methodology allows us to generalise the results to the study population and allows us to answer questions such as *How much? What? What is the relation?* For the specific case of a multidimensional poverty measurement that is representative of the population, quantitative methodology is critical to identify which parts of the population are living in poverty, what are the most urgent deprivations to address and how deep such deprivations are. In this regard, the objectives of generalisation and 'objective and fair' measurement is fulfilled.

Although the use of quantitative methods has been essential to inform the process of political decisions related to overcoming poverty, it is clear that the use of this research method in isolation presents some limitations. Specifically, the exclusive use of quantitative methods does not reveal enough information to understand the factors and processes operating within situations of poverty (Howe and McKay, 2007).

Given that poverty is a phenomenon with a high level of complexity and a multitude of factors at play, it is important to also approach it from a qualitative perspective. People experiencing poverty have the greatest knowledge of the hardships that affect their wellbeing. The qualitative methodology is intended to answer questions such as *How? Why?* And it seeks to understand the meaning of definitions, symbols and descriptions of particular cases (Neuman, 2010).

Although the use of qualitative methods strengthens the participation of communities and identifies their perceptions of the determinant factors to explore related to poverty, the main difficulty lies in the particularity of the results from participatory approaches, as these are not usually generalisable to the population.

Mixed methods research has become a fundamental tool to investigate complex issues such as poverty since 'they are useful when you have a need to explore and explain both, and when either qualitative or quantitative methods alone seem inadequate for the complexity of the research questions and topics' (Bronstein and Kovacs, 2013, p. 355). Creswell and Plano Clark (2011) suggest that mixed methods research went through a formative stage in the 1970s, a discussion of paradigms and procedures throughout the 1980–1990s, advocacy and expansion in the next period, and is currently in a reflective period (Bronstein and Kovacs, 2013). Mixed methods research has been called the 'third methodological movement', the 'third research paradigm', and 'a new star in the social science sky' (Creswell and Plano Clark, 2011, p. 52). It emerges as an alternative to develop research to understand the complexity of

social phenomena holistically. Morgan (2007, p. 73) aptly summarises the previous discussion:

I find it helpful to think of qualitative research as research that emphasizes a contextual-opinion-inductive approach, while quantitative research emphasizes a deductive-objective-generalising approach. Where we find problems is in the treatment of these major trends as absolute and defining characteristics for these two different approaches [. . .] the pragmatic approach [mixed methods] offers an effective alternative through its emphasis on abductive-intersubjective aspects and transferable of our research.

As such, the importance of combining methods has been widely covered. First, mixed methods research is desirable considering that both methods share the goal of understanding social phenomena in the best way, the common objective of disseminating knowledge for practical use as well as a shared commitment to rigour, awareness and critique in the process of research (Haase and Myers, 1988; Reichardt and Rallis, 1994). Second, a mixed methods approach allows for responding to a greater number of questions, permits a greater variety of perspectives on the problem analysed in terms of frequency, size and prevalence (quantitative), and depth and complexity (qualitative) (Creswell, 2005). Finally, as stated by Reams and Twale, the use of mixed methods emerges as necessary to 'uncover the most of information and perspective, increase corroboration of the data and render less biased and more accurate conclusions' (Reams and Twale, 2008, p. 133).

Although mixed methods have been used within development studies for more than two decades, no consensus currently exists in relation to a single definition on what these are. According to our experience, we believe that using mixed methods is not only the use of two different methodologies of data collection, but instead involves a complete integration of collection, interpretation and analysis of information for the study of the same phenomenon. In this sense, the definitions given by Sampieri et al. (2006), Tashakkori and Creswell (2007) and Bronstein and Kovacs (2013) mirror our understanding of mixed methods research. Sampieri et al. (2006, p. 546) suggest that mixed methods research could be defined as 'a set of systematic processes, empirical and critical research and involve the collection and analysis of quantitative and qualitative data, as well as their integration and joint discussion, to make inferences product of all the information collected and greater understanding of the phenomenon under study'.

Tashakkori and Creswell (2007, p. 4) define mixed methods as ‘research where the researcher collects and analyzes data, integrates results and draws conclusions using both quantitative and qualitative approximation in a single study [. . .] a key aspect of this definition is the integration’. Finally, according to Bronstein and Kovacs, research using mixed methods has been referred to as an intuitive way of doing research, pragmatic and practical, that draws on the strengths of both qualitative and quantitative methods to fully answer the research questions (Bronstein and Kovacs, 2013).

It follows that the use of methods in this study supports a holistic understanding of the mixed methods approach, where the integration of the methodologies starts from the formulation of the problem and enables a deeper understanding of a phenomenon as complex as poverty.

Diagnosis of poverty with mixed methods: The case of Colombian municipality

In Colombia, 20 million people live on less than two dollars a day (DANE, 2012). The latest results of poverty and extreme poverty incidence reveal stark differences between regions and geographical areas in the country. In 2010, the incidence of multidimensional poverty in rural areas was 53%, which is more than twice the percentage of poor people in urban areas (24%). Also the differences between regions are alarming; while Bogotá registered multidimensional poverty levels of 12%, the Atlantic region presented levels close to 46% (Angulo et al., 2011). Trying to understand these gaps within the country was one of the motivations for developing a diagnosis in a small municipality near Bogotá.

This research explores poverty as a multidimensional phenomenon in a small municipality composed of urban and rural areas in Colombia that did not have updated sources of information for measuring living conditions of the population at the municipal level (most recent local poverty measurements were based on the 2005 Colombian Census). To do so, we use a mixed methods design and the most up-to-date database (Survey for the Identification System and Classification of the Potential Beneficiaries for Social Programs [SISBEN] 2010) to explore poverty by combining both quantitative and qualitative research techniques.

The municipality in focus is Villapinzón, a community located in Cundinamarca (Colombia), which is 95 km away from Bogotá (the capital), the choice of which is closely linked to the programme of the School of Government Alberto Lleras Camargo, University of Los Andes, called Local Action Initiative (the acronym in Spanish is IDEAL).

The objective of this programme is to support both the administration and the citizens of small municipalities in developing local management capacity to overcome poverty, social policy and local planning, among others. The municipality is an interesting case study because it is part of the five hundred poorest municipalities in Colombia and, according to the National Planning Department, it had a high incidence of multidimensional poverty. Additionally, the size of the municipality in this case facilitated the research as well as future poverty monitoring efforts, as it is easier to survey and update information of relatively small populations. Furthermore, in this type of municipality (small population, low income, rural economy and approachable authorities) it is easier to establish a close relationship with the authorities, and know the results of the performance of mayors and governors, and also for citizens to learn about the implementation of public programmes and to exercise the necessary political control through the different mechanisms of participation.

This research was inspired by two shortcomings in poverty measurement in Colombia. Firstly, most government efforts are aimed at the development of public policies to overcome poverty at the national level. Secondly, there is a lack of updated measurements of the living conditions at the municipal level since local poverty measurements were only available from the 2005 Colombian Census. In this research, the SISBEN 2010 database was used. This data allowed for an elaborate poverty diagnosis and to operationalise the dimensions of the Multidimensional Poverty Index Colombia (MPI-Colombia), thereby supporting the identification of potential intervention areas that could contribute to the social policy agenda of the local authorities in this Colombian municipality. This approach is replicable and is therefore a starting point for the measurement of poverty in multiple parts of Colombia and other countries with similar conditions.

Methodology

The methodology of this research divides into five phases: (i) preliminary visits, (ii) participant observation, (iii) calculation of the MPI, (iv) focus group discussions and new poverty calculations and (v) semi-structured interviews. The research team consisted of two researchers with particular quantitative and qualitative skills that were drawn upon in isolation and combination throughout the research. The methodological approach allowed for the identification of new dimensions and risk factors that would not have emerged if multidimensional poverty phenomena had been analysed with qualitative or quantitative approaches separately.

Preliminary visits (Phase I) were aimed at familiarising ourselves with the leaders of the municipality and the general population. Seven families were prioritised by the *Red Unidos* (one of the main strategies to fight poverty and extreme poverty in Colombia¹). From these visits, one family was selected to participate in Phase II of the research. This family was composed of seven people that lived in the rural area of the municipality and agreed to participate in participant observation (Phase II). It is worth mentioning that the family did not allow the researcher to stay overnight, due to the distribution of space and overcrowding. This Phase I was important to build confidence through recognition within the context of research and breaking down power relations between researcher and researched (Crang, 2007; Ritterbusch, 2011).

Phase II aimed to identify, observe and collect qualitative information from a family living in poverty. The dimensions of the MPI were taken into account as well as new dimensions that emerged during the fieldwork. To this end the technique of Participant Observation was used; the expert in qualitative research methodology was embedded in the daily routines of this family living in poverty for a week.

Phase III of the research consisted of calculating the MPI as developed by Alkire and Foster (2007) and was led by the expert in quantitative research methodology. The MPI for the municipality is a replica of the national MPI conducted by the National Planning Department (NPD), recognising that it is the official measurement of multidimensional poverty used in the country and is part of the official indicators and goals of the National Development Plan. The MPI-Colombia measures five dimensions: household education conditions, childhood and youth conditions, labour, health and access to household utilities and living conditions.² Following Angulo et al. (2011), each dimension has a weight of 0.2 (for a total of 1.0), and each variable has the same weight in its respective dimension (see Table 8.1).

The MPI-Colombia uses two databases: the survey of Quality of Life in Colombia for 2008 (with updated information but representativeness for a larger geographic area) and the National Census for 2005 (with representative information at the municipal level but not updated). The main source of information used for this study was the SISBEN 2010 database. The SISBEN is the information system that identifies potentially beneficiaries of government subsidies. The database for the municipality selected has information from 16,173 people for 2010. The survey has 90 questions which are divided into seven modules: (i) identification, (ii) details of the dwelling unit, (iii) household data, (iv) socio-demographic background, (v) health and fertility, (vi) education and

Table 8.1 Construction of indicators to measure the Multidimensional Poverty Index

Dimension	Indicator	Technical description	Cutoff
1. Household education conditions	Low educational achievement	A household is deprived when the average of the educational achievement of household members aged 15 years or more is less than 9 school years.	9 years
	Illiteracy	A household is deprived when at least 1 person aged 15 years or more does not know how to read or write.	100%
2. Children and youth conditions	Child labour ^a	The index shows the percentage of children between 12 to 17 years in the urban area and 10 to 17 in the rural area that are out of the labour market. A household is deprived when this index is below 100%.	100%
	Non-attendance	A household is deprived if less than 100% of the children between 6 and 16 years old go to school.	100%
	Educational lag	Percentage of children between 7 and 17 years who do not have educational lag (number of approved years below the national norm). A household is deprived if any of the children between 7 and 17 has educational lag.	100%
3. Health	Without early childhood services	Percentage of children between 0 and 5 years with no access to childhood care (health, nutrition, education). Households with at least 1 child between those ages without childhood services are deprived.	100%
	Healthcare	A household is deprived if any of its members have no health insurance.	100%

4. Labour	Long-term unemployment Informal employment	A household is deprived when at least 1 person is unemployed for more than 12 months. A household is deprived when less than 100% of the Economically Active Population has formal employment.	100% 100%
5. Public services and housing conditions	Access to water service No sanitary service Inadequate floors Inadequate exterior walls Severe overcrowding	Rural area: a household is deprived if it gets water to prepare food from a well with no pump, rain water, river or other source. Urban area: a household is deprived if it has no public water service. Urban: a household is deprived if there is no connection to the sewerage system. Rural: a household is deprived if it has a toilet with no connection, latrine or do not have sanitary service. A household is deprived if it has dirt floors. Urban area: a household is deprived if the materials of exterior walls are made of wood, zinc, material, cardboard and so on. Rural area: a household is deprived if the material of the exterior walls is made of <i>bambú</i> , zinc or there are no walls. Urban area: a household is deprived when the number of people in each room is equal to 3 people. Rural area: a household is deprived when the number of people per room is 3 or more persons.	Urban: no water connection Rural: well water, rainwater, river, water truck Urban: there is no connection to the sewerage system Rural: no sanitary service Ground floors Rural and urban areas: inadequate wall materials Urban: 3 people per room Rural: 3 or more people per room

^a Child labour: children from 5 to 17 years who work at home for more than 15 hours a week, children until 14 years who work and children from 15 to 17 who do hard work (more than 36 hours per week and according to the Convention 182 of the ILO activities as worst forms of child labour).

Source: Adapted from Angulo et al. (2011)

(vii) occupation – income. In the SISBEN special emphasis is placed on questions related to housing conditions (health services, drinking water, number of rooms) and possession of goods (fridge, TV, computer, motorcycle), and less emphasis on questions related to health and education.

From the 15 variables used by the NPD for the calculation of the national MPI, it was possible to rebuild 14 for the municipality under study using the SISBEN database. The only variable that had no information in the SISBEN database and was not possible to estimate for this municipality was ‘Healthcare access when needed’, which measures the percentage of households with all members having access to healthcare.

The Multidimensional Headcount Ratio and the Adjusted Headcount were calculated, identifying the number of poor people in the municipality. The municipality has large differences in its geographic area (66% of the population lives in the rural area vs 34% in the urban area); for this reason calculations were performed differentiating by area of residence (urban/rural). Calculations were also performed by household composition: with children and without children. The Adjusted Poverty Gap (which adds in information about the depth of poverty) was also calculated, and a multi-variate analysis was done with the aim of studying the relationship between multidimensional poverty as a function of a group of potential risk factors or circumstances. Finally, the MPI was calculated for a second time but only for the population identified as multidimensionally poor in the municipality (people with 33% or more of deprivations in the Index). The focus of these calculations was to understand the contribution of individual dimensions to the index for MPI scores for this subgroup of the population.

Phase IV was developed by both researchers in the team and consisted of conducting focus group discussions (FGDs) with a sample of the urban population of the municipality. This Phase focused on the urban population considering that participant observation (in Phase II) was conducted in the rural area. The ‘snowball’ technique was used, making it possible to identify members with similar life experiences (Browne, 2003; Salganik and Heckathorn, 2004) and with personal information relevant for this research. Three focus groups were conducted, each with three to six participants. This sample was stratified by age, gender and education level but not socio-economic status of the participants.

The purpose of this activity was to generate a dynamic of participation and conversation with the community, in order to motivate the group to meet and discuss poverty in general through programmes, interventions related to overcoming poverty and the identification of alternative solutions. Participants answered questions such as: *What is poverty? How is poverty experienced? What is the most important factor*

in overcoming poverty? What can the government do to solve the problems of the community? The information was collected using a diary, focus group guides, videos, photos and audio. Along with the analysis of MPI, the focus group discussions were useful for ranking the most important dimensions for participants in order to contribute to the construction of the participatory poverty diagnosis. This phase allowed prioritisation of variables and dimensions of the MPI by the urban population, thus facilitating the identification and collection of qualitative information about the perception of poverty.

In this same phase, and following the approach made by Angulo et al. (2011) for Colombian MPI, the researchers estimated the incidence of poverty (Headcount Ratio) only for the poor population (people with 33% or more deprivations compared to the non-poor population) in order to prioritise the most important variables for the poor and compare the results with those found in the focus group discussions.

Finally, the two researchers conducted semi-structured interviews (Phase V) with the main decision-makers of the municipality in order to create awareness about and deepen the findings of the preceding phases. One of the main objectives of the interviews was to observe the viability of the public policy proposals coming out of the poverty research. As such, this phase allowed for deepening the dimensions of poverty that were identified as key through observing the responsiveness of local authorities to them, as well as programmes and policies designed and executed for overcoming poverty.

Key findings

Key findings following the combination of methods are discussed on a dimension by dimension basis.

Household education conditions

This domain pertains to past educational attainment of household members rather than current educational outcomes for children and youth.

Following findings from the qualitative components in Phases II and IV it was found that family support is a key aspect to achieving educational attainment. In the rural family, for example, we found that the absence of support from grandparents, parents and caregivers was a key factor for dropping out of school. Similarly, a young female participant (Phase IV) said the financial support from their parents was the reason for having finished her studies in fifth grade:

[When I got pregnant at age 15] I was not studying because my mom had just money to support me until fifth grade and I wanted to

continue studying but my family had no money to continue studying. (Woman, Focus Group, 21 years)

In general, many people in the municipality (both urban and rural areas) had low levels of educational attainment. Nearly 80% of surveyed household members over 15 on average received less than nine years of schooling. Large differences by area are identified in the results; the percentage of households with low educational attainment in urban areas is 58.1%, while for rural areas it is 94.0%.

Child and youth conditions

With respect to outcomes for children and young people, fieldwork revealed that several children in the municipality both work and study. This group, according to the discussion in Phase IV, is more likely to drop out of school. The most serious problem for children and youth in this municipality appeared to be child labour.

Other obstacles for children in rural areas in going to school or accessing early childhood services were found to be geographic isolation and beliefs about early education (Phases II and V). For example, children of rural families have to walk almost an hour to get to school every day. Regarding the access to early childhood services, it was found that none of the children of the family interviewed in Phase II had access to early education because the parents did not trust the quality of this service and children were taken care of by their mother or grandmother.

The variables that attract most attention in terms of their incidence are educational lag and lack of access to early childhood services. The first has an incidence rate of 57% and the second has an incidence rate of 31%. Differences between urban and rural areas are small – educational lag (58% urban and 56% rural) and without access to early childhood services (28% urban and 32% rural) – indicating that these variables must be taken into account regardless of the particular characteristics of the area that is inhabited. It is important to note that incidence rates for child labour were low, which is due to the way in which the question is asked in the SISBEN database. In the qualitative fieldwork phases, it emerged in all cases.

Health

Health issues in the municipality appear to be more of an access problem than a coverage problem, especially for rural areas. Observations from Phases II and IV indicate that access to healthcare is different by area.

In rural areas, Phase II showed that distance implies a barrier to accessing services. In terms of urban areas, participants felt that doctors do not pay enough attention to their cases, so it is more efficient to purchase medicines and have them at home.

Findings from the quantitative analysis in Phase III suggest that more than 20 out of 100 people have no health insurance. The percentage of people without health insurance is similar in rural and urban areas.

Labour

In the municipality it is difficult to get formal employment; adults in Phase IV indicated that there is not sufficient labour supply for all. Although both men and women experience difficulties getting work, women are more likely to be unemployed:

[How many cows milked today?], Five cows, [how much you earn for milking?], \$ 1.50, and my boss gives me a liter of milk. [But your husband earns more money?], Of course, they pay him 12 dollars per day. (Mother of the family, Participant Observation, 29 years)

Over 77% of households in the municipality have informal work, reflecting the labour instability that characterises large parts of the population in the municipality. The fieldwork for the qualitative component in Phase IV clearly showed that the dynamics of work (especially in rural areas) is based on the system of daily wages, where workers are not entitled to social security and do not have an employment contract. Results vary greatly depending on the area; the rate of informal employment is 59.5% in urban areas and 89.2% in rural areas.

Public services and housing conditions

For people in this municipality, the most pressing issue with respect to living conditions is overcrowding. In the case of the rural family (Phase II), it was found that the household has no access to sewage disposal and improved water. The focus group discussion with adults (Phase IV) suggested that government social programmes for housing are not designed according to the living conditions of the population:

It seems to me irrational, you must have a lot to build, or to have a housing allowance [. . .] the state should help us to get the land. (Women, Focus Group, 47 years)

Critical overcrowding (three people or more sleeping in the same room for the urban area and four or more people in the rural area) is very high in the municipality, revealing that 27 of every 100 households suffer from such deprivation. Sewage disposal (18%) and access to improved water (14%) are also variables that must be taken into account. When performing the analysis by area, greater deprivation can be observed in the rural area of the municipality (sewage disposal 26% and access to improved water 23%).

New dimensions

Finally, our research allowed for the identification of new dimensions of poverty. In the design of the focus group discussions (Phase IV), new dimensions that were observed in the participant observation (Phase II) and that were not reflected in the quantitative results were taken into account. We identified that *geographic isolation*, *domestic violence* and *teenage pregnancy* affected the household welfare.

The fieldwork in Phases II and IV revealed that geographic isolation is a major barrier to accessing services. Children of the family participating in Phase II walked 50 minutes to get to school, crossing a river in boots. Unfortunately, there is no information to quantify this variable beyond this single observation.

With respect to domestic violence, it was found that alcoholism is a factor that aggravates abuse, physical abuse and domestic violence. All participants in the qualitative components (Phases II and IV) referred to acts of violence in their families. Violence against women was mentioned in all the focus group discussions; most participants had been abused by their spouses.

In terms of teenage pregnancy, it was found that all women over 18 years who participated in the qualitative components (Phases II and IV) were mothers of at least two children. The average age of the mothers was 17 years. However, the quantitative SISBEN data only identified that 1.14% of women under 20 reported having been pregnant in the municipality. This value may be underestimating teenage pregnancy because the information specified in the database does not take into account women over 20 years who had children when they were teenagers.

Contributions to the methodological debate

Several contributions to the literature on mixed methods emerge from the research conducted for this small municipality in Colombia. We elaborate on the general advantages of the use of mixed methods in

our research. We analyse why it is advisable to develop mixed methods research with multi-disciplinary teams and researchers specialised in each of the methodologies. Also, we highlight the importance of paying attention to the timing and sequencing in the development of the fieldwork and the importance of integrating methods and data for a deeper understanding of the phenomenon.

This section proceeds as follows: we discuss (i) the advantages of using a mixed methodology for making a diagnosis, (ii) the identification of priority areas for state interventions and (iii) reflection on our study and policy uptake at the local level in relation to interviews with authorities.

Advantages of using mixed methods for measuring and analysing poverty

The use of mixed methods strengthened the results of this research for three main reasons: (i) *triangulation* of results, (ii) *timing* of integration of methodologies and (iii) a robust *methodological bilingualism*. These stages of analysis in the mixed methodology incorporate collection, analysis and integration of both techniques to answer one research question, so the usefulness of mixed methods results in a more comprehensive and inclusive understanding of the research problem (Bilinski et al., 2013).

Triangulation

Triangulation is a key element in the use of mixed methodologies. As said by Arias (2000) it is conventionally understood as the use of multiple methods in studying the same object in order to increase the validity of the results. However, it goes further and includes diversity of data, researchers, theories and methodologies (Arias, 2000). When an investigation has more than one type of triangulation it is called 'multiple triangulations' and it is ideal because it provides a 'more comprehensive and satisfactory sense of the phenomenon'. Thus, triangulation allows the researcher to 'take advantage of the strengths and offset the weaknesses of each [approach]' (Royce et al., 2010, p. 99).

In the case of this research, triangulation was understood as the use of different sources, triangulation of researchers and triangulation of methodologies. Triangulation of data sources was vital for complementing the information collected by using different methodologies. Researcher triangulation occurred as two researchers (with different backgrounds) explored the same phenomenon in the field. The methodological triangulation was carried out both in design and in data collection between the qualitative and quantitative methods (through different information gathering techniques to know the poverty condition of

this municipality). As a result, variables that were not evident from the quantitative data were picked up in the qualitative fieldwork as being important elements of living in poverty (teenage pregnancy and child labour). Similarly, while the quantitative methodology allowed a disaggregated analysis of the impact of different dimensions on conditions of poverty, qualitative methodology allowed deepening the enquiry into the causes of poverty that these people experienced. For example, geographic isolation was found as a determinant for the quality of life in the urban area.

The use of mixed methods facilitated the understanding of the complexity of the phenomenon of poverty in the municipality. The qualitative results allowed us to understand the causes and consequences of poverty. As said by Shaffer (2012), one of the great challenges when measuring poverty refers to the need to use categories 'with local significance' of poverty. In this study the concepts of poverty corresponded to the understanding of the people surveyed about their meaning of poverty. Meanwhile the quantitative results allowed us to speak of generalisable results to the entire population living in poverty in the municipality, essential for the designing of public policies to overcome poverty. At the local level, for example, dynamics that perpetuate the poverty status of the study population were identified such as alcoholism and domestic abuse. These dynamics that are not addressed in the national poverty measurement are part of the problems to solve in the case of this small municipality.

Timing

The timing or implementation of research methods refers to the sequence in which the information is collected, analysed and interpreted (Creswell et al., 2004). Morgan (1998) suggests that decisions about the timing need to be properly resolved when considering what is the complementary method and what is the central method. According to Bilinski et al. (2013) the methodological design of the mixed methods research requires deciding in advance the timing of the phases of the research. McWey et al. recognise how the timing of mixed methods can contribute more effectively to the research question: 'with this type of design, both quantitative and qualitative data are collected at the same time and given equal weight. Also with concurrent triangulation, the quantitative and qualitative data are analyzed separately, but are integrated when interpretations of the data are made' (McWey et al., 2009, pp. 80–81).

In our particular case, for the design and content of the various stages of the research it was essential to build the information about the

dynamics of poverty in the municipality in a parallel way. Both the complementarity of the information and the triangulation of data required a component of synchronicity in timing of our investigations. Although the sequence and development of the research seems a secondary factor for researchers, in the case of mixed methods, and particularly for our research, the strategy was to plan both *how* and *when* to integrate the data collected. It is worth specifying that both quantitative and qualitative data were collected and analysed with the same priority. While there were experts in both methodologies, none of them advocated any particular method, as these converged at the stage of interpretation, allowing a significant increase in the information available to analysis.

The sequencing of the various research phases led us to pay special attention to data that appeared to be contradictory between the quantitative and qualitative methodologies; dimensions such as child labour and teenage pregnancy showed the importance of using mixed methods and specifically timing in implementing the methodology. These dimensions emerged strongly in the qualitative work done in the early stages (Phase I – preliminary visits and Phase II – participant observation). However, when they were tested in the database they were found to affect a very small percentage of the population. The information gathered from the questionnaires of the survey did not capture the existence of child labour (it asked if people work or study without acknowledging that the two activities may be performed simultaneously) nor the magnitude of the phenomenon of teenage pregnancy (considering that the question ‘currently pregnant or had children’ does not identify adult women who got pregnant in their teens). Subsequently, the research team decided to deepen knowledge on these two variables in the later stages of fieldwork (Phase IV – focus group discussions and Phase V – interviews). Findings highlighted that child labour and teenage pregnancy are indeed important issues as perceived by research participants. As such, the policy recommendations made to decision-makers of the municipality included these two variables.

Two methodologies: One researcher? Methodological bilingualism

One of the major concerns of the sceptics in relation to mixed methods is related to the lack of tools and skills on the part of researchers to adequately develop the two types of methodologies. Denzin (2008, p. 11) says, for example, that the ‘methodological bilingualism is superficial, and perhaps even impossible’. Qualitative experts, as stated by Teddlie and Tashakkori, think that mixed methods researchers do not have adequate training in the ethnographic tradition or

appreciate the skills needed to conduct ethnographies (Teddle and Tashakkori, 2012). However, authors like Shulha and Wilson (2003) and Bliss (2008) state that this criticism is addressed by an approach to mixed methods through multi-disciplinary teams where not everyone is skilled in each methodology but everyone has a minimum knowledge of the other's arguments. Bliss says, 'familiarity in the two methodologies allows effective communication because they handle a common language and a viable number of conceptual similarities' (Bliss, 2008, p. 190).

In the case of this research, the team consisted of an expert in quantitative methods and an expert in qualitative methods. Each component of this research was peer reviewed by experts in quantitative and qualitative methodology respectively. Also, each researcher had a solid foundation in the other methodology before starting research, which facilitated communication and the choice of information gathering techniques; this was vital for the integration of the entire research process. Prior knowledge in the other methodology facilitated the sequence in which information was collected and improved the triangulation, allowing a fluid and proactive exchange of information.

Methodological bilingualism was also key when determining the priority areas of intervention and how the results were presented to local authorities. Both the statistical results and the voices of the participants occupied a central role in the diagnosis; this combination contributed to suggesting areas of intervention that are important for the population as a whole but also with a human face.

Identification of priority areas for state interventions

One fundamental aspect for the integration of quantitative and qualitative data in this research was the prioritisation of the dimensions of poverty in each of the methodologies. Following Tashakkori and Creswell (2007), mixed methods research goes beyond the reporting of quantitative and qualitative findings; this methodology must connect and integrate the information collected by these two approaches. The conclusions must then present a much more complete picture of the phenomenon studied. The integration must compare, contrast and build upon the conclusions from each of these methods (Tashakkori and Creswell, 2007). As such, the great contribution of this prioritisation is that it was used to generate and develop analysis in order to provide richer data. It also strengthened the significance of our research since this integration was important to generate new approaches to the study of poverty in the municipality from the combination of the two sources of data input.

Table 8.2 Prioritising mixed diagnostic variables

Priority order	Qualitative component	Quantitative component
1	Having a job	Low academic achievement
2	Firm material walls	School lag
3	Overcrowding	Without access to services for childhood
4	Health, nutrition and initial care	No insurance
5	Have bathroom and sewer	Rate of informal employment
6	Knowing how to read and write	Overcrowding

To this end, the six most important variables for each component were prioritised on the basis of both quantitative data (Phase III) and qualitative data (Phase IV) – see Table 8.2. In the quantitative component, the prioritisation was performed according to the incidence of the dimensions contained in the MPI for the population multidimensionally poor. The incidence rate of the dimensions of the MPI was calculated only for the poor (only people deprived in 33% or more of all dimensions in the municipality³). With respect to the qualitative component, a weighting exercise was developed to identify the relative importance of dimensions of MPI in terms of the quality of life of the people who participated in our research. Specifically participants answered to the question: What is most important to me?⁴

In this phase of the investigation, the gathering of information was made through each one of the methodologies. It was found that five of the six most important variables of wellbeing associated with the MPI were prioritised by both methodologies (from a total of 15 variables). At the moment of the data integration it was found that priorities which were the same in both components were: Education (low achievement, educational lag and knowing how to read and write), Services for Early Childhood, Work (informal and having a job) and Household Conditions (overcrowding, walls material and having bathroom and sewer). The only variable that was not prioritised by the two methodologies was health insurance, only present in the quantitative prioritisation.

Policy uptake at the local level

From the prioritisation of the variables of mixed diagnosis it was possible to make recommendations regarding priority areas for intervention. As explained before, the prioritised dimensions combined into: housing conditions (overcrowding), education (low educational level and illiteracy) and labour (rate of informal employment). This prioritisation allowed the government authorities to develop programmes and projects (described later) that emphasised people's resources and supported them in improving their living conditions. Clarity of information combined with detailed analysis allowed the Mayor to take efficient decisions addressing real needs and the technical and managerial capabilities of the municipality. As a result of the prioritisation of needs arising from this research, government authorities prioritised labour and housing conditions.

Two years after the diagnosis, the municipality is developing a policy advocacy framework in partnership with the University of Los Andes, a project with fungi as an alternative crop and source of nutrition, and access to employment. Besides offering advice about mushroom growing, the University supports the marketing thereof, to improve the income of people living in poverty. The authorities have also created strategies to improve housing conditions with the help of academia and the private sector. These initiatives aim to improve the use of public resources, including basic services such as drinking water. Likewise, strategies have been designed to optimise the overcrowded conditions of households in the municipality.

Considerations and conclusions

The contribution of the research presented in this chapter is threefold: Firstly, it provides a detailed and up-to-date understanding of the poverty situation at municipal level in Colombia. Secondly, it allowed for the identification of priority areas for state interventions and input into policy processes. Thirdly, it provides insight into the value added by using mixed methods in the study of poverty at subnational level and its contribution to local policy take-up. The research responds to two shortcomings noted in addressing Colombian social policy: first, currently in Colombia most government efforts are aimed at developing public policies to overcome poverty nationwide. Second, there is a lack of current sources of information to measure the living conditions of the population at municipal level.

As noted throughout this document, the use of mixed methods proved to be useful. Triangulation of results and methodological bilingualism strengthened the credibility and usability of this research, which would

not have been possible without the participation of specialists in each methodology and without time synchronisation of methods.

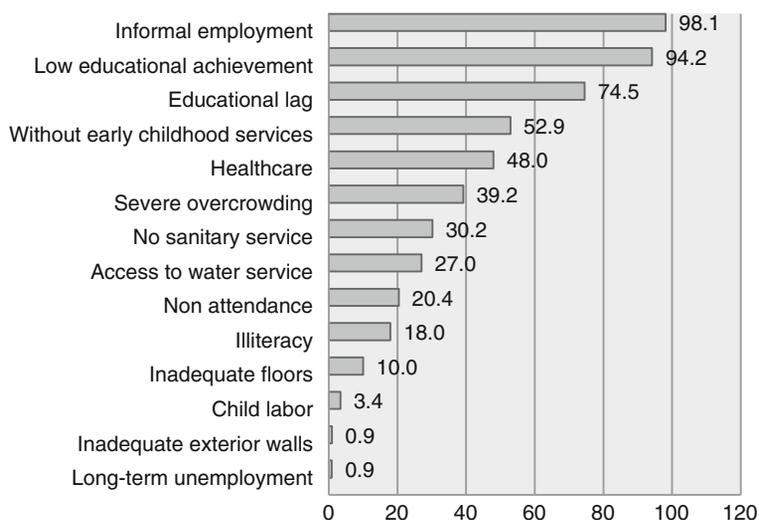
For the study of poverty in this chapter, mixed methods were particularly effective in identifying information that was not present in the use of quantitative or qualitative techniques in isolation. The methodological triangulation allowed the researchers to identify variables that were not evident from the quantitative data and were picked up on in the qualitative fieldwork as being important elements of living in poverty and for policy recommendations: the effect of geographic isolation, domestic violence and teenage pregnancy on household welfare. These dynamics are not addressed in the national poverty measurement but are part of the problems to solve in the case of this municipality. The integration of methodologies was vital to strengthen the results and to reach local authorities with a convincing diagnosis of the poverty of the municipality.

Regarding timing, it was essential to build the information about the dynamics of poverty in the municipality in a parallel way. Both the complementarity of the information and the triangulation of data required a component of synchronicity in timing.

Methodological bilingualism was key when determining the priority areas of intervention and how the results were presented to local authorities. To have the ability to deliver information to the Mayor that was representative of the entire population of the municipality, but also highlight the experiences of poverty in the everyday life of some of the town's inhabitants was key for the decision-makers to undertake public policy actions once the diagnosis was known.

The identification of priority areas for state interventions was also a fundamental exercise regarding the importance of using mixed methods. Five of six dimensions of wellbeing associated with the MPI were prioritised in the same way by both methodologies, finding significant matches in key results obtained by both methods. The impact of this study is not confined to the identification of priority areas for overcoming poverty areas but, on the contrary, was the starting point for the Mayor to give priority in its municipal development programme to a strategy for eradicating extreme poverty. Two years after the diagnosis, the municipality is developing programmes related to employment and housing conditions.

We consider that poverty studies can greatly benefit from the application of different methodologies to the same phenomenon. The subsequent results are more information-rich, are more respectful of each methodology-specific epistemology, place equal emphasis on quantitative data and narrative information and provide a broader range of policy alternatives. In sum, this provides a more satisfying sense of the complex phenomenon studied.



Appendix 8.1 Multidimensional Poverty Index for poor population

Notes

- 1 For more information visit: <http://www.colombiaenaccion.gov.co/2011/06/07/red-unidos/>
- 2 The criteria used by the government of Colombia to set these dimensions are: (i) frequent usage (national or international), literature review, discussion with experts and other indicators; (ii) indicators can be affected by public policies and (iii) availability of information (in the survey of Quality of Life in Colombia). For more information visit: <https://colaboracion.dnp.gov.co/CDT/Estudios%20Economicos/382.pdf>
- 3 Results are presented in Appendix 8.1.
- 4 Each participant had three votes, researchers presented the MPI for people to vote for what they considered the most important (could use all the votes in one variable or distribute them according to their preferences). In the end, a rapporteur chosen by them did the prioritisation according to votes and made a classification of the variables that were fundamental to the group.

References

- Alkire, S. and Foster, J. (2007). *Counting and Multidimensional Poverty Measurement*. OPHI Working Papers No. 7. Oxford: Oxford University Press, pp. 3–38.
- Angulo, R., Díaz, Y. and Pardo, R. (2011). Índice de Pobreza Multidimensional Para Colombia (IPM-Colombia) 1997–2010. Archivos de Economía.
- Arias Valencia, M. (2000). La triangulación metodológica: sus principios, alcances y limitaciones. *Redalyc – Red de Revistas Científicas de América Latina, el Caribe, España y Portugal*, 18(1), 13–26.

- Bautista Hernández, E. and Torres Penagos, M. (2013). Diagnóstico mixto para la superación de pobreza en Villapinzón, Cundinamarca: Identificación de áreas prioritarias de intervención. *Documentos EGOB No. 4*, Bogotá.
- Bilinski, H., Duggleby, W. and Rennie, D. (2013). Lessons Learned in Designing and Conducting a Mixed Methods Study to Explore the Health of Rural Children. *International Journal of Health Promotion and Education*, 51(1), 1–10.
- Bliss, L. (2008). Review of Jennifer Greene's Mixed Methods in Social Inquiry. *Journal of Mixed Methods Research*, 2(2), 190–192.
- Bourguignon, F. (2001). *Qualitative and Quantitative Approaches to Poverty Analysis: Two Pictures of the Same Mountain?* Paper presented at the Qual-Quant. Qualitative and Quantitative Poverty Appraisal: Complementarities, Tensions and the Way Forward. Cornell University.
- Bronstein, L. and Kovacs, P. (2013). Writing a Mixed Methods Report in Social Work Research. *Research on Social Work Practice*, 23(3), 354–360.
- Browne, J. and Russell, S. (2003). Recruiting in Public Places: A Strategy to Increase Diversity in Qualitative Research Samples. *Qualitative Research Journal*, 75–87.
- Carvalho, S. H. W. (1997). Combining the Quantitative and Qualitative Approaches to Poverty Measurement and Analysis: The Practice and the Potential. *Technical Paper No. 366*. World Bank, Washington DC.
- Castro, R. (1996). En busca del significado: supuestos, alcances y limitaciones del análisis cualitativo. México.: I. N. d. S. P. Centro de Investigaciones en Sistemas de Salud (Ed.).
- Crang, M. and Cook, I. (2007). *Doing Ethnographies*. Norwich: Geobooks.
- Creswell, J. (2005). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Boston: Pearson.
- Creswell, J. and Plano Clark, V. (2011). *Designing and Conducting Mixed Methods Research*. Washington, DC: Sage.
- Creswell, J., Fetters, M. and Ivankova, N. (2004). Designing a Mixed Methods Study in Primary Care. *Annals of Family Medicine*, 2, 7–12.
- DANE (17 de Mayo de 2012). *Portafolio*. URL: <http://www.portafolio.co/economia/disminuyen-pobreza-indigencia-y-desigualdad-gobierno> (accessed 12 August 2014).
- Denzin, N. (2008). The New Paradigm Dialogs and Qualitative Inquiry. *International Journal of Qualitative Studies in Education*, 21, 7–12.
- Haase, J. E. and Myers, S. T. (1988). Reconciling Paradigm Assumptions of Qualitative and Quantitative Research. *Western Journal of Nursing Research*, 10, 128–137.
- Hernández Sampieri, R., Fernández Collado, C. and Baptista Lucio, P. (2006). *Metodología de la Investigación* (Vol. 4ta Edición). México: McGraw Hill Editores.
- Howe, G. and McKay, A. (2007). Combining Quantitative and Qualitative Methods in Assessing Chronic Poverty: The Case of Rwanda. *World Development*, 35(2), 197–211.
- McKay, A. and Lawson, D. (2003). Assessing the Extent and Nature of Chronic Poverty in Low Income Countries: Issues and Evidence. *World Development*, 31(3), 425–439.
- McWey, L., Bolen, M. et al. (2009). I Thought I Was the Adult in This House: Boundary Ambiguity for Parents Involved in the Foster Care System. *Journal of Social Service Research*, 35(1), 77–91.

- Morgan, D. (1998). Practical Strategies for Combining Qualitative and Quantitative Methods: Applications to Health Research. *Qualitative Health Research*, 8, 362–376.
- Morgan, D. (2007). Paradigms Lost and Paradigms Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research*, 1(1), 48–76.
- Munem, W. L. (2010). *Social Research Methods: Qualitative and Quantitative Approaches: Quantitative and Qualitative Methods*. Boston: Pearson Education.
- Reams, P. and Twale, D. (2008). The Promise of Mixed Methods: Discovering Conflicting Realities in the Data. *International Journal of Research & Method in Education*, 31(2), 133–142.
- Reichardt, C. S. and Rallis, S. F. (1994). Qualitative and Quantitative Inquiries Are not Incompatible: A Call for a New Partnership. *New Directions for Program Evaluation*, 61, 85–91.
- Ritterbusch, A. (2011). A Youth Vision of the City: The Socio-Spatial Lives and Exclusion of Street Girls in Bogota, Colombia. PhD, Florida International University, FIU Electronic Theses and Dissertations.
- Royce, D., Thyer, B. and Padgett, D. (2010). *Program Evaluation: An Introduction* (5th ed.). Belmont, CA: Wadsworth.
- Salganik, M. and Heckathorn, D. (2004). Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling. *Sociological Methodology*, 34, 193–239.
- Shaffer, P. (2012). Ten Years of ‘Q-Squared’: Implications for Understanding and Explaining Poverty. *World Development*, 45, 269–285.
- Shulha, L. M. and Wilson, R. J. (2003). Collaborative mixed methods research. In A. Tashakkori and C. Teddlie (Eds), *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks: Sage, pp. 645–647.
- Tashakkori, A. and Creswell, J. (2007). Editorial: The New Era of Mixed Methods. *Journal of Mixed Methods Research*, 1(3), 3–7.
- Teddlie, C. and Tashakkori, A. (2012). Common ‘Core’ Characteristics of Mixed Methods Research: A Review of Critical Issues and Call for Greater Convergence. *American Behavioral Scientist*, 56(6), 774–788.

9

Challenges and Insights from Mixed Methods Impact Evaluations in Protracted Refugee Situations¹

Sally Burrows and Marian Read

Introduction

A series of four mixed methods impact evaluations and an overall synthesis report, conducted in 2011 and 2012, explored the contribution of food assistance to short, medium and long-term outcomes and impacts in WFP/UNHCR operations in protracted refugee situations. Refugee groups in four countries were studied:

- Myanmar Rohingya refugees in Bangladesh, arriving from the early 1990s.
- Refugees from the Central African Republic in southern Chad, arriving from 2002.
- Eritrean and Somali refugees in Ethiopia, arriving from 1990.
- Congolese refugees living in camps in Rwanda, arriving from 1994.

The overall objective of the evaluation series was to provide evidence and inspiration for future strategies to improve the contribution of food assistance for refugees in protracted situations. The main intended audiences were policy and strategy makers and programme implementers within WFP and UNHCR, governments hosting refugees, donor agencies and other relevant United Nations and non-governmental agencies. The contribution of this chapter is to illustrate how evaluators working in ‘real world’ situations and especially in humanitarian contexts experience and manage time, budget, data availability and political constraints (Bamberger et al., 2012). Following Leeuw and Vaessen (2009), it shows how a mixed methods approach can be used to address threats to validity and to ‘assess different facets of complex outcomes or impacts, yielding more breadth, depth, and width in the portrait than one method

alone can' (39). This chapter also explores the potential of contribution analysis in humanitarian contexts as the evaluations assessed the effects against the intended results set out in a Theory of Change derived from WFP and UNHCR documentation. Expected outcomes and impacts were improved nutrition and food security in the short term and increased self-reliance linked to local integration in the host country or voluntary repatriation or resettlement in a third country (known as 'durable solutions') in the long term. Each evaluation sought to measure effects of food assistance and also to identify the principal contextual and implementation factors influencing those effects and understand how they interacted together as part of a causal package, recognising that impact in complex contexts is often due to a combination of causes. The evaluations used a mixed methods approach, collected primary qualitative and quantitative data and analysed secondary sources of information. Data collection methods including sampling are discussed later in this chapter. These evaluations, conducted jointly by WFP and UNHCR, employed a wide range of strategies to enhance usability and followed a standard process at the preparation, inception, fieldwork and reporting phases, thereby facilitating the synthesis of findings.

The growing interest in impact evaluation, mixed methods and contribution analysis, combined with the challenges inherent in real world and especially humanitarian contexts, presents opportunities for strengthening their use in future evaluations. The series presented in this chapter provided insights into how to apply a mixed methods approach to complex questions in data-poor humanitarian contexts, where impact is rarely linear and few evaluations of impact are conducted. It contributed to the ongoing methodological discussion in the field of impact evaluation by providing an innovative example of how to use an integrated mixed methods design at all phases, and, how to integrate a 'contribution analysis' approach into the evaluation design to strengthen evidence of causal chain linkages. It also demonstrated how strategic use of evaluations can be strengthened by synthesising cases where the programme theory and implementation were similar but the contexts quite different. The absence of a conventional counterfactual, limited time series data, complexity of the Theory of Change to be tested, diversity of the contexts, limited budget and time to conduct the evaluations, and constraints of working in a humanitarian context were all challenges that needed to be addressed in preparing an appropriate evaluation design and throughout the evaluations.

The first part of the chapter discusses the institutional context of the evaluation series, including WFP's Evaluation Quality Assurance System

(EQAS), and presents relevant literature on the mixed methods and contribution analysis approaches. It then describes how the EQAS and mixed methods and contribution approaches were applied through the different phases of the evaluation process and assesses the threats to validity of findings and conclusions. Finally, it draws some conclusions and discusses future opportunities.

Institutional context for support in protracted refugee situations

The international community has aimed for a long time to support refugees in leading self-reliant independent lives. In 2002, after years of joint assistance to refugees living in protracted situations, WFP and UNHCR adopted a new Memorandum of Understanding – updated in 2011 – renewing their commitment to shift from a ‘care and maintenance’ approach that assumed the situation would be temporary, towards a policy of actively supporting ‘self-reliance and durable solutions’. WFP is responsible for mobilising cereals, edible oils and fats, pulses or other sources of protein, salt, sugar, high energy biscuits, and fortified blended foods for general, selective and school feeding programmes for protracted refugee situations. UNHCR is responsible for ensuring adequate supplies of non-food items and services, particularly those relevant to the safe and effective use of food items, such as cooking utensils, fuel, water and sanitation, medicines, soap, shelter, and appropriate health services. WFP has piloted and adopted new approaches and tools for food assistance, including providing cash or vouchers for refugees to purchase food in local markets. However, in-kind food rations were regularly distributed to refugees in all the country cases studied. UNHCR has used cash grants in repatriation programmes, but only recently started to give serious consideration to its use in refugee camps.

Given how inter-connected the agencies’ responsibilities vis-à-vis the food components are, an evaluation undertaken jointly by WFP and UNHCR was considered the optimal way to credibly evaluate the impact of food assistance. This was the first time impact evaluations had been undertaken on this topic. The series complemented UNHCR’s studies of other aspects of protracted refugee situations.

This series evaluated four situations with WFP/UNHCR collaboration and assessed the effects and contribution of food assistance for refugees living in camps, usually for more than 20 years, in Bangladesh, Chad, Ethiopia and Rwanda. The Theory of Change posited that WFP/UNHCR combined inputs and activities contributed to increased refugee

self-reliance over three stages of evolution: short term – lives saved, improved food consumption, safety and protection; medium term – improved food basket and nutritional status; and livelihood capacity established, setting the basis for long term outcomes of self-reliance and durable solutions – integration, resettlement or repatriation.

WFP's Evaluation Quality Assurance System for impact evaluations

The WFP Office of Evaluation's 'Evaluation Quality Assurance System (EQAS) for Impact Evaluations' (WFP Office of Evaluation, 2011) is a working tool that guided evaluation managers and independent evaluation teams through the entire evaluation process. Originally developed in 2009, and periodically revised, it complements relevant evaluation literature by focusing on the observance of evaluation principles of independence, credibility and utility and good practices of partnership and transparency. This was the second series that applied EQAS for impact evaluations, the first focused on the impacts of WFP-assisted school meals programmes. EQAS required the use of a mixed methods approach. It outlined for the individual evaluations and series the processes and quality standards to be adhered to, and it provided detailed guidance and templates for each process step and major written products. To enhance transparency and utilisation, EQAS explicitly required stakeholder participation throughout the evaluation process and preparation of a formal management response to the evaluation recommendations. EQAS covered five major evaluation phases: Preparatory, Inception, Evaluation including Fieldwork, Reporting and Dissemination.

Relevant literature on the mixed methods and contribution analysis approaches

Michael Bamberger, Jim Rugh and Linda Mabry in the second edition of *RealWorld Evaluation* (Bamberger et al., 2012) include new perspectives on the debate over the 'best' evaluation designs and give emphasis to mixed methods evaluation. They state '[m]ixed-method evaluation involves the use of a multi-disciplinary team so the evaluation design and hypothesis formulation, as well as data collection, analysis, dissemination and use, all draw on the theories and methods of two or more relevant disciplines' (Bamberger et al., 2012: 319). Mixed methods designs 'involve the planned use of two or more quantitative and qualitative methods of data collection and analysis' (Bamberger et al., 2012: 353).

This builds upon other influential evaluation literature, such as Frans Leeuw and Jos Vaessen's 'Impact Evaluations and Development' (2009) that also identifies the use of a mixed methods approach to address threats to validity and to 'assess different facets of complex outcomes or impacts, yielding more breadth, depth, and width in the portrait than one method alone can' (39). Stern et al. (2012) argue similarly for a broad range of designs and methods for impact evaluation, stating that '[t]he choice of methods (and overall design) needs to follow from the kinds of questions that are being asked; whilst taking into account the settings in which they are to be applied' (9).

Bamberger et al. discuss various options for sequencing qualitative and quantitative methods and identify concurrent designs as one option. In concurrent designs, the quantitative and qualitative approaches are used at the same time. 'An example of a concurrent design is where quantitative and qualitative data are collected simultaneously, using triangulation to compare information on outcomes, impacts and other key indicators from different independent sources' (Bamberger, 2012: 9). An advantage of sequential designs compared to concurrent designs is 'that logistics are easier to organize' but concurrent designs 'also make it more difficult to handle feedback, as adjustments would have to be made more quickly than for sequential designs' (Bamberger, 2012: 9). While each country evaluation employed concurrent designs, in some cases the quantitative data was in practice collected immediately prior to qualitative data (sequential), allowing for some adjustment of the latter (see discussion in section EQAS, Mixed Method and Contribution Analysis Approaches). Moreover, the overall series was conducted sequentially and evaluation teams were able to learn from the experience of the previous country evaluations thereby mitigating some of the challenges of concurrent design.

When initiating the evaluation series it was considered unlikely that evaluators could identify comparison groups 'without' treatment allowing for experimental design since registered refugees, living in camps in countries that receive WFP food assistance, are universally entitled to food rations. Given this fact, plus the nature of the evaluation questions and the complexity of the context, a theory-based design was preferred, using a mix of methods appropriate to that design (Stern et al., 2012). Mayne proposes an approach to 'attribution through assessing the contribution a programme is making to observed results' (2008: 1). He identifies a set of steps to be followed that do not require experimental data collection methods to establish causal linkages. The main steps focus on the development of a Theory of Change and risks to it, gathering the

existing evidence on the Theory of Change and assembling and assessing the contribution story, and challenges to it (Mayne, 2008). The contribution story is tested by asking questions about: how strong or weak the evidence is regarding the links in the results chain; the overall credibility of the contribution story; the agreement or not of stakeholders with the story; the validation or not of the key assumptions; and the clarity and understanding of the impacts of other influencing contextual factors (Mayne, 2008: 3).

EQAS, mixed methods and contribution analysis approaches

How EQAS, mixed methods and contribution analysis approaches were applied to the individual country evaluations and the synthesis is outlined below. The text following the Table 9.1 explains in greater detail each of the five evaluation phases: (a) Preparatory, (b) Inception, (c) Evaluation including Fieldwork, (d) Reporting and (e) Dissemination.

Preparatory phase

In line with EQAS, evaluation managers developed a Theory of Change for the whole series and identified the main analytical frameworks to be used. According to Bamberger et al. '[p]rogram theory is particularly useful for RealWorld Evaluation because it helps to identify the critical issues on which scarce evaluation resources should focus, and where possible a program theory model should be developed during Step 1 (scoping) of the evaluation' (2012: 181). The theory was drawn from agencies' policy, programme and Memorandum of Understanding documents, including existing logical frameworks for the countries included in the series. The main intervention studied was the contribution of food assistance to 'care and maintenance' of refugees in protracted displacement as well as to the intended shift to promoting self-reliance. Conceptual frameworks, concepts and definitions used were nutrition, food security, self-reliance and protection as they pertain to refugee contexts (WFP/UNHCR, 2013b). The evaluations tested evidence related to the Theory of Change and its assumptions about context and process.

Developing a Theory of Change retrospectively to incorporate agencies' policy statements broadened the evaluation scope to a strategic level. In Ethiopia, Rwanda and Bangladesh where governments had not formally supported employment or farming opportunities for refugees, short term results remained the focus. The evaluations explored what were the longer term and unintended results of the food assistance, as

Table 9.1 EQAS, mixed methods and contribution analysis approaches applied to the WFP/UNHCR Evaluation Series in protracted refugee situations

WFP's Evaluation Quality Assurance System for impact evaluations	Bamberger et al.: elements of a fully integrated mixed methods approach	WFP/UNHCR mixed methods impact evaluation series in protracted refugee situations
Evaluation phases	Mayne: contribution analysis	Four country-level evaluations
<i>1. Preparatory phase</i>		
A. Develop series Evaluation Quality Assurance System, select countries, complete terms of reference. Recruit multi-disciplinary evaluation team.	Attention to composition of research team.	Joint WFP/UNHCR evaluation managers. Joint Technical Reference groups. Expert reviewers. Multi-disciplinary evaluation teams, same leaders for two evaluations with different team members.
B. Develop Theory of Change for series and contextualise for individual country-level evaluations.	Integrate different Conceptual Frameworks and Planning approaches. Formulate hypothesis drawing from all disciplines. Step 1: Set out the attribution problem. Step 2: Develop the Theory of Change and the risks to it.	Series Theory of Change developed. Nutrition, food security, self-reliance/livelihoods conceptual frameworks used. Specific contexts, processes and assumptions added to series' Theory of Change for each country.

(continued)

Table 9.1 (continued)

WFP's Evaluation Quality Assurance System for impact evaluations	Bamberger et al.: elements of a fully integrated mixed methods approach	Mayne: contribution analysis	WFP/UNHCR mixed methods impact evaluation series in protracted refugee situations
			Four country-level evaluations
			Synthesis
<i>2. Inception phase</i>	Undertake field visit and literature review. Develop evaluation matrix. Prepare inception report. Review inception report and provide feedback.	Integrate Data Collection methods.	No inception phase.
<i>3. Evaluation phase including Fieldwork</i>	Conduct concurrent quantitative and qualitative data collection. Fieldwork debrief with country-based stakeholders, evaluation managers and technical reference group.	Systematic use of triangulation during data collection.	Inception reports included quantitative and qualitative sampling strategy and research methods and evaluation matrix, showing how the different methods inter-related to answer the evaluation questions.
		Step 3: Gather existing evidence. Assess the logic of the links in the Theory of Change.	Data collected was triangulated to confirm or corroborate results evidenced by one method with evidence from another method. In this way they checked for (a) any unexpected fundamental inconsistencies between findings emerging from

Stakeholders included refugees (both official and unofficial), national and local governments, WFP and UNHCR at many levels including executive bodies and implementing partners. Indirect stakeholders included other UN agencies, donor agencies and host communities.

different methods, requiring further data checking & analysis or (b) different perspectives emerging from different stakeholders and sources, which deepened understanding of causal linkages. Initial findings and conclusions presented to in-country field based informants and other stakeholders for feedback.

4. Reporting phase

Draft report analysis, conclusions and recommendations quality checked. Triangulated through review and comments by reference group and stakeholders. Revised, full report and summary completed. Management response prepared.

Integration of different approaches during Data Analysis including returning to the field to verify initial findings.

Step 4: Assemble and assess the contribution story, and challenges to it.

Data analysis included four cases which sequentially tested and verified the Theory of Change.

Data analysis integrated quantitative and qualitative findings and triangulated results. Contribution story was developed and assessed by evaluation team. Evaluation managers, technical reference group and stakeholders reviewed draft report and provided written feedback to team.

(continued)

<p>Step 6: Revise and strengthen the contribution story.</p>	<p>The contribution story was based on a synthesis of four evaluations. The standing High Level Meeting of Senior Managers from both agencies considered the findings and recommendations of the Synthesis Report and agreed a follow-up action plan.</p>	<p>As country evaluations were conducted sequentially, each successive team reviewed the contribution story in light of findings from previous evaluations as well as country specific findings, especially concerning the contextual and implementation factors influencing the causal chain.</p>
<p>5. <i>Dissemination phase</i> Web publishing. Dissemination of products. Presentation to Governing Bodies. Monitor follow-up to Management Plan.</p>	<p>Synthesis Report and management response were presented to Executive Bodies and follow-up action taken.</p>	<p>Each evaluation and management response was formally presented to the WFP and UNHCR's respective Governing Bodies. At country level, management and teams from WFP and UNHCR are responsible for follow-up based on management response, which WFP monitors. All products are available to public on web.</p>

implemented, from the perspective of the affected populations and in comparison to the intended theory, while acknowledging the constraints imposed by the context. In Chad, the results of a self-reliance strategy were closely studied. The government supported a policy of integration and allocated land to refugees, and phased reductions in food distributions had been implemented alongside the provision of seeds and tools in some camps. During fieldwork, evaluators encouraged stakeholders to discuss self-reliance opportunities and challenges in all countries. This laid the basis for understanding differences between theory failure and implementation failure.

In accordance with EQAS guidance, the preparatory phase included establishing evaluation management structures, quality assurance processes, reference groups and expert reviewers. This approach is consistent with evaluation literature where this key stage clarifies, among other issues, the activities needed to evaluate and who is responsible for each (Bamberger et al., 2012: 427). WFP and UNHCR evaluation managers agreed that all major decisions would be made jointly, including finalisation of terms of reference, country selection criteria and case selection, evaluation team recruitment and acceptance of inception and evaluation reports. This approach increased transaction costs, especially in the early stages, but strengthened quality of decisions and management process. They also agreed that communications to core stakeholders would be done jointly from the two agencies' evaluation offices and a joint funding formula for the series was adopted. From the outset WFP and UNHCR management were jointly committed at both country and headquarter levels to use the evaluation findings and recommendations. The joint evaluation management facilitated continued engagement by both agencies and follow-through on this commitment.

EQAS Quality Assurance sets out the roles and responsibilities of the evaluation managers, evaluation teams, reference groups and clarified and communicated norms and processes. Evaluation literature stresses the importance of evaluation offices working closely with agency management and other stakeholders to produce more realistic terms of reference and evaluation design and to help ensure that evaluations are designed to be useful and that findings and recommendations are actually used (Bamberger et al., 2012: 451). Reference groups included management and technical staff from WFP and UNHCR and stakeholders from case countries. Expert reviewers – an evaluation expert and a refugee expert – were engaged for two of the evaluations to provide feedback at evaluation design and report stages.

Contracting external evaluators, an EQAS practice, is consistent with approaches described in the literature as external evaluators usually have more credibility and lend the perception of objectivity to the evaluation (Morra-Imas and Rist, 2009: 18). Using multi-disciplinary teams to conduct the evaluations also served to facilitate triangulation at all stages. Evaluation teams included international and national members with expertise in anthropology, sociology, nutrition, economics, evaluation methodology and experience in refugee contexts, food security and livelihood analysis. The combination of international and national team members added value to the evaluation process, especially field-work and analysis. The national members brought deep contextual knowledge (political, social, economic, etc.) and cultural understanding essential to quality of design, appropriate analysis and triangulation. For example, how to pose certain questions and lines of enquiry so that they were received in the way intended and thus more likely to elicit accurate responses and locally appropriate interpretation of responses. Teams with previous WFP impact evaluation and research experience were used for each evaluation. Some team members had previously worked together in multi-disciplinary teams, which facilitated their internal communication and work processes. Evaluation managers promoted inter-disciplinary discussion within the team during the design phase and in triangulation of findings in the analysis and reporting phases as part of the quality assurance process. They also facilitated communications between the team and stakeholders throughout the entire process.

Inception phase

Methods of data collection

During the Inception phase evaluation managers and team undertook joint field visits and completed the literature reviews. The evaluation team then developed the evaluation matrix and prepared the inception reports. Reference groups, expert reviewers and evaluation managers reviewed the draft inception reports and provided detailed feedback and quality assurance. This EQAS phase is consistent with the evaluation literature that highly recommends using an evaluation matrix to organise the evaluation purpose and questions to match what is to be evaluated with the appropriate data collection techniques (Morra-Imas and Rist, 2009: 241).

As the evaluations were undertaken consecutively, the first country evaluation team (Ethiopia) designed and tested the mixed methods approach and tools from which the following countries built their

evaluation design, making adaptations as needed related to reflect, for example, cultural differences for both refugees and host populations, variation in government policies and practices relating to refugees in camps, and livelihood and migration activities of selected comparison groups. Maintaining flexibility in evaluation tools used allowed for findings to be nuanced to the context and helped to make the findings more reliable and useful to the country-level stakeholders. All the subsequent evaluation designs used the series' Theory of Change as their foundation.

During the inception stage the evaluation teams met in-country stakeholders, partners and beneficiaries; reviewed and assessed secondary data sources, identified locations of camps and data available on refugee households, and identified possibilities for comparison groups; and prepared for the logistics of fieldwork. Each inception report included the evaluation design, summarised in an evaluation matrix, indicators, quantitative and qualitative methodologies and data triangulation approaches, information on sampling (quantitative and qualitative), draft survey and interview instruments and made reference to the conceptual frameworks to be used in the analysis of the data. The inception mission was also the stage that the teams were 'formed', comprising up to six international team members with up to three national members,² and roles and responsibilities for each team member were clearly identified.

Data collection methods and sampling strategies

The sampling strategies for the different qualitative and quantitative methods were selected to permit generalisation and triangulation of findings, as illustrated on a country by country basis in Table 9.2. Countries are listed in the sequential order that the evaluations were conducted showing an evolution of the sampling approach during the conduct of the series. While refugee populations in each country fluctuated over time, the downward trend in randomly selected households generally reflected smaller populations in Chad and Bangladesh. In both Chad and Bangladesh host population groups were surveyed using random selection. In Chad the households in neighbouring villages were surveyed to compare the refugees to the local population and to provide an indication of the level of self-reliance that could be realistically reached given the local context. In Bangladesh a small sample of households from host communities provided descriptive comparators. Participatory rural appraisal exercises with local Bangladeshi communities sought to develop appropriate instruments taking into consideration cultural and social norms as well as the political context.

Table 9.2 Data collection methods and samples, by country

Data collection method	Ethiopia	Rwanda	Chad	Bangladesh
Primary Data, Quantitative Methods				
Quantitative	1,155 refugee households randomly selected in 2 camps	1,200 refugee households randomly selected in 2 camps	641 refugee households randomly selected in 3 camps.	349 registered refugee households randomly selected in 2 camps
Household Survey: refugees	(a) Inter-camp and intra-camp comparison between three different ethnic groups across two camps	(a) Inter-camp and intra-camp comparison across two camps	(a) Inter-camp comparison of refugee groups that had received different rations over time and comparison with 246 households randomly selected in host population	(a) Inter-camp comparison of refugee groups and comparison with 620 unregistered refugee households randomly selected in 3 makeshift camps
Household Survey: (a) Comparison Groups				(b) 100 poorest households non-randomly selected in host population
(b) Others				26 FGDs with refugees from 6 different locations
Primary Data, Qualitative Methods				
Focus Group Discussions (FGDs)	32 FGDs with 256 refugees from 4 camps and members of host populations	38 FGDs with refugees and members of the host population in/ around 3 camps	39 FGDs with refugees from 5 different refugee camps, 2 with host population	

(continued)

Table 9.2 (continued)

Data collection method	Ethiopia	Rwanda	Chad	Bangladesh
Key Informant Interviews	48 key informant interviews with implementing organisations and donors Interviews with 3 small business owners in 2 camps	54 key informant interviews with implementing organisations and donors	50 key informant interviews with implementing organisations and donors 24 household interviews in the camps	40 key informant interviews with implementing organisations and donors and refugee representatives 12 household interviews in the camps 6 host communities to include direct stakeholders with more programmatic and political insights
Participatory Rural Appraisal				

All countries: observation of conditions in camps; secondary data review including country operation documents; monitoring and Joint Assessment Mission reports, Health and Nutrition surveys; and external evaluations.

Source: Table developed by authors

Recognising the impossibility of using conventional counterfactuals and 'double difference' approaches, other comparisons were constructed based on what was appropriate in each context. In all countries inter-camp comparisons were possible and in two of the four countries 'opportunistic' comparison groups were identified where both the context and treatment differed. The following describes the stratified random sampling approaches used in all four countries for the quantitative household survey.

In Ethiopia and Rwanda the focus was on encamped and officially recognised refugees. All refugees received the same WFP food rations and UNHCR assistance. Therefore, survey design allowed for analysis of cross-sectional differences among camps and, to a lesser extent, among different socio-economic and ethnic groups within the refugee population. In Ethiopia a total of 1,155 households were surveyed, stratified by two camps and by ethnic group. In Rwanda 1,200 households stratified by two camps were surveyed.

In Chad there had been significant expectations that encamped refugees would grow their own food and in some camps there had been a stepped reduction to half rations over time. Survey design allowed for inter-camp comparison of refugee groups that had received different rations over time. A total of 641 refugee households were surveyed, stratified by groups receiving different quantities of food and living in three different camps which were differentiated by the ethnicity and wealth of resident households (WFP/UNHCR, 2013a). In addition, 241 households in the local host communities were surveyed.

In Bangladesh there were a large number of ethnic Rohingya people, living in unofficial make-shift camps. They were judged by UNHCR to be refugees within the criteria of the 1951 Refugee Convention, but not acknowledged as such by the government of Bangladesh and thus officially disqualified from receiving humanitarian assistance including food rations. This permitted a quasi-experimental design, allowing for comparison of refugee groups in two camps that received regular food rations with unregistered refugees in three makeshift camps with no food rations. Registered refugees (349 households) and unregistered refugees (620 households) in five different locations and poorest households within the host population (100 households non-representative sample) were surveyed.

An example from the Rwanda evaluation illustrated the purposive sampling strategy and the topics covered in interviews and focus group discussions. Within each camp the team interviewed groups of male and female refugees, food distribution committees, male and female refugee

leaders, youth groups and group leaders (men and women) and refugees representing different vulnerable groups. The team also interviewed one focus group each of men and women (non-beneficiaries) within three different host communities that had social and economic relations with refugees. The key informant interviews with implementing agencies elicited opinions, perspectives and strategic thinking on the potential for durable solutions within refugee contexts (WFP/UNHCR, 2012: 11).

Evaluation phase including fieldwork

In the EQAS Evaluation phase, the evaluation team conducted quantitative and qualitative data collection concurrently, while, consistent with evaluation literature, the evaluation managers followed the progress of the evaluation and provided feedback and guidance to evaluators during all phases of implementation (Bamberger et al., 2012: 450). Overall, the wide variety of data collection methods used provided a wealth of information and perspectives that were triangulated in the data collection and analysis stages. Emerging findings were presented during a fieldwork debrief, with country-based stakeholders, evaluation managers and the reference group, which provided an opportunity for validation of the findings and deepening of analysis and interpretation. This feedback and later comments on draft reports indicated where the evaluation teams needed to improve the accuracy of their analysis and sharpen their reporting so that the messages they intended to convey were indeed the messages received. The evaluation teams retained full authority to decide which comments they accepted (fully or partially) and which they rejected and gave transparent explanation of their decision.

Quantitative data collection methods

For this series two different quantitative methods were used to generalise results. The household surveys captured information related to:

- Access and receipt of the food basket and non-food items;
- Food security and coping strategies;
- Asset building or asset retention;
- Livelihood activities, including agricultural and other income-generating activities;
- Security and protection.

In addition, nutrition surveys had been conducted in refugee camps in each country over several years. The evaluations analysed nutrition

trends including measures of Global Acute Malnutrition (GAM), Severe Acute Malnutrition (SAM) and stunting rates.

Qualitative methods

Qualitative methods were used to gain a deeper understanding of perceptions and experiences related to nutrition and food security, self-reliance, livelihoods and protection results from refugees and other stakeholders including host communities (non-beneficiaries) and agency representatives. They also provided insights into (a) unintended results and sensitive issues not well captured by quantitative methods, (b) the contextual factors outside WFP and UNHCR control affecting the results observed and (c) the implementation factors within the agencies' control or influence that affected the results. All four evaluations included: qualitative interviews, including focus group discussions with refugees (and also unregistered refugees in Bangladesh) and with members of host communities; key informant interviews with WFP, UNHCR, government and non-governmental organisations; observation during transect walks of camps; and reviews of literature and secondary data. In each country, both qualitative and quantitative fieldwork was conducted within a period of a few weeks and in some cases quantitative data was collected immediately prior to qualitative data (sequential). This permitted teams to adjust the qualitative instruments to gain more relevant contextual information, enhance understanding of sensitive and emerging topics, such as the situation of unregistered refugees in and around camps and undocumented incidence of sexual and gender-based violence, and to identify and include previously unknown informants. National team members, in particular, helped to build trust with local communities and refugees by explaining the evaluation's purpose, objectives and target group (WFP/UNHCR, 2012: 1).

The Ethiopia evaluation provides an example of how data collection and analysis stages were combined, findings triangulated and initial results validated. For focus group discussions and interviews, each day of data collection was followed by a day of entering data into content analysis matrices, an iterative process that allowed the team to pursue questions, fill gaps and to conduct systematic checks on accuracy, consistency and reliability of collected data to ensure quality. Information from interviews was organised by different themes relevant to the Theory of Change as well as to those that emerged from the participants themselves. Initial themes included household food security, food distribution, non-food item distribution, livelihood programming, gender-based violence in camps, sexual exploitation of children, opportunities

for employment, relationship with host communities and institutional performance. Once all interviews in a camp were completed the matrices were shared with all team members to identify recurring and newly emerging themes. Through team discussions the major findings were triangulated and themes recorded and then analysed by ethnic group, gender and age group, where appropriate. Differences in perception between stakeholders were also noted. Additional key informant interviews were carried out with WFP and UNHCR regional and headquarter staff and a debriefing session was held at country level with a range of stakeholders to share preliminary findings and validate results (WFP/UNHCR, 2011b: 9).

Reporting phase

At the EQAS Reporting phase each evaluation team prepared a draft report including triangulated findings from primary and secondary sources, conclusions and recommendations. Overall, the guarantor of the validity of the findings was the range of sources and methods used to collect, assemble and triangulate the evidence. Consistent with evaluation literature the EQAS process ensured that the draft report underwent a thorough quality check by evaluation managers to ensure that it satisfied the terms of reference, evaluation findings were defensible and recommendations were realistic (Bamberger et al., 2012: 450). The evidence presented in the draft report was further triangulated through review and comments by reference groups, stakeholders and peer reviewers (where used). Following the review, a revised, full report was completed with annexes. In addition, a summary document was finalised for presentation to the Executive Bodies. WFP and UNHCR Management prepared a formal response to the recommendations that was reviewed by the Executive Bodies together with the evaluation summary report.

Both quantitative and qualitative analysis methods were used to deepen analysis and findings were subsequently compared. The qualitative analysis assisted interpretation of quantitative findings and deepened the analysis and understanding of causal links (WFP/UNHCR, 2011b: 9). The degree of corroboration of findings across methods also determined the degree of certainty with which evidence was presented. Conventional quantitative methods produced frequency tables, cross-tables and regression analysis, while qualitative data content matrices were further analysed. The quantitative methods analysed the main outcome results of household food security, measured using the standard indicators of Food Consumption Score (FCS), Household Dietary Diversity Score (HDDS) and the Coping Strategies Index (CSI), as well

as asset scores to measure household levels of wealth.³ Nutrition data were analysed for trends and compared to global benchmarks for malnutrition. Content analysis derived from interviews and focus group discussions was systematically compared to validate findings, check on inconsistencies, and highlight different interpretations. It was also used to test the validity of the assumptions in the Theory of Change. A similar triangulation was done with evidence found in secondary source documents, especially WFP/UNHCR Joint Assessment Mission reports. However, if a finding could not be sufficiently verified, a note of caution was included in the report (WFP/UNHCR, 2012: 7).

The Ethiopia evaluation provides an example of how the quantitative and qualitative data were combined and the Theory of Change was tested. The team matched the qualitative data with the quantitative data to determine if results were consistent. This triangulation process enabled the team to explain some of the findings that were emerging in the quantitative survey. For example, in Ethiopia, polygamous marital relations were found to occur in far greater numbers among the Somali refugees than in the surrounding general population, who are also Somali. Focus groups explained that this was an important food access strategy for refugees as each woman in a polygamous marriage takes a ration card under her name. When there were significant inconsistencies, both the qualitative data and quantitative data were re-examined. Through this iterative process, topics for further quantitative analysis were identified and the team examined emerging evidence relating to outputs, outcomes and impacts identified in the Theory of Change and determined whether their assumptions were correct. The qualitative data provided significant insights regarding the assumptions (discussed in WFP/UNHCR, 2011b: 9). The validity and reliability of the findings generated through these data collection, analysis, triangulation and verification processes were strengthened by having multiple disciplines and international and national members on the evaluation team.

Contribution analysis

For contribution analysis, Mayne suggests that after gathering the existing evidence on the Theory of Change, to assemble and assess the contribution story, and the challenges to it, and then to seek additional evidence and revise and strengthen the contribution story (2008: 1). This approach was used in the analysis and reporting phases for each country evaluation and to a certain extent subsequently for the synthesis report, where the four evaluations were systematically analysed.

As the Theory of Change had three time dimensions, the analysis examined evidence related to short, medium and long-term outcomes and impacts and the programme logic. Following this analysis, the synthesis report concluded that the contribution story was plausible in the short term but not sufficient in the medium and long-term perspectives. Contribution analysis offered a structured and systematic approach to analysing a theory with complex and not entirely linear influences. To understand the challenges to the contribution story, assumptions regarding the contextual factors outside agencies' control and implementation factors within their control were examined. For example, the assumption that food distributed to refugees on a biweekly or monthly basis was taken home and consumed was found to be only partly true as significant quantities of food and other items distributed by the agencies were used for income and collateral.

Prior to the finalisation of the reports, several iterations of the findings and recommendations in presentations and reports were widely shared with agency staff and management, government representatives, donors and implementing partners' field staff to identify factual errors, major omissions or misinterpretations and to complete the quality assurance steps. Upon completion of each evaluation report, country-level management and teams from WFP and UNHCR, in consultation with the government, considered the report's recommendations and prepared a management response that included follow-up activities to be implemented over the course of a few years.

The synthesis report facilitated cross-country learning and led to strategic level action at the management and governance levels. The standing high-level meeting of senior managers from both agencies considered the findings and recommendations and agreed to a follow-up action plan.

Dissemination phase

At this final EQAS phase the evaluation managers, consistent with evaluation literature, disseminated the evaluation results to key stakeholders and other audiences and promoted the implementation of recommendations and the use of evaluation results (Bamberger et al., 2012: 451). The Directors of Evaluation, from WFP and UNHCR, made formal presentations to the Executive Bodies for each country evaluation and the synthesis report together with managers from both agencies, who presented the joint management response including follow-up action plans. Just prior to the Executive Bodies' meetings, the country evaluations' full and summary reports and the synthesis report were published

on WFP and UNHCR public websites. To enhance accessibility and use by diverse audiences, short evaluation briefs for each evaluation and a colourful booklet of the synthesis report and management response, including photographs, were also published on public websites and disseminated. WFP systematically monitors the follow-up action taken vis-à-vis the management responses over a period that spans several years.

The synthesis report (WFP/UNHCR, 2013a) presented evidence from the four country impact evaluations, analysing patterns and drawing out systemic issues emerging from the series as a whole and testing the logic and assumptions of the common Theory of Change. Agency-wide utilisation at the policy and strategy levels in both WFP and UNHCR was greatly enhanced by presenting evidence that could be generalised across the four different country cases and by drawing attention to the weaknesses in the programme theory related to medium- and long-term outcomes and impacts.

The overarching conclusion from the series is that the intended evolution towards self-reliance had not occurred. The international community's response to refugees in protracted crises is failing to deliver. To resolve the issues blocking progress, there is need for a change in paradigm with concerted action among all actors, backed by political will and financing to enable refugees to make productive contributions to the countries where they live, and to support other long-term durable solutions where possible.

Management response and follow-up action

In presenting their response to their governing bodies, management of WFP and UNHCR agreed that the synthesis provided 'compelling evidence of the need for all parties to renew their commitment to supporting the attainment of durable solutions for refugees in protracted situations' (WFP/UNHCR, 2013c: 29). Subsequently, UNHCR in its Global Appeal for 2014 and 2015 stated that:

[B]ased on the mixed-method impact evaluations [in 2011–12, assessing the contribution of food assistance to durable solutions in protracted refugee situations, WFP and UNHCR agreed at a high-level meeting in January 2013, to develop a strategy for the transition to self-reliance, using a holistic approach and establishing the partnerships necessary to achieve it at the corporate and country levels, including with full engagement of UN Country Teams, development actors, host and donor states. The strategy will be implemented in three countries: Chad, South Sudan and Uganda, using new economic

assessment criteria and methodologies to help target interventions. (UNHCR, 2013)

In following up recommendations of the individual impact evaluations WFP and UNHCR have taken measures to introduce cash or voucher transfers to refugees in Ethiopia, Rwanda and Chad replacing food distributions, initially on a pilot basis, while in both Chad and Bangladesh efforts are being made to support vulnerable host populations as well as refugees. In Bangladesh the agencies continue to make concerted efforts to have the government's policies affecting support to the unregistered refugees changed. In Chad self-reliance strategies based on the updated livelihood approach are being pursued.

An assessment of the threats to validity of the findings and recommendations

This section presents an analysis of the threats to validity of the findings and recommendations of the evaluation series. It follows Bamberger et al.'s method for conducting a summary assessment of an evaluation's design, implementation, products and communication strategy (2012: 126). Undertaking an assessment when the evaluation series is complete provides insights into how the methodology used supported validity and suggests how the methodology could be strengthened in future evaluations.

Evaluators distinguish between internal validity (conclusions about the contribution of project and programme interventions to explaining observed changes in the beneficiary population may not be valid) and external validity (conclusions about the replicability of the project in other contexts may not be valid) (Bamberger et al., 2012: 127). Internal validity includes objectivity, internal design validity/reliability, statistical conclusion validity and construct validity; whereas external validity is about transferability, utilisation validity asks 'was the evaluation designed to address the right questions?' (Bamberger et al., 2012: 129).

Internal validity/objectivity

Overall, the conclusions of the four evaluations and the synthesis were drawn from available evidence and were free of researcher bias. The EQAS and mixed methods approach ensured that: (a) methods and procedures were well described; (b) evaluation data was retained and made available for the synthesis and future use within the agencies; (c) data presented supported the conclusions; (d) multi-disciplinary teams with

international and national members were used, with appropriate technical and evaluation skills and knowledge and (e) a range of qualitative and quantitative methods were used to control for bias.

With hindsight, objectivity could have been enhanced if the contribution analysis approach had been followed more closely for each individual impact evaluation and analysis of a competing hypothesis had been more explicitly explored and recorded in the series. For example, this could perhaps have involved a case where registered refugees living in protracted situations did not receive WFP food assistance but rather support from another agency or government such as a social safety net, employment or agricultural land that logically could have resulted in the same outcomes and impacts.

Internal design validity/reliability

Overall the adherence to the evaluation phases set out in EQAS and use of similar methods and tools to collect and analyse data in each country evaluation ensured a process that was consistent, coherent and reasonably stable over time. The findings were credible to the agencies and populations studied and to a wide range of readers. Reasons were found why the causal relationships outlined in the Theory of Change were valid for the short-term but not valid for the medium and long-term periods. EQAS, mixed methods and contribution analysis approaches ensured that: (a) there was sufficient information to provide a credible and valid description of the refugees and their context; (b) triangulation among complementary methods and data sources produced generally converging conclusions which were well linked to the Theory of Change; (c) negative evidence was sought and found, though rival explanations could have been more explicitly considered; (d) where accounts of different observers did not converge this was recognised and addressed; for example, in the Bangladesh evaluation if a finding could not be sufficiently verified using primary or secondary data, as was the case for some 'anecdotal' findings, a note of caution was included in the report (WFP/UNHCR, 2012: 7); (e) peer and colleague reviews were used systematically throughout the evaluation process and (f) the steps for confirmation of the Theory of Change were applied.

Internal design validity could have been strengthened if more information on attrition of refugees had been available. Data on what had happened to individuals or households who had left the camps was patchy. Tracer studies of some households no longer in the camps were considered, but would have required considerably more time and financial resources with high uncertainty of success in tracing refugees.

Threats to statistical conclusion validity

Evaluation designs for quantitative sampling ensured that the samples were large enough to detect statistically significant differences between refugee groups living in different camps and with the comparison groups of unregistered refugees (Bangladesh only). Statistical tests ensured that observations were independent of each other.

Limitations to the data collected by the evaluation teams were taken into account during the analysis stage and were systematically recorded in the reports. For example, Bangladesh, Chad and Ethiopia faced the constraint of inaccurate camp population databases. Enumerators often found particular households were not where they were supposed to be or were no longer in the camp (WFP/UNHCR, 2011a). Possible biases in quantitative survey data also arose from the timing of surveys in Chad and Rwanda, which could have had an impact on accessibility, respondent availability, food consumption and dietary diversity (depending on the season and last food ration distribution).

The main threats to statistical conclusion validity stemmed from the nature and context of the refugee situation and the lack of conventional counterfactual and baseline data for the refugee populations. Future evaluations should ideally be designed over a longer timeframe so that 'pre-test' and 'post-test' surveys could be undertaken.

Construct validity

EQAS, mixed methods and contribution analysis approaches ensured that multiple indicators and methods were used to measure the complex constructs defining processes, outcomes and impacts as well as contextual variables set out in the Theory of Change. The construction of the Theory of Change from literature and testing it in the first country, and further developing it through the inception missions, clarified the constructs to be used to define the outputs, outcomes and impacts. To maximise coherence in how the constructs were interpreted by evaluation teams, two different evaluation companies were used for the four evaluations (two each) and two team leaders led two evaluations each. Prior to the commencement of their work team leaders were briefed well by evaluation managers on the definitions of the constructs. This facilitated a degree of consistency within and between teams in their interpretation.

The main threat to construct validity was that although WFP and UNHCR had an agreed definition of 'self-reliance' and 'protection', there are nevertheless constantly evolving differences in interpretation in different cultural contexts and the diverse teams used a variety of

indicators to study them. These differences particularly influenced the extent of comparability of the individual evaluation findings for the synthesis report. There are trade-offs in how these threats could be handled in future evaluations. An evaluability assessment during the preparatory phase clarifying and pre-defining all the constructs could help to ensure consistency across cases. Strict adherence to pre-defined global construct definition and measurement, however, could undermine the validity of the evidence if room is not left to adjust the constructs in line with the context.

External validity

Using EQAS, mixed methods and contribution analysis all helped to build the external validity in this series. The quantitative sample design enabled some generalisations, but was always balanced against the need to maintain some flexibility of design to be appropriate to context in order to maintain internal validity. Analysis of the assumptions, especially of host government policies towards refugees, demonstrated that the context had a strong influence on outcomes, especially on the potential of refugees to achieve self-reliance or local integration. A very wide range of readers/stakeholders reported the findings to be consistent with their own experience at country level, and at agency management and governance levels.

Threats to external validity related to the small sample of countries selected for the series, which did not include protracted situations where refugees had become self-reliant through local integration or repatriation. A future series might consider including a case where self-reliance had occurred that could provide useful information on the contribution of food assistance to these impacts.

Utilisation

The two key factors contributing to utilisation were the series model itself and adherence to the EQAS processes. The problematique of increasing self-reliance for refugees in protracted situations is highly complex. The political, social and economic issues involved are not amenable to simple solutions and require simultaneous action to common ends on the part of host governments, donor governments, UNHCR and WFP and refugees themselves. Individually, each evaluation contributed solid and motivating evidence to help build political will to bring the parties in each country together to seek solutions. Collectively, and with common and systemic lessons and patterns synthesised, the series of reports from multiple countries provided evidence to feed into renewal of WFP and

UNHCR's global strategy. The evaluations provided guidance for future action through the formulation of specific recommendations for each country and for the agencies generally at strategy and policy setting levels and the reports for each evaluation were presented to different stakeholders in country, at the governing body level and to inter-agency management groups.

The EQAS process required identification and analysis of stakeholders at the very start of the evaluation process and formation of reference groups from among them. Reflecting the series model, individual reference groups were formed in-country for each evaluation. Another reference group comprised regional and headquarters level staff from both agencies. The pre-existing inter-agency group of senior managers responsible for operations comprised a third reference group. Each reference group reviewed and commented on the main evaluation outputs of terms of reference and draft evaluation report. The evaluation teams considered the comments and recorded whether the comments were accepted (partially or fully) or not accepted and why. In this way, the independence of the evaluation was maintained, but relevance, accuracy, transparency and utility of the evaluation products were enhanced. Engagement of these key stakeholders throughout the evaluation process contributed to the commitment to follow up.

The series model and EQAS processes ensured that the findings were useful to WFP and UNHCR country and headquarter level staff and management and indirectly to refugees, host governments and other stakeholders. Follow-up actions taken by UNHCR and WFP management related to recommendations of each individual evaluation as well as the synthesis were tracked over years. The follow-up project documents suggest that the series had significant influence on subsequent project design and funding requests. The UNHCR Global Appeal 2014 and 2015 specifically cite the series and mention a new strategy being undertaken by WFP and UNHCR to enhance refugee livelihoods in three pilot countries. Government representatives from at least one of the countries also expressed interest in the reports from other countries for peer-to-peer learning. The series also contributes to the existing body of knowledge on protracted refugee situations, as the reports are available to the public on WFP/UNHCR websites.

Utilisation could have been improved further through clear processes for dissemination of the final version of the evaluation reports and other products in the countries where the evaluations took place. This was left up to the agency staff based in the field and was not monitored. Having a dissemination strategy for evaluation products for local

governments, local UN and implementing partners, refugees and host populations could enhance evaluation utilisation and implementation of recommendations.

Conclusions

This series of four mixed methods impact evaluations and a synthesis report analysed the contribution of food assistance in WFP/UNHCR operations in protracted refugee situations. The series provided a case study of how mixed methods and contribution analysis approaches, coupled with a systematic evaluation quality assurance system (WFP's EQAS) can be used to ensure credible evaluation processes and products that resulted in good levels of utilisation by a range of stakeholders. Within this chapter insights about the elements of a credible evaluation process were identified by examining how EQAS provided a strong framework within which the various elements of the mixed methods and contribution analysis approaches were applied. The examination also showed how these approaches effectively mitigated the challenges of evaluating in humanitarian contexts, where the absence of a conventional counterfactual, limited time series data, a complex Theory of Change and limited budget and time to conduct the evaluations were constraints. The insights gained from EQAS, mixed methods and contribution analysis approaches are noted below, followed by discussion of opportunities for strengthening impact evaluation, mixed methods and contribution analysis in real world contexts.

Insights regarding EQAS

EQAS created an enabling environment by setting out how the series was designed and conducted. Built into EQAS was a set of processes that promoted adherence to good evaluation principles. EQAS ensured that skilled international and national evaluation teams brought a multi-disciplinary approach to the series. EQAS focused from the outset on how the evaluation series would be used for programme and policy improvements at multiple levels. Consistent with EQAS' transparency and partnership principles, follow-up action was facilitated through stakeholder participation from the outset in evaluation reference groups and by mandating a formal management response with follow-up action presented to Executive Bodies. Adherence to EQAS' phases, templates and quality assurance guides on content and process, by evaluation managers, evaluation teams, reference groups and other stakeholders, ensured that the strengths of quantitative and qualitative approaches

were incorporated into each phase. Consistently following EQAS' five phases (preparatory, inception, evaluation, reporting and dissemination) helped to ensure common evaluation designs across different country contexts, integrated triangulation throughout all phases and enhanced transparency in the evaluation process for stakeholders. The common EQAS system facilitated partnership of WFP and UNHCR at multiple levels from evaluation managers, who jointly initiated and managed the country evaluations and synthesis, to the joint inter-agency management group which took responsibility for implementing strategic and policy recommendations.

Insights regarding the mixed methods approach

EQAS clearly mandated a mixed methods approach. The evaluation designs used were consistent with mixed methods approaches described in relevant evaluation literature, such as Bamberger et al.'s 2012 *RealWorld Evaluation*. The strengths of mixed methods were realised in the four evaluations and the synthesis report. It allowed for examining different facets of outcomes and impacts and resulted in a comprehensive portrayal of refugees in protracted situations through triangulation at multiple levels. Quantitative methods greatly enhanced the series' credibility. Collecting primary household level data and using existing nutrition evidence addressed selection bias, allowed for generalised results, permitted replicability by using standardised approaches and strengthened validity.

The qualitative methods added breadth and depth, strengthened findings' validity and assisted generalisation. Evaluators adapted qualitative instruments to address emerging issues and used purposive sampling to assess different facets of outcomes and to examine different stakeholder views. Qualitative methods facilitated collecting information on sensitive topics and difficult-to-reach groups. Evaluators noted that using qualitative methods to collect data regarding programme implementation and context proved critical in explaining the results, including the relative importance of various influencing factors in highly complex environments.

Multi-disciplinary evaluation teams, comprising international and national members, employed a range of analytical frameworks, enhancing the validity of the process used and findings generated. These benefits outweighed the transaction costs involved in team members reaching mutual understanding. Evaluation design included multiple data sources and methods of data collection, repeated observations over time and analysis incorporating a range of perspectives. Conducting

primary quantitative and qualitative data collection concurrently resulted in iterative data analysis and validation between team members while fieldwork was ongoing.

Insights regarding contribution analysis approach

As outlined in Mayne (2008), the contribution analysis approach provided a systematic way to assess the Theory of Change. The particular humanitarian context and timing of the evaluations led to the expectation that quasi-experimental design would not be possible. A Theory of Change, developed during the preparatory phase and contextualised for each country, provided a common programme cause and effect logic to be tested. It also identified common assumptions which placed emphasis on collection, analysis and reporting on process and contextual information. At the synthesis level, when programme theory was tested across evidence from four countries, the congruence of results led to strong conclusions about the plausibility of the existing theory, which informed recommendations and subsequent follow-up action by stakeholders.

Opportunities for strengthening impact evaluation, mixed methods and contribution analysis in real world contexts

A first area of opportunity is to adopt an evaluation quality assurance system that ensures quality, promotes utilisation and provides a framework for the mixed methods approach. In particular, utilisation can be enhanced when the evaluation system ensures that major stakeholders are engaged throughout the evaluation process and a formal management response to evaluation recommendations is mandated within the organisation.

A second opportunity is to ensure that a systematic mix of quantitative and qualitative methods is used in complementary ways appropriate to the evaluation questions to be answered and to the context, using multi-disciplinary teams, and closely following relevant literature providing guidance on the methodology. In humanitarian cases where an experimental design method often cannot be used, mixing methods in a theory-based approach and incorporating contribution analysis steps and processes can provide robust causality arguments.

A final opportunity is to conduct a series of impact evaluations using the same Theory of Change and, as far as possible, methodology, drawing evidence from a number of cases. Common and systemic lessons and patterns can be synthesised from multiple cases and effectively used to inform organisational-level policy and strategy, even where

there is not sufficient consistency of method for statistical comparison. This approach is particularly powerful when evaluations are conducted jointly with partners thereby widening the focus and strengthening use and follow-up. Furthermore, publicly disseminating the reports enables a contribution to the existing body of knowledge on the topic evaluated.

Notes

- 1 Disclaimer: The views expressed in this chapter are those of the authors and do not necessarily reflect the views of the United Nations World Food Programme or the UNHCR.
- 2 Excluding multiple enumerators working on the quantitative surveys.
- 3 FCS combines food diversity, food frequency and the relative nutritional importance of different food groups. HDDS represents the average number of food groups, consumed by households during a 24-hour period, taken from a list of 12 food groups. CSI measures how often 12 strategies were used by households to overcome difficulties in accessing food.

References

- Bamberger, M. (2012). *Introduction to Mixed Methods Impact Evaluation*, Impact Evaluation Notes No. 3. Washington, DC: InterAction. URL: <http://www.interaction.org/sites/default/files/Mixed%20Methods%20in%20Impact%20Evaluation%20%28English%29.pdf>
- Bamberger, M., Rugh, J. and Mabry, L. (2012). *RealWorld Evaluation* (2nd ed.). Thousand Oaks: SAGE.
- Leeuw, F. and Vaessen, J. (2009). *Impact Evaluations and Development. The Network of Networks on Impact Evaluation*. Washington, DC: Independent Evaluation Group.
- Mayne, J. (2008). *Contribution Analysis: An Approach to Exploring Cause and Effect*. Netherlands: Institutional Learning and Change (ILAC) Brief 16.
- Morra-Imas, L. and Rist, R. (2009). *The Road to Results*. Washington, DC: The International Bank for Reconstruction and Development/The World Bank.
- Smith, L. and Subandoro, A. (2007). *Measuring Food Security Using Household Expenditure Surveys*. Washington, DC: International Food Policy Research Institute.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., and Befani, B. (2012). *Broadening the Range of Designs and Methods for Impact Evaluations: Report of a Study Commissioned by the Department for International Development*. DFID Working Paper 38. London, UK: DFID.
- UNHCR. (2013). *Global Appeal 2014 and 2015*. Geneva: UNHCR.
- WFP Office of Evaluation. (2011, updated in 2013). Evaluation Quality Assurance System – Impact Evaluations. Rome: WFP. <http://documents.wfp.org/stellent/groups/public/documents/reports/wfp231096.pdf>, accessed 19 August 2014.
- WFP/UNHCR. (2011a). *The Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations: Its Impact and Role ETHIOPIA: A Mixed Method Impact Evaluation Vol. I Full Report*. Rome: WFP.

- WFP/UNHCR. (2011b). *The Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations: Its Impact and Role. Ethiopia: A Mixed Method Impact Evaluation Vol. II*. Rome: WFP.
- WFP/UNHCR. (2011c). *The Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations: Its Impact and Role in Rwanda (2007–2011), Annexes*. Rome: WFP.
- WFP/UNHCR. (2012). *The Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations: Its Impact and Role in Bangladesh: A Mixed Method Impact Evaluation Vol. I*. Rome: WFP.
- WFP/UNHCR. (2013a). *Synthesis Report of the Joint WFP and UNHCR Impact Evaluations on the Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations*. Rome/Geneva: WFP/UNHCR.
- WFP/UNHCR. (2013b). *Joint Assessment Missions: A Practical Guide to Planning and Implementation*. Geneva: UNHCR. www.alnap.org/resource/8835, accessed 19 August 2014.
- WFP/UNHCR. (2013c). *Summary Report of the Joint UNHCR/WFP Impact Evaluation on the Contribution of Food Assistance to Durable Solutions in Protracted Refugee Situations – Chad*. Rome: WFP.
- WHO. (2000). *The Management of Nutrition in Major Emergencies*. Geneva.

10

Competing Interpretations: Human Wellbeing and the Use of Quantitative and Qualitative Methods

J. Allister McGregor, Laura Camfield and Sarah Coulthard

Introduction

The renewed focus on human wellbeing in international policy and academic debate encourages us to think again about how quantitative and qualitative methods are combined in understanding development and poverty. During the 2000s the Q-Squared movement provided a wide range of valuable and insightful material that explored the possibilities for and obstacles to combining qualitative and quantitative research methods (Kanbur and Schaffer 2007). Since then a growing number of national and international statistical agencies have recognised the importance of seeking to assess the impacts of governance and public policy in terms of human wellbeing and in doing so they are adopting a hybrid framework that takes account of the objective and subjective dimensions of human wellbeing (OECD 2011; UK ONS 2011). This chapter presents the argument that if development policy and practice is to promote a focus on human wellbeing then it is necessary to consider what types of data are required to do that. The current 'business as usual' in international development focuses mainly on material wellbeing and tentatively and occasionally stretches out to encompass human development outcomes (Alkire and Foster 2011). This chapter argues that in order to assess a more holistic conception of human wellbeing (McGregor and Sumner 2010), it is necessary to generate three types of data: objective, subjective and inter-subjective.¹ To achieve this, both quantitative and qualitative methods are required and it is necessary to reconsider how these are sequenced and mixed to generate the data.

Following the publication of the Final Report of the Commission on the Measurement of Economic Performance and Social Progress in 2009 there has been tremendous upsurge of initiatives to find ways to assess

development in terms of its impacts on human wellbeing. Chaired by Joseph Stiglitz, Amartya Sen and Jean-Paul Fitoussi, the report concludes with a recommendation and challenge to policymakers, academics and civil society actors to 'shift emphasis from measuring economic production to measuring people's well-being' (Stiglitz et al. 2009, p. 12). After all, as they put it in a subsequent publication, 'What we measure affects what we do. If we have the wrong metrics, we will strive for the wrong things. In the quest to increase GDP, we may end up with a society in which citizens are worse off' (Stiglitz et al. 2009, p. xvii). But, what does measuring wellbeing actually mean? In this chapter we outline the key elements of a wellbeing approach, discuss some of the conceptual and methodological challenges, present commonly used wellbeing frameworks and provide an example of its use in an evaluation in Zambia. In doing so we demonstrate how a wellbeing approach can potentially bridge the disconnect between macro- and micro-level policy thinking, address ineffective governance caused by limited engagement of the potential beneficiaries of development policies and provide a more rigorous and politically sensitive framework for participatory work in development. We begin by outlining the distinctive characteristics of a human wellbeing approach.

Our conceptualisation of wellbeing draws on work initially developed by the UK ESRC Research Group on Wellbeing in Developing Countries (WeD, see Gough and McGregor 2007). In this framework, wellbeing is understood as

[a] state of being with others which arises where human needs are met, where individuals and groups can act meaningfully to pursue their goals, and where they are satisfied with their way of life. (McGregor 2008)

This approach requires us to engage with human wellbeing both as an outcome and as a process, taking into account three inter-dependent dimensions: the *material*, the *relational* and the *subjective*. These are grounded in specific cultural, social and ecological contexts and in a particular time. Although the third dimension is labelled 'subjective wellbeing' this is to distinguish it as a necessary inquiry into people's own perceptions of their 'quality of life'. While the material and relational dimensions can be assessed using objective data, each of the components of these can also be assessed subjectively. Thus, for example, in relation to material wellbeing, whether people feel their income is secure and sufficient to meet their needs is measured, as opposed to how

much income they have. Looking at both the objective and subjective component not only provides more information, but also highlights situations where the two may be in conflict. For example, objectively people may have a large income, but it is not reliable or is earned through work that is perceived as demeaning.

Material wellbeing relates to the extent to which men, women and children meet their basic human needs (in terms of food, source of income, housing, access to services, etc.), on a secure and sustainable basis. The material dimension therefore focuses on practical welfare and standards of living (White 2010). These are often assessed using a household survey; however, a wellbeing approach would also entail preparatory qualitative research to identify what local people consider to be the most important material needs and the thresholds at which they consider them to be met (Camfield et al. 2008).

Relational wellbeing refers to personal and social relations. Rather than viewing relationships as something (an asset) that someone possesses, men, women and children are viewed here as becoming 'who and what they are in and through their relatedness to others'. Relatedness is also central to the construction of distinction and difference and can encompass illbeing as well as wellbeing.

Subjective wellbeing refers to people's perceptions of what constitutes a good or improved quality of life, and the (changing) extent to which people are satisfied that they are able to realise the goals that they believe are necessary to lead a good life. This dimension of wellbeing refers not only to the usual basket of 'livelihood assets', but also incorporates an inquiry into non-material goals including cultural values, norms and belief systems. On an individual level it accounts for notions of self, individual and shared hopes, fears, aspirations, expressed levels of satisfaction or dissatisfaction, trust and confidence.

Calls for greater attention to the specificities of human wellbeing have been present in the study of international development since its emergence in its modern form in the post-colonial era. In his seminal article, 'The Limitations of the Special Case', Dudley Seers (1967) argued that a blind acceptance of a conceptual framework for development built from the experiences of a limited number of special-case economies (Western industrialised countries) was inhibiting our understanding of the realities and policy problems of most other (developing) societies. Seers' pleas for what was effectively an 'anthropologising' of economics fell on deaf ears and post-war neoclassical economics increased in its sophistication and developed into an ever more powerful and hegemonic conceptual machine in the international development arena. Aside from

the dominance of this science and the GDP paradigm in both academic and policy spheres, a key element in its maintenance has been the development of a national and supra-national statistics industry mainly oriented to serving this intellectual and ideological pre-disposition (Merry 2011; Shore and Wright, 2015).

The observation by Stiglitz et al. on the relationship between metrics and goals affirms a view that has been presented in many forms in international development studies since the work of Seers, but which has not been so concisely and powerfully expressed: that economic models and metrics of development and the policies that flow from them are incomplete in the ways that they comprehend the challenges that are encountered in the day-to-day lives of people in developing economies and societies.² For development to be effective in terms of improving peoples' lives it must necessarily entail considerations of money, jobs and growth, but for development policy and practice to be effective in a broader sense it must take account of much more than that. Following Stiglitz's observation about the misdirection caused by the wrong metrics, it is important that the development industry finds ways of measuring what matters for the people who experience and make the development processes, particularly in social, economic and cultural contexts. In 2013, United Nations General Assembly adopted a resolution that signalled an acceptance of the consensus that a focus on GDP alone is not adequate for the measurement and promotion of human prosperity, and that 'a more inclusive, equitable and balanced approach is needed to promote sustainability, eradicate poverty and enhance well-being' (UN General Assembly Resolution 65/309, 'Happiness: towards a holistic approach to development', p. 1).

Why measuring what matters, matters

In this chapter we focus our discussion on approaches to understanding and measuring human wellbeing at the grass-roots level of development policy and practice. Many of the initiatives that have followed the publication of the Stiglitz-Sen-Fitoussie Report have concentrated on measuring development performance at the level of the nation state (see, for example, OECD 'How's Life?' 2011). It is important, however, that the flaw that is illustrated by the theoretical divide between macro- and micro-economic thinking (Coase 1998, Chang 2003, Pouw and McGregor 2014, p. 9) is not replicated in the wellbeing paradigm and that work at the 'top' level does not become conceptually disconnected from the work at the 'bottom', that focuses on the lived realities of

people, households and particular communities. While there have been significant advances in the top level measurement initiatives, there has been much less discussion about how the concept of wellbeing is made relevant and operational for the public policy frontline: in this case for development policymakers and practitioners working at the level of the project or programme (Rojas 2008, Camfield and McGregor 2009, White 2014).

In order to understand why taking account of human wellbeing matters in international development it is important to recognise that public policy efforts are a key component of contemporary governance. As a particular form of public policy, development programmes and projects can be regarded as specific and organised efforts to create better outcomes for people and/or for society (they are not always the same). Like all such policy measures, their success or failure depends on the extent to which the programme or project engages effectively with the values and behaviours of those people for whom they are intended or upon whom their realisation depends. The relationship between the majority of the people and the development policies of their countries is currently at a low ebb. As participants in a 2012 Summit reflecting on the challenges and opportunities for international development in the 21st century put it, 'people have become bystanders in their own development' (Bellagio Initiative 2012a, p. 46). Melamed (2011) observes that there is often a 'disconnect' between the views and priorities of poor people and those involved in development decision-making (Bebbington et al. 2007). This suggests that in the interests of greater aid effectiveness, but also for improving the contribution of development policy processes to effective governance, there is a need for methods and practices that are founded in greater transparency and accountability and which enable better mutual comprehension.

The struggle to have the idea accepted that people might participate directly in shaping their own development policies has been a long one and has gained some success (Chambers 1997). 'Participation' has been taken on board at a level of high rhetoric and then in some large and high-profile set-piece studies (e.g. 'Voices of the Poor'). This has led to it being co-opted into some development routines such as the staged elements that have produced Poverty Reduction Strategy Papers (PRSPs – see Booth 2005 for a critical review), but knowledge generated by participatory means is not consistently incorporated in everyday policy and practice routines. Participatory methods may or may not be used at different stages in the policy or the project cycle and the data that they generate tends to have less traction in policy or project decision-making

processes (Eyben 2013). In short, participatory methods have not made sufficient headway in the transition to becoming an accepted, legitimate and routinised element of development policy and practice, although they have tended to have more traction within some NGOs in relation to needs assessment, planning and evaluation (Catley et al. 2008). Moreover, while participatory approaches have enabled the articulation of the types of things that matter for people in their own development, producing lists that are effectively criteria for wellbeing is only one part of what might be required in an assessment of what really matters for people. If we scan across from the treatment of wellbeing within participatory work in international development to a rich tradition of quality of life, work in health studies and sociology, then we can identify at least three different elements that must be further developed in order to make a human wellbeing approach relevant for policy and practice: identifying systematically what is important to people for them to live their lives well in their particular context (so addressing the critique of lack of rigour which is often levelled at participatory work); finding ways of assessing how well people are doing in their achievements of the things that they regard as important; and developing assessment methods that acknowledge and address wellbeing conflicts. These elements are discussed below.

The *first* element is to understand how to identify in a systematic way what is important to people for them to live their lives well in their particular geographical and societal context. The challenge here is to do this in a way that transcends the idiosyncrasy of numerous participatory studies of individual communities. This highlights the perennial challenge of finding a workable balance between universal and local approaches to understanding the world (see McGregor 2004). In other words, it is necessary to find ways of eliciting what it is that matters for people in particular social, economic, cultural and geographical contexts, but also to make these things comprehensible in terms of a universal framework of interpretation. As we will contend, a potential solution to this conundrum lies in a methodology that is founded in an iterative process between 'top-down' and 'bottom-up' interpretations of the world. This would involve a grounded adoption and adaptation of a universal framework that nevertheless enables us to take account of the context-specific forms of the 'things' that matter. As we explain, mixed methods have an essential role to play in this process of bringing together competing interpretations of what matters for a good life.

The *second* is to find ways of assessing how well people are doing in their achievements in respect of the things that they regard as important

for them to live well. For policy process purposes this is most often done using quantitative data as this enables communication at the population scale that is required; however, there will be some policy purposes for which a qualitative assessment of how well people are doing in their lives is particularly appropriate. The nature of the assessment that is required and the extent of its breadth and/or depth depend heavily on the purpose for which the assessment is being undertaken. While national level policy assessments will likely require large sample data covering a relatively limited range of variables, other more detailed policy purposes, such as the evaluation of the differential impacts of a policy intervention in particular community contexts, will require detailed and context-specific content. The sampling standards for such purposes would be shaped by the numbers of variables being explored and would have a strong orientation to revealing patterns of differentiation within a study population. Bearing this in mind, a key feature of any universal wellbeing assessment framework that is to be adopted is that it should be flexible so as to be adapted to different policy needs and be sensitive to differing levels and contexts in which the policy problem is being addressed. For example, qualitative methods could be used to develop and test a measure that captures what is important for wellbeing in particular contexts and for particular individuals, but could still be used comparatively across a larger sample (Camfield et al. 2008).

The *third* element is to develop assessment methods that acknowledge and address wellbeing conflicts, as opposed to developing measures that homogenise and obscure these (e.g. by using income as a proxy for everything that contributes to wellbeing): an approach to understanding development processes that focuses on the different human wellbeing aspirations and strategies of those involved and reveals the essentially contested and conflictual nature of development, making apparent the role of power (for empirical examples of this see McGregor et al. 2014). Conflicts are evident at the *intra-personal* and the *inter-personal* levels and each must be considered in relation to time – both now and in relation to the future.

At an *intra-personal* level, people are often required to make decisions about trade-offs between different elements of their own wellbeing. For example, in circumstances of impoverishment a person may have to trade-off their dignity and sense of self-worth by begging in order to eat or trading-off time with family and children (relational wellbeing) in order to keep a job (material wellbeing – economic security). Such intra-personal trade-offs can take place both in the present and over time: a person may choose a path of ‘deferred gratification’ where consumption

in the present is forgone in the hope of greater gratification at a later point in time.

At the *inter-personal* level, within any given community (whether a household, village or nation state) there are also different and competing notions of (and strategies for) wellbeing. For example, recipients of emergency cash transfers in Zimbabwe traded relational wellbeing for material wellbeing as they experienced resentment from non-recipients neighbours and exclusion from communal work activities (MacAuslan and Riemenschneider 2011). The struggle to find ways of living well and living together in these communities inevitably involves people trading their wellbeing off against the wellbeing of others in relationships that are implicitly or overtly political. The aspect of time in inter-personal trade-offs neatly reformulates the problem of sustainable development as was formulated in the Brundtland Report – where the wellbeing choices of some people in the present may adversely affect the wellbeing possibilities either of others elsewhere in this time period or of others sometime in the future (McGregor 2014). In such circumstances, governments are often called on to act in order to encourage people in the present to trade-off some of their wellbeing now for the sake of the wellbeing of people in the future (e.g. by reducing carbon-intensive energy consumption).

The issue of wellbeing conflicts at the *intra-personal* level refers to the challenges we each face of finding ways to live well as individuals, while conflicts at the *inter-personal* level imply ‘the governance problem’ – of finding ways and arrangements to live well together (Deneulin and McGregor 2010). Of course, the irreducibly social nature of human wellbeing means that trade-offs around intra-personal and inter-personal levels are inter-dependent and cannot be meaningfully separated except for abstract analytical purposes (as in maximisation calculus). As we noted above, public policy has a distinctive role in this challenge because policy choices then involve the authoritative allocation of resources to favour or support the wellbeing aspirations and strategies of some in society over that of others.

Given the complex and contested nature of wellbeing, mixed methods have a particular role to play in giving insight into the societal processes that support some notions of wellbeing over others (Corsin-Jimenez 2008, Davey and Selin 2012). These processes are the main subject of interest for policy analysis and although some insights may be gained through having better quantitative data and analysis of that, the real life and dynamic processes with which policy and practice must engage can only be revealed substantively through qualitative analysis. This is the

argument for mixing methods as the perceived shortcomings of qualitative data – for example, generalisability, idiosyncrasy – can be addressed by drawing on good practice in quantitative work without losing their distinctive insights.

Objective, subjective and inter-subjective

A possible source of confusion for the progress of wellbeing thinking and the development of metrics is the wide range of different frameworks that are out there. While many of these are broadly similar they nevertheless make different claims about what things matter, how we should organise our thinking about them and what data is needed to understand and assess human wellbeing. The differences in the kind of data that the frameworks call for is an important distinguishing feature and deserves careful consideration, but before we do this it is worth highlighting a basic source of confusion in the current debate. While the mixed and sometimes interchangeable use of the terms ‘dimension’ and ‘domain’ may look like a trivial semantic problem, it is a cause of miscommunication that has the potential to be an obstacle to the coherent development of wellbeing measures.

In the quality of life literature, arising out of the fields of health, social psychology and sociology, the term ‘dimension’ is used to distinguish between objective and subjective aspects of life: thus it refers to an ‘objective dimension of quality of life’ and a ‘subjective dimension of quality of life’ (see Axford et al. 2014). This is not the same sense in which the term dimension is used in the multidimensional poverty and international development literatures. In that literature the term is used to encompass dimensions of poverty beyond income or consumption and is therefore closer to the terminology of ‘domains’ or ‘domains of life’ (Rojas 2004) used in quality of life approaches.

We draw a distinction between indices of multidimensional wellbeing and of multidimensional poverty, for example, the Multidimensional Poverty Index (MPI) developed by Alkire and Foster, as the former deals with only a limited range of things that might be considered to represent poverty or refer to the human conditions which reproduce poverty. As we will discuss later, some of these other dimensions can be regarded as ‘objective’ aspects of poverty (as in the UN’s human development approach and MPI) while others can be considered more subjective (as is illustrated by the ‘Voices of the Poor’ category dealing with peace of mind, harmony, etc.). Others still are social in character (e.g. social exclusion and security). An important difficulty with the use of the term

'dimension' in 'multidimensional' approaches to wellbeing, poverty or development is that these mix up different categories of things in ways that can cause confusion for empirical study.

There appear to be two main explanations for the different numbers of dimensions or domains. The first is that each of these frameworks has been developed from a different intellectual tradition and second is that they have been developed with different purposes and audiences in mind. If, for example, we take the frameworks that arise out of health or social psychology such as the World Health Organisation Quality of Life (WHOQOL) or the SF-36 it is not surprising that they are relatively poorly developed in terms of their interest in and appreciation of social and economic factors that affect a person's overall wellbeing. These frameworks have been mainly interested in what is happening inside people's heads either to assess their mental health or to decide on what is an appropriate treatment for a person. In this respect they are able to limit their frameworks to identifying 'domains' within the dimension of subjective wellbeing that are most germane to health.

Those developing multidimensional poverty frameworks are primarily interested in providing a view of poverty that is expanded beyond a narrow income or consumption measure and their intended audience is a set of poverty policymakers who are profoundly wedded to objective indicators. It is not a surprise, therefore, that their framework would be limited both in terms of what they want to include in it and what measures they want to use to provide an analysis of the situation, in part due to problems of data availability (Alkire 2008). There has been considerable neglect until recently in economics and other disciplines of psychological or cognitive issues and physical health has been more a matter of concern than mental health. At the conceptual level these frameworks have been constrained by their underlying ontology (see Kanbur and Schaffer 2007 on the role of 'inter-subjective observability'³), but also by the hegemony of the idea of 'parsimony' (unwillingness to accept notions of complexity) and also by the limits of their mathematical manipulation (too many variables produce intractable mathematical problems).

These examples also indicate that there is a data dimension to this confusion between dimensions and domains, and this arises from the different academic traditions that are contributing to the current debates. Broadly speaking economics and most public policy analysis traditions (e.g. social policy and development studies) have been strongly invested in working with what they regard as objective data⁴ – data that (notionally) can be objectively verified and are understood to be a measure of

some reality in the world. In psychology, sociology and in other social science disciplines there has been longstanding acceptance of subjective data – data that arises from the views and experiences of the respondents of the study which can be quantitative or qualitative. However, as we develop approaches to generate a more comprehensive view of human wellbeing, it becomes apparent that a combination of objective and subjective data is required. Some things that are important as manifestations or indicators of wellbeing or are important to achieve wellbeing can be comprehended objectively; others can be comprehended both in objective and subjective terms, while some others can only be comprehended subjectively.

As an illustration, if we take Martha Nussbaum's (2000) 'list' of dimensions of a good life then clearly we can comprehend many important aspects of a person's state of 'bodily health' in objective terms, but for many purposes it will also be important to take into account the person's own subjective assessment of their 'bodily health'. However, when we turn to another of her dimensions 'emotions', unless we move into objective measurement by neuroscience or bio-chemical assessment, this is mainly comprehended in subjective terms. Indeed, Nussbaum's list illustrates a third dimension of data that is often overlooked in most debates over the role of objective and subjective data in the assessment of 'multidimensional' human wellbeing; these are those things that can only be comprehended inter-subjectively and therefore lend themselves particularly well to a mixed methods approach.

There are some things that many would agree are important for human wellbeing that cannot be well understood in either objective or subjective terms, or, as we have argued earlier, through quantitative or qualitative data alone. A strong argument for mixed methods is presented by the existence of inter-subjective phenomena, which are items or characteristics that only exist in terms of a socially accepted, agreed understanding between people. For example, Nussbaum's dimension of 'affiliation' provides a good illustration of this in as much as it deals with issues of relationality and identity – it is agreed that you are someone's friend or that you are an accepted member of an identity group. It is at this point that researchers start to mix methods as more qualitative and open-ended methods can capture the quality of relationships and the recursive nature of identity while quantitative methods such as social network analysis can show the extent and density of people's networks. In the following sections we discuss the characteristics of potential wellbeing frameworks and how these might then be adapted using a mix of methods.

Possible wellbeing frameworks

In 2012, in a major international conference to evaluate how the global community has been responding to the Stiglitz-Sen-Fitoussie challenge, Martine Durand, the Chief Statistician of the OECD, felt compelled to state that

[w]e are witnessing a convergence in our understanding of well-being with a common core set of well-being dimensions, and national priorities reflected in more specific domains and measures. (Plenary Address, 4th OECD Global Forum, New Delhi 2012)

Perhaps the first and main point of convergence has been an acceptance that human wellbeing must be understood as ‘multidimensional’ and cannot be captured adequately by considering a single measure such as happiness (e.g. Bellagio Initiative 2012b). This runs counter to the work of the happiness economists such as Richard Layard, Andrew Oswald and Bruno Frey, and Alois Stutzer whose arguments suggest that a single happiness score contains enough information for it to be relied on in policy decision-making.⁵ This work harks back to the days of Jeremy Bentham’s balance of pain and pleasure and implies that it is not necessary to know the detail of what lies below a person’s feeling of wellbeing or happiness and that a single happiness or utility score can adequately capture the underlying complexity, contradiction and fluidity of human judgements about the quality of their lives across its whole range. This is not a widely accepted view when considering the possibility of the detailed use of wellbeing measures in particular policy spheres. Although a single overarching concept and measure of subjective wellbeing may have some uses as a broad brush comparator or as a variable to slot into regression equations, it cannot provide the detail or richness of information required for most policy purposes (see Alexandrova 2005).

A second fundamental point in the emerging consensus is that it is necessary to take account of both objective and subjective aspects of wellbeing in some kind of integrated framework⁶ although there may be technical challenges related to aggregation. That said, although it is widely agreed that wellbeing is multidimensional, we quickly move into the more perplexing discussion about which dimensions matter and how we define these. This discussion hinges around which (or whose) framework is to be adopted. Currently there are a large and growing numbers of different multidimensional frameworks and each offers

slightly different lists of what are proposed to be the set (or list) of universal dimensions.⁷

Put simply, there are two ways in which these lists have been established: one is in a 'top-down' manner in which the particular ingredients of what is required for wellbeing are identified predominantly from a particular philosophical position, conceptual framework or ideology; the other is in a 'bottom-up' way, in which the 'list' of what is required for wellbeing is built up from observation or engagement with the people whose wellbeing you are concerned to understand.⁸ The desire for universality has tended to favour the 'lists' produced by top-down approaches and most of the multidimensional frameworks on offer arise from particular theoretical and philosophical positions.

Acknowledging the tension and the politics that lie between these two approaches, Sen had declined to propose such a list, arguing that any such list should arise from a process of public reasoning within a particular polity. However, many of his fellow academics have been less hesitant. Nussbaum, for example, argues that there are 10 universal capabilities (what she calls 'spheres of existence').⁹ The Human Development Index has provided another multidimensional list based on Sen's capabilities but like the MPI this is not a multidimensional 'wellbeing' measure as there are only three 'dimensions' (health, education and standard of living), and the rationale for these three is drawn from the human development and capabilities framework. In the MPI these three dimensions are then assessed using 11 objective indicators.

The 'Voices of the Poor' study conducted by the World Bank provides an interesting mid-point between narrow multidimensional poverty frameworks and broader efforts to assess wellbeing (Narayan et al. 2000). In a study that took evidence from over 60,000 poor people in many countries of the world, they confirmed that people viewed and experienced poverty as multidimensional. From this relatively 'bottom-up' perspective they assert that there are eight dimensions of experienced poverty: (1) material wellbeing (i.e. having enough food, assets and work); (2) bodily wellbeing (health, appearances and physical environment); (3) social wellbeing (being able to care for, bring up, marry and settle children); (4) self-respect and dignity; (5) peace, harmony and good relations in the family/community; (6) security (civil peace, a physically safe and secure environment, personal physical security, security in old age and confidence in the future); (7) freedom of choice and action and (8) psychological wellbeing (peace of mind, happiness, harmony including a spiritual life and religious observance).

These dimensions are reflected in the many wellbeing frameworks arising out of social psychology. For example, Bob Cummins (1998), who has spearheaded one of the most globally prominent subjective wellbeing frameworks, proposes seven domains that contribute to quality of life (material wellbeing, health, productivity, intimacy, safety, community and emotional wellbeing – spiritual wellbeing was added later). In this approach these are all assessed subjectively. Other prominent frameworks arising out of social psychology or health such as the WHOQOL assessment tool or the SF-36 from the Medical Outcomes Study have between four and eight domains.

Coming from a tradition in social policy that focuses on human needs, Doyal and Gough (1991) suggest that there are only two fundamental and universal human needs and these are health and autonomy. However, in an elaborate schema, these two basic needs are met in different social contexts through achievements in respect of 11 universal intermediate needs: adequate nutritional food and water, adequate protective housing, non-hazardous work and physical environments, appropriate health care, security in childhood, significant primary relationships, physical and economic security, safe birth control and child-bearing and appropriate basic and cross-cultural education (Doyal and Gough 1991).

The Stiglitz-Sen-Fitoussie Commission concluded that a 'multi-dimensional' approach to measuring wellbeing was required and also taking a top-down approach they proposed that there were eight dimensions that needed to be taken into account (material living standards, health, education, personal activities, political voice, personal relationship, environment and physical and economic security). Following on from this the OECD's 'How's Life' Framework (2011),¹⁰ which is an element of their Better Life Initiative and is a direct descendent of the Stiglitz Commission formulation, expands the number of dimensions to 11. As is shown in Figure 10.1, these 11 'dimensions' are divided between two 'pillars' that are labelled: 'Material Conditions' and 'Quality of Life'. The quality of life pillar mainly contains human development dimensions but draws on other frameworks and includes a dimension labelled 'Subjective well-being', despite also using subjective data to assess the material and human wellbeing dimensions. A third and important element of the 'How's Life' Framework is labelled 'sustainability' and this elaborates the context for material conditions and quality of life in that it identifies four types of 'capitals' that are seen as providing the enabling relationships through which present wellbeing can be translated (or not) into future wellbeing outcomes. This third element provides the dynamic

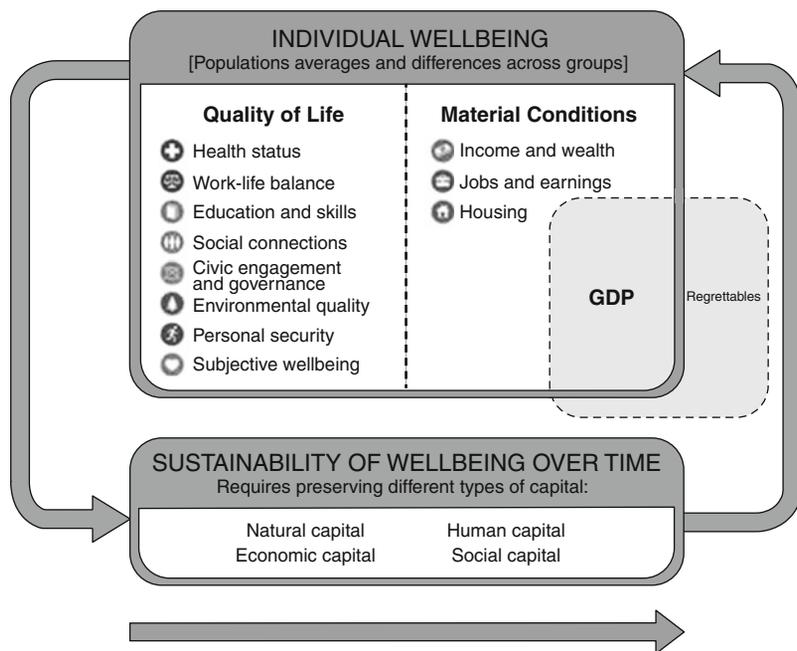


Figure 10.1 The OECD wellbeing framework: better life initiative

that is necessarily part of any understanding of how human wellbeing is produced and how it is distributed across populations in a society.

Providing structure for assessing human wellbeing in development policy and practice

The purpose of frameworks such as the ‘How’s Life’ Framework developed by the OECD is much grander, more ambitious and more complex than the frameworks that arise out of health and social psychology or those that are concerned only with multidimensional poverty. Since the purpose of such a framework is to measure development or societal progress in terms of whether it is improving human wellbeing, it cannot just be confined to what is happening to the poor (albeit that providing an insightful analysis of poverty dynamics and distributions can be one important product of such an approach) nor can it be interested only in measuring wellbeing in terms of what is happening inside peoples’ heads; it is also interested in what is happening to their bodies and their

social relationships. It is this combination of considerations of the material aspects of wellbeing, the assessments of perceived quality of life and the relationships that give these their dynamic, that suggest a distinctive and three-dimensional framework.

When we consider the many other 'multidimensional' frameworks currently being offered (discussed in the preceding section) we find that most operate with some combination of three dimensions: a dimension that refers primarily to the material conditions of life; a dimension that refers to people's broader quality of life, which includes human development components such as education and health, but also extends to 'conditions of being' like security and subjective wellbeing; and a third dimension that can be understood as a relational dimension, which refers to the relationships that a person must be able to enter into in order to continue meeting the needs that are important for their wellbeing (e.g. the relationships in society that allow them to act with autonomy or the relationships amongst friends and family that provide care).

There are different variants of this conceptual cluster and it is tempting to try to make the variants map on to each other. For example, we could see these three dimensions as mapping onto the types of data that might be used to assess how well a person is doing in their life. The material dimension is easily thought of as being assessed primarily by objective data (e.g. how much income a person has) but as we have noted when one looks deeper into this dimension we see that there are different issues about the 'qualities' of this income that are not readily assessed solely by objective data. From more subjective and relational perspectives it is important to consider whether the person regards this as a satisfactory level of income or whether it is a 'reliable' and 'predictable' income since this will provide an insight into the extent to which the person might be motivated by opportunities to generate more income. Perceived reliability could be explored quantitatively, for example, by asking people to report how reliable they think their income is on a scale of one to five. However, in order to understand the response we would need an understanding of prevailing values and norms in the societal context, which could be explored through ethnographic research. In fact, in order to develop this scale it would be beneficial to know what the word reliability means in this context and how it is likely to be responded to, which could be done during an initial pilot, possibly using cognitive interviewing techniques. The 'reliability' or 'predictability' of the income will also depend on the quality of relationships that the person has with the source of that income. If,

for example, the person has a secure and legally protected employment contract that establishes clear rights to a regular and predictable income then that must be considered an important aspect of their material conditions. However, in most developing countries, the majority of people do not have a contract for income that has such a secure relational quality and in such circumstances the unreliable and unpredictable quality of income becomes a key motivational factor for much of the behaviour (economic, political and social) of the person and their household. These types of insights have been accessed through mixed methods evaluations of development interventions or grants which have demonstrated conclusively that the process by which a 'good' is delivered is as important in shaping its outcome as the good itself (e.g. see Molyneux 2006, Adato 2007 and Roelen and Devereux 2014 in the context of cash transfer schemes). Nonetheless, combining objective with subjective data can be challenging, due to the contradictions that can occur between the two data types. For example, asking people whether they are satisfied that their amount of food is sufficient may give a very different picture to objective assessments of their actual state of nourishment, as Sen (1981) depicts in his account of 'starving widows' during the West Bengal famine who subjectively described themselves as faring well. More recent work on the development of the Human Kind Index in Scotland reported a high level of self-reported satisfaction with health (88% responded that their [self-assessed] health was good or very good in the 2007/2008 survey and 93% did so in 2009/2010), despite Scotland's well-known health problems and health inequalities (Dunlop et al. 2012). The presence of adaptive preferences that these examples suggest is one of the strongest arguments for combining objective with subjective assessments, and alerts us to the dangers of relying on subjective assessment alone.

If we return to the perspective of the person whose wellbeing we are trying to assess then we can remind ourselves that in order to achieve wellbeing in any particular societal context, people need *to have* certain things, they need to be able *to do* certain things and these then combine to allow and/or deny the person the possibility *to be* who they aspire to be.¹¹ In order to assess how well people are achieving this kind of rounded and social notion of wellbeing it is necessary to use objective, subjective and inter-subjective data to establish whether they have the things that they need (e.g. objectively enough clothes to protect against physical harm from the weather) and also that they have the things that they believe they need (e.g. subjectively, the clothes that will enable them to go out in society without shame). Equally important is whether

they are able to do the things they need to do (e.g. in inter-subjective terms, are they socially excluded), and whether as a result of this they are able to be the person that they want to be (e.g. is their self-esteem and sense of autonomy low).¹² Many of these aspects of life are culturally specific, for example, autonomy, and are not always successfully interpreted through universal measures, which suggests that an important role for qualitative research is bringing universal and local interpretations into dialogue with each other.

A Wellbeing-Focused Evaluation (WFE) of social protection: Sequencing and mixing qualitative and quantitative methods

The final part of the chapter briefly describes how a mixed methodology built around a three-dimensional wellbeing framework has been implemented in an evaluation of a social protection programme in Zambia. Social protection programmes and policies are forms of intervention that provide poor and vulnerable people with inputs of some kind (usually money or food, but also advice, education, medicines, nutritional supplements, empowerment, etc.) to enable them to meet their needs and to participate better in society and in the development process. In the case of Zambia, the government has been assisted in its social protection policy development by international aid donors and has experimented with a range of different social protection models (Harland 2014). As part of its move towards creating a national social protection policy and a unified programme, the Government of Zambia (GoZ) commissioned a series of evaluation studies in order to gather learning from its existing programmes. While much of the learning is taking place through the form of large-scale randomised control trial studies (American Institute of Research – AIR), the GoZ alongside UNICEF also commissioned a study of the ‘wider impacts’ of two of the Cash Transfer social protection schemes (the Child Grant scheme and the ‘10% inclusive’ scheme).

The overall goal of the Zambian social protection programme was to reduce extreme poverty and the inter-generational transfer of poverty. The ‘10% inclusive’ cash transfer programme has been implemented since around 2004 in four districts in Zambia: three in Southern Province (Kalomo, Kazungula and Monze) and one in Eastern Province (Chipata). The 10% inclusive scheme is so-called because it targets the poorest 10% of households in the communities and seeks to address the poverty that is believed to be causing them to be excluded from community life. It

focuses particularly on households that are destitute or incapacitated: by age, by disability, by being single headed households or by virtue of them being households caring for orphans.¹³

The Child Grant Cash Transfer Programme (CGP) has been implemented since 2010 in three districts of the country which have amongst the highest rates of stunting, wasting, mortality and morbidity among children under five years old in the country: Kaputa in Northern Province and Kalabo and Shang'ombo in Western Province. The CGP has a universalist targeting strategy, going to any household with a child under five years old with the particular intention of breaking the inter-generational transmission of poverty.

In response to a call to understand 'wider impacts' of social protection, the commissioned study proposed a framework for evaluation that focuses on human wellbeing. Using the type of three-dimensional framework described above, this evaluation study has explored impacts and effects in terms of: material conditions (the extent to which the transfer is enabling material needs to be met); quality of life (with a particular focus on whether people receiving the transfer perceive themselves as experiencing a better quality of life); and relational wellbeing (whether the recipients are better able to do things that are important to them). Accordingly, the study has defined 'wider impacts' in two ways: the effects of the social protection schemes on the wellbeing of recipients and their households beyond the immediate material impacts; and the positive and negative effects that the cash transfer schemes have had on social, economic and political relationships in recipient communities. In order to remain true to the idea of assessing the impact on the wellbeing of the recipients in the three-dimensional way explained earlier, the studies used a sequenced combination of qualitative and quantitative research methods. The purpose of the sequencing was to ground the assessment of wellbeing in the things that mattered to people themselves: for a good life in the different areas; to deepen the insights into the range and interactions of impacts in the material, relational and subjective dimensions; and to produce robust quantitative data on the different but interacting forms of objective, subjective and inter-subjective impact (Figure 10.2).

This framework is dynamic and takes account of the cumulative and compounding effects of the cash transfer intervention (feedback loops).¹⁴ These include consideration of relationship changes that take place as a result of the material, relational and subjective impacts of the programmes (e.g. positively where recipients perceive themselves as more empowered, or negatively where cash transfer receipt is seen in the community as stigmatising). This approach enables the study to

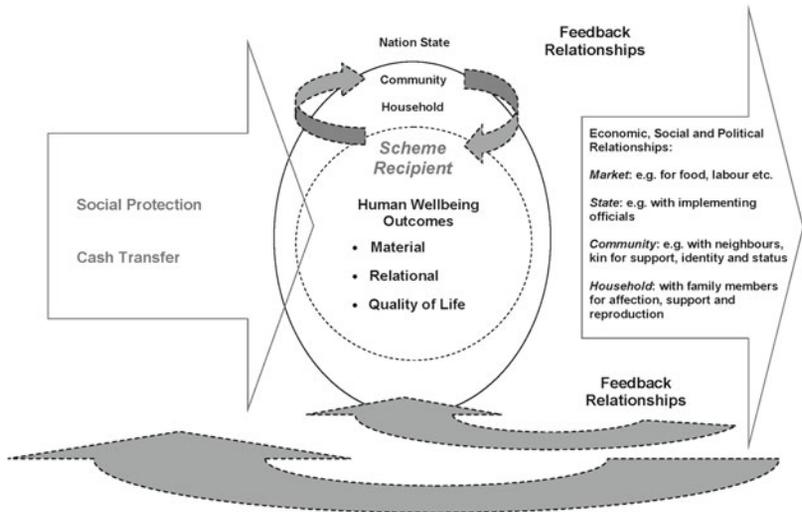


Figure 10.2 A wellbeing evaluation framework for social protection

Source: OECD, 2013

identify unforeseen or unintended consequences of Social Cash Transfer (SCT) schemes that are often not readily in the purview of narrower evaluations.

The framework directs the exploration of the impacts and effects of the SCT programme to three key levels:

- (i) The impacts and outcomes that arise at the individual and household levels from receipt of the cash transfer.
- (ii) The broader outcomes and impacts that arise in the community as a result of the community being included in the scheme.
- (iii) The direct and indirect experiences that recipients and non-recipients have of the programme processes.

The fieldwork for all the studies was carried out in two stages. Six communities were selected for study in the two different scheme areas: five of these were recipient communities and one was a non-recipient community in the same or an adjacent district, included as a point of reference. The first fieldwork visit involved mainly qualitative methods, involving a combination of participatory methods, focus group discussions and semi-structured interviews. These were used to establish relationships

of trust and understanding to enable further study in the community, to carry out some basic scoping of key characteristics of the community and to undertake the grounding work for the design of the research survey instrument to be used in the second fieldwork stage.

As a first step in operationalising the wellbeing framework, there are two clear outputs from the first stage of the study. The first was a 'community profile' – this is a descriptive report on the community context in which the recipients are living and receiving the cash transfer. It includes a description of basic community demographics, an insight into the distributions of key resources and a preliminary identification of the key social, economic and political relationships at work in the community that pertain to poverty and vulnerability. The community profile provides both practical and analytical information for the evaluation: on the practical side it provides a better picture from which to make decisions about sampling for the survey instrument implementation, while on the analytical side it provides a first exploration of the relationships and structures within which cash transfer recipients are located. The second output of the first stage of the study was a set of lists from each of the communities of the things that people regarded to be important for them to live their lives well in these particular communities. This information was gathered through a combination of Focus Groups Discussions (FGDs) and semi-structured interviews. FGDs were organised in each community for male and female recipients and also for non-recipients. The objective of both the FGDs and the interviews is to have people generate from amongst themselves lists of the things that a person must *have*, be able to *do* or be able to *be* in that community in order to be considered to be living well. The information generated by these methods is then collated from across the study communities and forms the basis for the design of a survey instrument which is then systematically applied back to a sample of recipients and non-recipients in the second stage of the study.

The second stage of the fieldwork took place some six months later. This second fieldwork visit focused mainly on the application of the survey instrument that had been designed in the intervening period. This survey instrument consisted of a collated list of the 'items' that people said were important to *have*, to *do* and to *be* in order to live well in the community. It asked a mixed sample of recipients and non-recipients in the recipient communities and a set of non-recipients in the non-recipient community about their objective circumstances in relation to these items and their subjective view of the importance and level of satisfaction in relation to their achievement in respect of these items. The

survey instrument was mounted on tablet computers and was implemented with the sample group by local research assistants. The application of the quantitative survey was paralleled by a small amount of further qualitative work. Key informant interviews and a small number of further FGDs were used to have people in the communities explain what they regarded to be possible pathways out of poverty for recipients in these communities. This involved describing the different possible ways that people could seek to improve their wellbeing in these communities. It involved descriptions of what things people had to have in order to progress and an explanation of the types of relationship they needed to be able to access in order to achieve these outcomes.

From an analysis of qualitative and quantitative data, both of the social protection schemes were found to have positive impacts on the objective and subjective wellbeing of recipients. There were also positive spill-over effects on the wellbeing of non-recipients in the recipient communities. These wellbeing impacts spanned all three of the dimensions of wellbeing discussed above: material, relational and subjective.

In objective terms, the study broadly affirmed the findings of the Randomised Control Trial and recipients had been able to spend more on meeting the needs that they prioritised. This included spending on food and other small purchases such as soap. It particularly included expenditures that enabled their children to stay in school. Recipients in the '10% Inclusive' scheme spent their cash transfer either on meeting these basic needs or hiring labour to work on their land. 'Child Grant' recipients were more likely to report investments in productive activities – both new and existing businesses/services. This finding was affirmed by the parallel pathways out of poverty focus groups. These showed that for Child Grant recipients the pathway focused more on being able to make progress in agriculture or business, while for '10% inclusive' recipients these pathways were seen as less feasible and that their aim was to stay out of destitution. These results appear intuitive since the incapacity of 10% recipients usually constrains their ability to engage directly in income-generating activities and the receipt of a cash transfer does not change these constraints.

In terms of impacts on subjective assessments of wellbeing, both schemes were perceived to have reduced the dependency of recipients on others. This was widely held to be a positive development and important for the psychological wellbeing of recipients. The overall quality of life scores for recipients are broadly similar across both schemes and for all sites and were higher than for non-recipients of a similar economic status.

In terms of relational wellbeing impacts, a key impact in both schemes is that recipients perceive they have an improved ability to participate in their community and be respected. In all sites there was widespread reporting that begging had reduced after the introduction of the schemes. One of the main concerns voiced about such social protection schemes was that they undermine or displace traditional and community-based systems of support – in effect undermining existing social capital. The analysis of the relational impacts of the cash transfers in these studies suggests that this has not been the case. Although some aspects of support systems may become monetised (e.g. the incapacitated hiring labour to work on their land rather than wait for that to be supplied voluntarily), the general sense was that monetisation was not a major issue. The 10% recipients in particular reported that they still needed help from more able-bodied family, friends and community members but that this continued to be given. At the same time many of the recipients in the 10% scheme perceived themselves, and were perceived by others, to be less of a burden on these informal systems of support. The 10% recipients in particular seemed to welcome the greater autonomy that this provided and in both schemes poorer recipients were proud to be able to report that they were even at times able to assist others. This brief insight gives an example of what a methodology for a Wellbeing Focused Evaluation (WFE) might look like. The findings that are revealed by the methodology offer a broader insight into the effects of development intervention than are usually offered by quantitatively dominated evaluation studies. These broader insights have important implications for how future social protection programmes in Zambia might be designed and implemented to effectively contribute to better human wellbeing outcomes.

Conclusion

We began this chapter by recognising that there is an ongoing shift towards assessing development and societal progress in terms of human wellbeing. The chapter explains some convergences around the kind of conceptualisation of wellbeing that is being adopted by national and international initiatives. This has involved recognition of the need for a hybrid conceptualisation of human wellbeing that takes account of both objective and subjective wellbeing. While many of these initiatives are taking place at the macro-scale of measuring wellbeing in the nation state and in relation to national development policies and trajectories, this chapter is concerned that these macro-level developments do not

lose touch with the understanding of wellbeing of people at the micro-level of the community and of the development project or programme. We argue that although the specific needs of measurement may be different at the two levels, the conceptual framework that underpins the development of measures and methodology should be the same.

In order to establish whether this might be possible we review a range of different wellbeing frameworks that are currently being used, explaining that most of these have been created from a top-down perspective, with only a few being informed by bottom-up processes. However, when we look at the different terminologies and different numbers of domains or dimensions that they identify, we nevertheless find that essentially they all encompass three broad dimensions: a material wellbeing, a relational wellbeing and a subjective wellbeing. This accords with and affirms the three-dimensional framework for understanding and studying wellbeing that was developed through a thoroughly ground-up process in the ESRC funded Research Group on Wellbeing in Developing Countries (WeD, see Gough and McGregor 2007, McGregor and Sumner 2010). We note that in evaluating each of the different frameworks it is important to take into account the standpoint of the developers and users.

When considering the efficacy of using a wellbeing framework at the grass-roots level and in relation to understanding the effects of specific interventions, policies or projects, we suggest that there are three basic tasks that a wellbeing methodology must address. *First* to understand how to identify in a systematic way what is important to people for them to live their lives well in their particular geographical and societal context; *second* to find ways of assessing how well people are doing in their achievements in respect of the things that they regard as important for them to live well; and *third* to develop assessment methods that acknowledge and address wellbeing conflicts that arise in relation to an intervention or policy, and as opposed to developing broad or proxy measures that homogenise and obscure differences. The multidimensional nature of the wellbeing framework requires the generation of objective, subjective and inter-subjective data and as argued by authors in other chapters in the volume (e.g. Dawson, Torres Penagos and Bautista Hernández) for other multidimensional purposes, this points to the necessity of mixing qualitative and quantitative methods. Moreover, to accomplish these three basic tasks the application of the different types of methods must be appropriately sequenced.

For a wellbeing approach to work at the project, policy or intervention level it is necessary to find a workable balance between universal and

local approaches to understanding the world (see McGregor 2004). In other words, it is necessary to find ways of eliciting what it is that matters for the wellbeing of people in particular social, economic, cultural and geographical contexts, but also to make these things comprehensible in terms of a universal framework of interpretation. We contend that a solution to this conundrum is to be found in a methodology that builds an iterative process between 'top-down' and 'bottom-up' interpretations of the world. This involves using a sequence and mix of qualitative and quantitative methods to identify what things matter to people for their wellbeing in that particular context and then using that to adapt a universal framework around which quantitative and comparable data can be generated. The conceptual underpinnings and motivations behind the OECD How's Life Framework (OECD 2011) coincide to a large extent with the purposes that lie behind the desire to carry out wellbeing assessments at the frontline of development and as such this is the most likely candidate framework for adoption and adaptation.

As we explain, mixed methods have an essential role to play in this process of bringing together competing epistemologies that might differently interpret what matters for a good life. The mix also allows us to address the issues that are raised by the third category of data: the inter-subjective. Inter-subjective data has the characteristic that it only exists in terms of an agreed understanding between people (e.g. being someone's friend or a member of an identity group). It is an often neglected category of data but one which is essential for understanding wellbeing in specific societal and cultural contexts. While global wellbeing frameworks such as the OECD How's Life Framework can come up with labels that give meaning to the dimensions that must be taken into account to assess wellbeing universally, the grass-roots implementation of any wellbeing assessment must also explore local meanings. For example, a house in rural Bangladesh must have different qualities from a house in urban London if it is to be considered to be an adequate form of shelter. We argue that an important reason for mixing methods is that qualitative methods can capture something of the *quality* of the things and relationships that are important for wellbeing. Quantitative methods can explore the extent to which that quality is being achieved in either objective or subjective terms.

In the concluding part of this chapter we have presented an example of a WFE to show how we can practically assess whether an intervention has enabled people to *have* certain things, to *do* certain things and combine these havings and doings to *be* who they aspire to be. The insights drawn from the evaluation of the impact of social protection

on wellbeing and relationships in communities in the Zambia case study employs a methodology that reflects the three-dimensional wellbeing framework. It combines different methods in a sequenced way to generate the different types of information that when taken together afford us a rounded insight into the wellbeing effects of the cash transfer programmes. Perhaps such insights could have been generated using different methods and under the banner of a different methodology but seldom are these methods and methodologies brought together in a way that enables us to explore the impacts on a multidimensional notion of wellbeing. If we remind ourselves that the overarching purpose of development interventions is to improve human wellbeing, then it would seem that different types of interpretation of effects must be brought together by the combined use of qualitative and quantitative methods that generate three different categories of data (objective, subjective and inter-subjective). Such a methodology has the benefit of re-engaging people in the development process and allowing them to have a say in what matters for their wellbeing. It is only then that we will begin to understand in a more holistic way what makes for a good society in which human beings can survive and flourish.

Notes

- 1 Inter-subjective refers to those things that are collectively agreed to be 'true' or 'real'. It denotes those things or qualities of things whose reality is founded in agreement based in the subjective views of two or more people. A striking example of this would be the accountancy practice of placing a value on 'goodwill' when recording the assets of a business.
- 2 Following the logic of the Seers' argument, this chapter takes the view that all societies are developing societies.
- 3 Popper (1959, p. 103) argued that "a basic statement" (a statement of scientific fact) must also satisfy a material requirement . . . this event must be an "observable" event; that is to say, basic statements must be testable, intersubjectively, by "observation".
- 4 There are serious doubts as to whether much of the data that they regard as 'objective' can in fact be considered truly objective. For example, income data is notoriously difficult to collect accurately and the data that is presented on this is usually deduced from other statistics or is the result of self-report in survey. Neither can be regarded as an objective measure.
- 5 See, for example, Layard (2010) and Oswald et al. (2010).
- 6 For example, see the discussion paper for session 3.1 at the 2013 OECD Global forum on development (www.oecd.org/site/oecdgfd/) or one of the nine key policy briefing papers that informed the United Nations Conference on Sustainable Development (Rio+20) where it was agreed that 'Both objective and subjective factors are critical for the overall well-being of individuals'

- (RIO+20 POLICY BRIEF #6: Human well-being for a planet under pressure, http://www.planetunderpressure2012.net/pdf/policy_wellbeing.pdf).
- 7 For a review of the different domains of various prominent frameworks, see Samman (2007).
 - 8 See Alkire (2008), or Rodriguez-Takeuchi (2014), for a more sophisticated account of this process.
 - 9 Life, bodily health, bodily integrity, senses, imagination and thought, emotions, practical reason, affiliation, other species, play and control over one's environment.
 - 10 <http://www.oecd.org/statistics/howlife.htm>
 - 11 This definition arises from the work of the Wellbeing in Developing Countries ESRC research group described in the introductory section (see also Gough and McGregor 2007) and draws on the work of Sen, among others.
 - 12 Work by Carol Graham (e.g. 2010) is particularly illustrative of situations where objective and subjective assessments diverge. She identifies populations that objectively have enough 'stuff' but feel dissatisfied with what they have because of their aspirations, and others who are satisfied with life even though they objectively have much less.
 - 13 According to the Harmonised Manual of Operations for the Zambia Social Protection programmes: "Destitute" means that the household struggles to survive, adopts negative coping mechanism, has less than 3 meals a day, indecent shelter and clothing, limited access to education and health and only irregular and insufficient support. "Incapacitated" means that the household has either no household members who are fit for work and of working age or that there is a very high dependency ratio (at least 3 unfit members for every fit member)' (MCDMCH 2013, p. 6). An unfit household member is defined as a person who is younger than 19; older than 64 or; 19–64 and chronically sick, or still going to school (MCDMCH, 2013, p. iv).
 - 14 For a related framework that draws on a similar view of wider impacts and those acting over time and introduces feedback loops, see Devereux et al. (2013).

References

- Adato, M. (2007). Combining Survey and Ethnographic Methods to Evaluate Conditional Cash Transfer Programs. *Q-Squared Working Paper No. 40*. CA: University of Trent.
- Alexandrova, A. (2005). Subjective Well-Being and Kahneman's 'Objective Happiness'. *Journal of Happiness Studies*, 6, 301–324.
- Alkire, S. (2008). Choosing Dimensions: The Capability Approach and Multidimensional Poverty. In Kakwani, N. and Silber, J. (eds), *The Many Dimensions of Poverty*. New York: Palgrave Macmillan, pp. 89–119.
- Alkire, S. and Foster, J. E. (2011). Counting and Multidimensional Poverty Measurement. *Journal of Public Economics*, 95(7–8), 476–487.
- Axford, N., Jodrell, D., and Hobbs, T. (2014). Objective or Subjective Well-Being? In A. Ben-Arieh et al. (eds), *Handbook of Child Well-Being*. Netherlands: Springer, pp. 2699–2738.

- Bebbington, A., Hinojosa-Valencia, L., Munoz, D., and Rojas-Lizarazu, R. E. (2007). Livelihoods and Resource Accessing in the Andes: *Desencuentros in Theory and Practice*. In Gough, I. and McGregor, J. A. (eds), *Wellbeing in Developing Countries: From Theory to Research*, pp. 176–198. Cambridge: Cambridge University Press.
- Bellagio Initiative. (2012a). *Synthesis Report: The Future of Philanthropy and Development in the Pursuit of Human Wellbeing*. IDS-Resource Alliance-Rockefeller Foundation. <http://www.bellagioinitiative.org/resource-section/bellagio-outputs/#final>
- Bellagio Initiative. (2012b). *Final Report: Human Wellbeing in the 21st Century – Meeting Challenges, Seizing Opportunities*. Institute of Development Studies, Rockefeller Foundation. http://www.bellagioinitiative.org/wp-content/uploads/2012/09/BELLAGIO_WELLBEING_SPREADS.pdf
- Booth, D. (2005). Missing Links in the Politics of Development: Learning from the PRSP Experiment. *Working Paper 256*. London: ODI.
- Camfield, L., Crivello, G. and Woodhead, M. (2008). Well-Being Research in Developing Countries: Reviewing the Role of Qualitative Methods. *Social Indicators Research*, 90(1), 5–31.
- Camfield, L. and McGregor, J. A. (2009). Editor's Introduction: Quality of Life and International Development Policy and Practice. *Applied Research in Quality of Life*, 4(2), 128–134.
- Catley, A., Burns, J., Abebe, D. and Suji, O. (2008). *Participatory Impact Assessment: A Guide for Practitioners*. Boston, MA: Feinstein International Center, Tufts University.
- Chambers, R. (1997). *Whose Reality Counts? Putting the Last First*. London: Intermediate Technology Publications.
- Chang, H.-J. (2003). *Kicking Away the Ladder. Development Strategy in Historical Perspective*. London: Anthem Press.
- Coase, R. (1998). The New Institutional Economics. *American Economic Review*, 88(2), 72–74.
- Corsin-Jimenez, A. (ed.) (2008). *Culture and Well-Being: Anthropological Approaches to Freedom and Political Ethics*. London: Pluto Press.
- Cummins, R. (1998). The Second Approximation to an International Standard for Life Satisfaction. *Social Indicators Research*, 43, 307–334.
- Davey, G. and Selin, H. (2012). *Happiness Across Cultures: Views of Happiness and Quality of Life in Non-Western Cultures*. Springer: Dordrecht.
- Deneulin, S. and McGregor, J. A. (2010). The Capability Approach and the Politics of a Social Conception of Wellbeing. *European Journal of Social Theory*, 13(4), 501–519.
- Devereux, S., Roelen, K., Béné, C., Chopra, D., Leavy, J. and McGregor, A. (2013). *Evaluating Outside the Box: An Alternative Framework for Analysing Social Protection Programmes*. IDS Working Paper 431/CSP Working paper 010. Brighton: Institute of Development Studies and Centre for Social Protection.
- Doyal, L. and Gough, I. (1991). *A Theory of Human Need*. London: Macmillan.
- Dunlop, S., Swales, K., Trebeck, Glinski, J. (2012). *The Oxfam Humankind Index for Scotland*. Oxfam Research Paper, April 2012, First results. <http://www.oxfam.org.uk/scotland/blog/2012/04/--/media/6A6B095DB10E432A88DEBCA5C9F0F365.ashx>

- Eyben, R. (2013). Uncovering the Politics of 'evidence' and 'results'. A Framing Paper for Development Practitioners. Downloaded 3 January, 2015, <http://big-pushforward.net/wp-content/uploads/2011/01/Uncovering-the-Politics-of-Evidence-and-Results-by-Rosalind-Eyben.pdf>
- Gough, I. and McGregor, J. A. (eds) (2007). *Wellbeing in Developing Countries: From Theory to Research*. Cambridge: Cambridge University Press.
- Graham, C. (2010). *Happiness Around the World: The Paradox of Happy Peasants and Miserable Millionaires*. Oxford, UK: OUP.
- Harland, C. (2014). Can the Expansion of Social Protection Bring about Social Transformation in African Countries? The Case of Zambia. *European Journal of Development Research*, 26, 370–386.
- Kanbur, R. and Schaffer, P. (2007). Epistemology, Normative Theory and Poverty Analysis: Implications for Q-Squared in Practice. *World Development*, 35(2), 183–196.
- Layard, R. (2010). Measuring Subjective Wellbeing. *Science*, 372(576), 534–535.
- MacAuslan, I. and Riemenschneider, N. (2011). Richer but Resented: What Do Cash Transfers Do to Social Relations?. *IDS Bulletin*, 42(6), 60–66.
- MCDMCH. (2013). The Wider Impacts of Social Protection: Report on the 10% Inclusive SCT: Kalomo, Monze and Choma. Final Draft Report by the Institute of Development Studies (IDS) in association with RuralNet Associates Limited; the Platform for Social Protection, Department for Development Studies, University of Zambia. Ministry of Community Development, Mother and Child Health, Government of Zambia. Lusaka, Zambia.
- McGregor, J. A. (2004). Researching Well-Being: Communicating between the Needs of Policy Makers and the Needs of People. *Global Social Policy*, 4(3), 337–358.
- McGregor, J. A. (2008). Wellbeing, Poverty and Conflict, WeD briefing paper January 2008. Downloaded 20 May 2015, <http://www.welldev.org.uk/research/bp/bp1-08.pdf>
- McGregor, J. A. (2014). Poverty, Wellbeing, and Sustainability. In Atkinson, G., Dietz, S., Neumayer, E. (eds), *Handbook of Sustainable Development*. Cheltenham: Edward Elgar.
- McGregor, J. A., Salagrama, V., and Bahadur, A. V. (2014). Fisheries in Coastal India: Extraction, Livelihoods and a Way of Life. In Gilberthorpe, E. and Hilson, G. (eds), *Natural Resource Extraction and Indigenous Livelihoods Development Challenges in an Era of Globalization*. Surrey, UK: Ashgate, chapter 5, p. 93.
- McGregor, J. A. and Sumner, A. (2010). Beyond Business as Usual: What Might 3-D Wellbeing Contribute to MDG Momentum. *IDS Bulletin*, 41(1), 104–112.
- Melamed, C. (2011). *Does Development Give Poor People What They Want?* Background note, ODI. <http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/6606.pdf>
- Merry, S. E. (2011). Measuring the World: Indicators, Human Rights, and Global Governance. *Current Anthropology*, 11(S3): S83–95.
- Molyneux, M. (2006). Mothers at the Service of the New Poverty Agenda: Progres/Oportunidades, Mexico's Conditional Transfer Programme. *Social Policy and Administration*, 40(4), 425–449.
- Narayan, D., Patel, R., Schafft, K., Rademacher, A. and Koch-Schulte, S. (2000). *Voices of the Poor: Can Anyone Hear Us?*. Oxford: Oxford University Press.

- Nussbaum, M. (2000). *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.
- OECD. (2011). *How's Life? Measuring Well-Being*. Paris: OECD Publishing. <http://dx.doi.org/10.1787/9789264121164-en>
- ONS. (2011). *Measuring what Matters: National Statistician's Reflections on the National Debate on Measuring National Well-Being*. London: ONS.
- Oswald, A. et al. (2010). Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the USA. *Science*, 327(576), 576–579.
- Popper, K. (1959). *Logic of Scientific Discovery*. London: Hutchison and Co.
- Pouw, N. and McGregor J. A. (2014). *An Economics of Wellbeing: What Would Economics Look like if it were Focused on Human Wellbeing?*. IDS Working Paper 436. Brighton: IDS.
- Rodriguez-Takeuchi, L. K. (2014). *Incorporating People's Values in Development: Weighting Alternatives*. Methodological Background Note ODI.
- Roelen, K. and Devereux, S. (2014). *Evaluating Outside the Box: Mixing Methods in Analysing Social Protection Programmes*. CDI Practice Paper 06. Brighton: Institute of Development Studies.
- Rojas, M. (2004). The Complexity of Well-Being a Life Satisfaction Conception and a Domains-of-Life Approach. *Paper for the International Workshop on Researching Well-Being in Developing Countries* Hanse Institute for Advanced Study Delmenhorst, near Bremen, Germany, 2–4 July 2004.
- Rojas, M. (2008). Experienced Poverty and Income Poverty in Mexico: A Subjective Well-Being Approach. *World Development*, 36(6), 1078–1093.
- Samman, E. (2007). Psychological and Subjective Wellbeing: A Proposal for Internationally Comparable Indicators. *Oxford Development Studies*, 35(4), 459–486.
- Seers, D. (1967). The Limitations of the Special Case. In Martin, K. and Knapp, J. (eds), *The Teaching of Economic Development*. London: Gass.
- Sen, A. (1981). *Poverty and Famines: An Essay on Entitlement and Deprivation*. London: Oxford University Press.
- Shore, C., and Wright, S. (2015). Governing by Numbers: Audit Culture, Rankings and the New World Order. *Social Anthropology*, 23: 22–28.
- Stiglitz, J. E., Sen, A., and Fitoussi, J.-P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. *Sustainable Development, Commission on the Measurement of Economic Performance and Social Progress*, p. 292.
- White, S. (2010). Analysing Wellbeing: A Framework for Development Practice. *Development in Practice*, 20(2), 158–172.
- White, S. (2014). *Wellbeing and Quality of Life Assessment: A Practical Guide*. Rugby, UK: Practical Action Publishing.

11

Conclusion

Laura Camfield and Keetie Roelen

In the Conclusion we return to the nine chapters and review the evidence that they provide to support the claim made in the Introduction that mixed methods in research on poverty and vulnerability can offer greater credibility, acknowledgement of complexity and usability. We also summarise the chapters' contributions to measurement, evaluation and increased uptake of research by policymakers.

Chapter 2 by Fahmy, Sutton and Pemberton in the opening section on Poverty Measurement uses qualitative data from 14 focus group discussions to assess the credibility of evidence derived from a large-scale UK-based survey used to establish public agreement on what constitutes the necessities of life (those things everyone should be able to afford and no-one should have to do without). The focus group discussions suggested that public understandings of 'necessity' were diverse and not always consistent with understandings held by researchers, or the way in which the term is used in survey measurement. For example, the researchers observed considerable variation in understandings of what was necessary, and that this was not related to the background of the respondents. This variation was not always resolved through discussion, thereby challenging the idea of the existence of a societal consensus.

Based on qualitative investigation, the chapter makes useful recommendations for researchers designing or analysing questionnaire-based surveys. For example, being aware that many respondents may confuse the original question wording of 'should' (e.g. should people live on two meals a day) with 'could' or 'would' (e.g. could people live on two meals a day, would they live on two meals a day if they had to), making it more of an empirical than a normative question, holds important implications for both survey design and the use of survey data by policymakers. The chapter also draws out challenges that respondents faced

in answering questions without adequate contextual information, for example, deciding whether a car was necessary or not would require consideration of household circumstances, composition, availability and affordability of public transport, among others. Fahmy et al. argue that survey developers should include techniques such as cognitive interviewing and behavioural coding to increase their understanding of what people's responses mean.

In terms of mixed methods research, Fahmy et al. advocate using mixed methods for complementarity – to draw out different aspects of research problems – rather than for triangulation, as the former does not assume that different measures are commensurate, given their different philosophical and epistemological underpinnings. They suggest that policymakers should approach the findings of poverty consensus surveys more cautiously and should not assume either stability over time or that answers capture what respondents' opinions might have been after greater reflection (e.g. in the context of individual interviews versus that of a group discussion).

Chapter 3 by Edmiston also uses qualitative data alongside survey data to explore people's lived experiences of poverty. He aims to increase the credibility of studies of these experiences by showing how the 'phenomenology of deprivation' is as reflective of the structures of citizenship as it is of individual experiences. Individual interviews also enable him to capture the recursive dimension of poverty where "self-conscious" emotion such as shame is constructed in reference to an individual's expectations and circumstance as well as the expectations and circumstance of others'. Interviews also allowed him to pay greater attention to people's agency in order to understand 'how people negotiate a socio-economic and political landscape that has increasingly come to structure their marginality'. Edmiston draws on the concept of citizenship to access the relational dimensions of poverty and vulnerability, including relationships with institutions. The use of the concept of citizenship avoids the trap of constructing inclusion and exclusion solely in relation to participation in market exchange. Edmiston argues that such a narrow view pushes researchers towards income measures at the expense of understanding how poor people relate to each other and to members of other economic groups.

Methodologically, Edmiston claims that subjective measures of social status are often as accurate as objective ones and more useful in predicting sequelae to poverty such as increased morbidity. By combining quantitative analysis with a second phase of qualitative data generation and analysis, he was able to explore the meanings of the quantitative

differences observed in the first stage of the study (e.g. why deprivation reduced support for certain social and economic citizenship rights). Edmiston's study challenges the convention of measurement that poverty is a fixed or isolated condition, arguing instead that it is a signifier of socio-economic and political relations within a particular context. Paradoxically, an 'objective' analysis of these relations cannot be done using purely quantitative measures. He also shows the effect of positionality on responses in that affluent respondents were much less aware of the systemic features of structural inequality or the ways in which marginalised people nonetheless exercised agency within this.

The chapter makes the important point that while subjective perceptions and lived experiences cannot be understood in isolation from institutional arrangements, this does not undermine their objective significance. In fact, it is only possible to understand the causes and consequences of policy by capturing the socio-economic and political relations that shape it. The chapter exemplifies Edmiston's expressed aim of using mixed methods to offer a 'new methodological framework through which to assess the efficacy of poverty alleviation processes in a comparative context'.

Dawson's chapter starts from the contradiction that while in Rwanda national indicators show rapidly declining poverty, studies that consider the control that people have over their land, property and the way they live suggests a parallel process of declining wellbeing. He argues for a greater use of mixed methods to capture specificity (depth) as well as complexity (breadth), which can only partly be captured through quantitative multidimensional measures. The application of his approach to some of Rwandan government's policies such as villagisation, which were celebrated internationally due to their success according to a narrow set of poverty indicators, highlights the way they were perceived negatively by the population and impacted severely on the poorest households (e.g. the financial burden involved in moving to areas with housing built to a higher standard). In pursuing the theme of differential benefit, Dawson used cluster analysis to identify four different groups who had and were having very different experiences of Rwanda's development. While the groups were intentionally not identified according to existing social categories, further analysis found a strong ethnic dimension to the clustering to the extent that none of the Twa households belonged to the two wealthiest categories.

Dawson draws a distinction between approaches that aim to achieve consistency of measurement (e.g. by measuring whether a household has sufficient income to meet their needs with a standard 'basket of goods') and those that try to reveal the specificities of context (e.g. the effects

of social complexity and the variations in factors that characterise poverty and wellbeing in particular contexts). He notes the dangers of making judgements on poverty and wellbeing status based on a limited set of variables that may exclude some of the things that most matter to people, for example, political freedom. During his fieldwork, participants provided numerous examples of the ways in which their behaviour and practices were subject to controls they felt were excessive, even extending to how they constructed their houses and reared their livestock.

While Dawson's study was primarily research-based, the implications for evaluation and policy from his use of mixed methods are drawn out by an exploration of the trade-offs entailed in apparently beneficial poverty alleviation initiatives and the differential effects they can have on different groups within the community. It also shows how perceptions of success are shaped by the choice of measure. For example, the 2001 Participatory Poverty Assessment that was conducted by the government contained no reference to politics, ethnicity or culture, except in the sense of the latter being a barrier to modernity. A combination of methods and measures can reduce such risks while at the same time providing greater insights into differential experiences of benefits.

The final chapter in this section, by Tincani and Poole, combines quantitative methods from the ecological sciences and participatory qualitative research tools to gain greater understanding of the complexities in building resilience for rural livelihoods in Burkina Faso. By focusing on polygamous families – related households living together in family compounds – the study introduces an additional layer of complexity in capturing intra-household distribution of food and other resources as each member of the family who identified themselves as contributing to the family's food security was interviewed repeatedly over a period of 14 months. Assets and entitlement to the assets of others were also measured for every household member. In addition to looking at individual decision-making, the study looked at the families as a unit and compared them across sites. The timeframe enabled the study to capture the differential effects of the four main seasons on household activities and consumption, which would not have been possible in a study with a shorter duration. Due to the distinctive nature of the polygamous household form, it was not possible to demonstrate the expected relationship between diversification or asset levels and food security, and so the qualitative data was drawn on to explain the absence of such a relationship. For example, asset-rich households might have made poor investment decisions, illustrating that it is not just the existence of assets but how they are managed that matters.

Tincani and Poole use the Local Adaptive Capacity framework developed by the African Climate Change Resilience Alliance to investigate climate change adaptation at the level of local government institutions and apply it instead to individuals and their households to test its utility at the micro-level. This is an entry point for mixed methods as the framework includes softer indicators such as innovation, which would be difficult to capture accurately using a purely quantitative approach. The use of multiple approaches enabled triangulation and also improved the quality of measures used. For example, the gender of the livestock was included in asset inventories when it was demonstrated that this was important because male animals were usually sold.

Tincani and Poole also highlight the agency of household members in, for example, proactively reducing their consumption in the dry season to enable them to eat more in the lean season when they needed their strength for farming, or using networks of relatives to sell their produce in major urban centres in the south. They conclude that the more nuanced understanding of resilience and the strategies used to achieve it provided by a mixed methods approach produces findings that are more credible and usable for policymakers as they reflect the complex local realities.

The first chapter in the second section (Evaluation Research), by Copestake and Remnant, focuses on the development of a qualitative impact protocol to generate and analyse qualitative data in a credible and cost-effective way, offering an alternative or complement to a largely quantitatively dominated stream of research. The protocol developed by Copestake and Remnant needed to be something that small NGOs could use to evaluate their projects – so an ‘appropriate technology’ – but also something that external audiences would judge to be rigorous. They increased the credibility of their impact claims by employing multiple strategies to address attribution, so broadening the range of possible causal mechanisms explored and triangulating findings. Copestake and Remnant put forward the concept of ‘credible causation’ based on the reasonableness of claims, rather than rigour defined in narrow statistical terms. While this falls short of ‘scientific certainty’ they argue that this may be the most we can hope for in complex contexts, and is potentially more useful for policymakers than the results of randomised control trials (RCTs). In order to avoid suggestions that they had used the qualitative analysis to ‘cherry pick’ quotations to support particular perspectives, they inductively grouped perceived drivers of change. If these had been mentioned by more than two informants they systematically tabulated them and indicated the level of frequency with which they were mentioned.

Copestake and Remnant argue that mixed methods provide an opportunity to deal 'with "organised complexity" on its own terms, rather than through a process of deliberate reduction into a closed model with a more manageable number of variables'. For example, the qualitative analysis also illustrated the way that the same drivers, for example, advice from extension agents, were mentioned in relation to multiple indicators. One of the main challenges to the credibility of qualitative evaluation are claims of bias, which the authors argue can be partially addressed by triangulation with quantitative research studies subject to different forms of bias. Their pilot also experimented with a simple if logistically challenging way to reduce pro-project bias by blinding both respondents and researchers to the nature of the project being evaluated. Whilst this was successful, it introduced space for other forms of bias (e.g. pro-authority bias where the researchers were assumed to be evaluating government initiatives), illustrating the need for continual vigilance.

The authors critically reflect on the extent to which the evidence generated will be useful in evaluation and the role of mixed methods in this. For example, while respondents did not indicate the magnitude of the impact they experienced, this could be resolved by combining the qualitative data with data from household surveys in later phases. They note the problem of replicability in that an NGO might not be able to recruit a similar calibre of researchers or maintain the double-blinding. Given that the study was small, there was also a question of external validity; that is, how generalisable the findings would be to other villages. They addressed this by documenting key sources of variation between sub-areas within the project area and inviting knowledgeable stakeholders to sort villages into groups according to what they thought were likely to be the most important sources of variation in performance. This made it easy to see how the sampled villages compared with the rest of the project.

The second chapter in this section, by Devereux and Roelen, proposes a new way of studying the complex and non-linear nature of poverty reduction pathways, which aims to enable evaluations to capture unintended outcomes as well as intended impacts. They look at the complex dynamics within households, for example, between biological and non-biological children and between different generations. They also capture potential tensions between intended beneficiaries of grants (e.g. specific children) and their recipients (parents who have responsibilities extending to the whole household). The use of mixed methods enables them to capture effects on other actors within programmes that are not always

acknowledged; for example, traders who distribute grants on the proviso that a percentage of the grant is spent in their shop, or project officers whose different positionalities set up different relationships and deliver different sets of outcomes.

The chapter makes the case for mixed methods evaluation designs by highlighting two important and influential aspects that are often neglected by standard evaluation methodologies: programme processes ('design choices, implementation modalities and assumptions implicit in the theory of change') and social dynamics at the household, community and programmatic levels. Attention to the first explains why intended impacts might not be achieved, while the second shows how impacts might be compromised because they establish or modify a key set of social relationships. The evaluation design they propose goes beyond what might be expected from an evaluation in turning a critical eye on the project's theory of change, such as the linear and wholly positive relationships assumed in school feeding schemes. By outlining a methodology for including unintended social impacts in evaluation, despite their lack of amenability to quantification, they create a more nuanced understanding of the outcomes of different interventions and the ways these might vary across groups, even causing some participants to be worse off than they were before.

The final section on Research to Policy starts with a case study from Torres Penagos and Bautista Hernández illustrating the use of mixed methods to establish the extent and causes of multidimensional poverty in a Colombian municipality. The authors argue that the combination of methods enhanced their validity as not only did they provide a more complete picture of the municipality by doing so, but they were also able to triangulate different findings. They addressed the complexity of people's experiences of poverty and the structures that shape these through a combination of secondary analysis of national survey data, participant observation with a single family, focus group discussions addressing community problems and interviews with decision-makers.

Torres Penagos and Bautista Hernández argue that their findings were more persuasive and useful to the local authorities because they were able to provide an holistic picture of the condition of the municipality which accurately represented people's experiences. Their approach enabled them to look at relational dimensions of poverty such as the relationship between citizens and local authorities, in addition to more quantitative questions on the prevalence, depth and intensity of poverty. It also enabled them to draw on the knowledge the inhabitants

themselves held about poverty and develop poverty categories and indicators that held local significance. Nonetheless, data generation, analysis and presentation required a high degree of 'methodological bilingualism' and integration was only possible because they had two researchers with a solid foundation in each other's preferred methodology.

The chapter also contributes to a growing body of literature on the utility of the UNDP's Multidimensional Poverty Index (MPI) at different scales and in different contexts, and the value added by qualitative information in adapting and analysing the MPI. For example, they observed missing dimensions such as domestic violence (something that has also been noted by the developers of the MPI) and variations in the way that core dimensions were prioritised by respondents. Finally, their study noted discrepancies between the quantitative and qualitative data, for example, in relation to the incidence of child labour. This could be explained by the phrasing of the question in the survey data the MPI was based on which did not allow the possibility of children working and studying at the same time.

The final phase of the research involved interviews with decision-makers to raise their awareness of the research and check both the viability of the public policy proposals that were starting to emerge from it (e.g. strategies to address the high incidence of teenage pregnancy) and the decision-makers' responsiveness to these. They identified many problems such as alcohol abuse that were not captured by national measurements, but were nonetheless imperative to address in this community. The researchers subsequently worked with the local authorities to address overcrowding and improve the use of public resources such as drinking water, which they felt was made possible by the integrated use of information at the level of local decision-making.

The second chapter in this section, by Burrows and Read, describes a cross-country evaluation of the World Food Programme's response to protracted refugee situations. The mixed methods design shows how qualitative methods provided insights into unintended results and sensitive issues not captured by quantitative methods, contextual factors outside the interventions' control and implementation factors that were also influential. As in McGregor et al. and Devereux and Roelen (this volume), the most interesting insights include the effects of the intervention on social dynamics. For example, the finding that far more Somalis in camps were in polygamous marriages compared to the host population was explained by the fact that it enabled women to take ration cards under their own name and so was an important food access strategy.

In order to increase the credibility of the evaluation, the authors adopted a common evaluation quality assurance system (EQAS) that set

out the mandate and rules for all the evaluations, including the observation of common principles and practices of utility, partnership and transparency. For example, where accounts of different observers did not converge or there was insufficient evidence in support of an 'anecdotal' finding, this was noted in the report. EQAS provided detailed guidance and templates for the processes and written outputs at all stages of the evaluation, which also enhanced comparability across contexts and research teams. Data was triangulated between sources, but also by 'member-checking' when initial findings were presented to in-country informants and other stakeholders for feedback, including refugees and host communities. The mixed methods fieldwork was carried out sequentially so that other research teams could learn from the experiences of the other countries, which is particularly important for mixed methods designs.

As a result, the usability of the evaluations was high and led to both follow-up at country level (e.g. the provision of additional support to host populations in two of the countries) and strategic level action by the international NGO partners (e.g. the agreement to develop a strategy for the transition to self-reliance and establish the partnerships to achieve this). Usability was ensured by the participation of stakeholders in evaluation reference groups and formal management response presentations, and a focus from the outset on the dissemination and use of evaluation products for programme and policy improvements. Nonetheless, the authors acknowledge that more could have been done to ensure dissemination within the countries where the evaluations took place as this was left to field staff and was not monitored.

The main contribution of the study, alongside the welcome emphasis on usability, was a demonstration of the feasibility of a contribution analysis approach in strengthening evidence of linkages across the causal chain and the use of mixed methods within this approach. This was in a context of limited data, complex interventions, diverse contexts, limited resources and all the other constraints associated with working in a humanitarian context, which are not usually amenable to experimental designs. The contribution analysis enabled a rigorous testing of programme theories, for example, the assumption that food distributions were consumed was only partly true as they were also used for income and collateral.

The final chapter by McGregor, Camfield and Coulthard argues that the credibility of policies and frameworks that aim to shape or represent wellbeing is enhanced by a more sophisticated understanding of the different dimensions of wellbeing (subjective, material, relational) and the role of subjective and objective data within all of these. The

proposed approach supports triangulation by collecting objective and subjective data on every dimension ('how much income do you have?' and 'how satisfied are you with your income?'), as well as data that is both quantitative ('how was the income spent?') and qualitative (observation and conversations about the effects that this expenditure had on the household relationships). The emphasis placed on the relational dimension in the case study on a cash transfer scheme in Zambia meant that the evaluation captured the perhaps counter-intuitive finding that the scheme appeared to increase social harmony as well as individual confidence by reducing the dependence of poor households on other members of their networks.

Given the interest in and claims by agencies of improved wellbeing through intervention, the authors argue that the time has come for a Wellbeing Focused Evaluation (WFE) methodology, which uses mixed methods to bring together information on different aspects of people's lives that are considered valuable to them. As in Devereux and Roelen (this volume), the methodology would enable identification and interrogation of differences in benefit as particular programme processes interact with the characteristics of recipients. The mixed methodology proposed by the authors is intensive but feasible, and offers a rigorous model for the use of participatory research to inform evaluation and, ultimately, support community-level processes of deliberation over what counts as acceptable trade-offs. In the Zambian example, the WFE was conducted after an RCT and although the WFE throws some light on the RCT's findings (e.g. the nature of the increased expenditures observed by the RCT), the very different epistemological starting points made it difficult to integrate the two datasets in a more profound way.

As a body of work, the nine studies combined in this edited volume illustrate the ways in which the use of mixed methods can be undertaken in a credible manner, used to address complex realities and be useful to policymakers. The contributions show how mixed methods can enhance the credibility and usability of research and evaluation, while also enabling it to more effectively capture complexity than research using a single method. Nonetheless, the authors reflect frankly on the challenges they face in, for example, reconciling different epistemological standpoints and assuring quality, and share some of the strategies they have adopted to address these (cf. Torres Penagos and Bautista Hernández, and Burrows and Read, respectively). This process of critical reflection and mutual learning, which started at the London workshop in 2013 and continued through this volume, offers the best opportunity for mixed methods research and evaluation to reach its potential in addressing the 'wicked problems' of development.

Index

Page numbers in **bold** refer to figures, page numbers in *italic* refer to tables.

- absolute poverty 23, 35–6
abstract individual, the 24–5
accessibility 219
accountability 164–5, 235
accuracy, poverty measurement 48
action research 5, 141
ActionAid 162–3
adaptability 79
adults, definition 113n6
affective states 46
affluence 44
African Climate Change Resilience Alliance (ACCRA) 90, 91, 98, 265
alienation 43
Alkire, S. 64, 179, 231, 239, 240
anecdotal findings 221
Angulo, R. 183
Arias Valencia, M. 187
asset entitlements 99, 103, 104
attitudinal difference 44–5
attribution 3, 141, 265; cause-and-effect statements 133–4; explicit 136; implicit 136; outcomes 119–20, 121, 123, 125, 130, 133–4, 201–2; quantitative monitoring 134; self-reported 130, 133
auditability 66
Australia 13
autonomy 72–3, 105, 108–9, 110
Bamberger, M. 200, 201, 202, 220, 226
Bangladesh 13, 159, 210, 211–12, 213, 220, 221, 222
Befani, B. 3, 151
Belgium 13
Bentham, J. 242
bias 129, 134–7, 139, 141, 150, 220–1, 266
Bilinski, H. 188
blinding 124, 135, 136, 146n19
Bliss, L. 190
Botswana 160
Brannen, J. 11
Breadline Britain survey 12–13
Bronstein, L. 176, 177
Brundtland Report, the 238
Bundy, D. 153
Burkina Faso 94; application of mixed methods 109–11; asset base 98–9, 105–7, 107, 112; asset entitlements 99, 103, 104; autonomy 108–9; common wisdom 105; food production 94; food security 87–8, 94, 97–8, 264–5; food security level assessment 92; food security levels 95, 97; food security strategies 98; innovation capacity 99; livelihood diversification 101–2, 102; livelihood strategies 95, 97, 98, 114n16; local adaptive capacity (LAC) framework 98; methodological approach 91–5, 97–101; resilience in 87–113, 101–9, 102, 104, 264–5; study sample 94–5, 96, 97
Camfield, L. 2, 149–50
Campbell, A. 45
capabilities 36, 63, 243
capability approach 68
Castro, R. 175
cause-and-effect statements 133–4, 136
Central African Republic 197
Chad 197, 208, 210, 211–12, 213, 220
change: contribution analysis 217–18; drivers of 129, 131, 132–3, 141; theory of 150–2, 151, 152, 155–6, 168, 198, 201–2, 217–18, 222–3, 227
Chase, E. 38
child labour 184

- children, resilience 89–90
- citizenship: definition 33; and deprivation 33–53, 43–5; exclusion 44; and lived experiences 38, 43–5, 47; and participation 40; and poverty 43, 262–3; social 34, 41; social rights 33–4, 41; theories 2; topographies of 44; and worth 43
- Citizenship Survey 40–1
- clientelism 164
- Colombia 182–3; Adjusted Poverty Gap 182; child and youth conditions 184; educational attainment 183–4; educational lag 184; health and healthcare 184–5; interviews 183; labour 185; living conditions 185–6; methodology 178–9; mixed methods 180–1, 182–3; mixed methods poverty diagnosis 173–93; multidimensional poverty 177, 178, 180–1, 267–8; National Planning Department 178, 179; new dimensions of poverty 186, 193; Participant Observation 179; policy recommendations 192; poverty measurements 177, 178, 179, 182; preliminary visits 179; priority areas of intervention 190–2, 191, 193; public policy proposals 183; triangulation 187–8; Villapinzón 177–8
- Commission on the Measurement of Economic Performance and Social Progress, Final Report 231–2
- common wisdom 105
- community profiles 251
- comparability 223
- comparison groups 201, 213, 222
- complementarity 11, 262
- complexity 1, 3–4, 144n2
- conceptual approach 68–9
- concurrent design 201
- confirmation bias 134–7, 141
- confirmatory approaches 81
- consensual approaches: contribution 25–7; credibility 27; critiques 19–21; decision-making processes 21, 23–4, 24–5, 26; focus group methods 17–19; methodological issues 21–2; origins 13; poverty measurement 9–10, 12, 12–14, 36–7; valid consensus 20
- consensual deprivation index 17
- consistency 14, 48, 63, 64, 66
- Constas, M. A. 87
- construct validity 121, 222–3
- consumption poverty 63–4
- context specificity 78–9
- contextual information, need for 24–5
- contextual suitability 66
- contribution analysis 134, 198, 200–2, 217–18, 221, 227, 269
- Copstake, J. 135
- coping strategies 88–9
- credibility 1, 2–3, 27, 52, 82, 111, 192–3, 268–9; definition 122; elaboration 144n4; impact evaluation 121–2; impact of confirmation bias on 137; improving 2; lack of 2; lived experiences 48–52, 53; qualitative analysis 266; subjective perceptions 47
- credible causation 122, 265
- Creswell, J. 175, 176, 177, 190
- cultural values 68
- culture 67
- Cummins, R. 244
- data analysis 100
- data collection methods 9, 92, 209–10, 211–12, 214–16
- data sources 5, 10–11, 112, 187, 210, 205, 221, 226
- deferred gratification 237–8
- deliberative methods 27
- Denzin, N. 10–11, 189
- deprivation: and citizenship 33–53, 43–5; lived experiences 34–5, 39, 42–3, 43–5, 48–52; material 13; measurement 35–43; mixed methods approaches 48–52; objective 41; relative 2, 13, 23, 44; Rwanda 74; subjective perceptions of 49
- deserving poor 25

- development, wellbeing as
 a measure of 4
- development decision-making 235
- development interventions 125;
 accountability 164–5; aims 149;
 confirmation bias 134–7; data
 analysis 125, 125–7, 126, 127, 128,
 129–30; drivers of change 129–30,
 131, 132–3; institutional landscape
 127; intra-community dynamics
 161–2; Intra-household dynamics
 158–61; outcomes attribution
 120–1, 121, 123, 130, 133–4;
 participatory methods 235–6;
 programme actors 162–5;
 programme processes 150, 150–7,
 151, 152; programme staff 162–3;
 pro-project bias 129, 135;
 qualitative impact protocols
 evaluation 119–43; quantitative
 approaches 120–1; social dynamics
 150, 157–67, 158, 166; theory of
 change 150–2, 151, 152, 168; and
 wellbeing 231–2, 233–4, 234–9
- development policies 62, 66, 67,
 80, 82
- Devereux, S. 151
- difference-in-difference
 estimations 149
- disaster risk management 87
- disciplinary backgrounds 2
- dissemination strategy 224–5
- domestic violence 186, 193
- double difference approaches 213
- Doyal, L. 36, 244
- Durand, M. 242
- Duvendack, M. 149–50
- early childhood services, access
 to 184
- ecological resilience 88
- economic actors 163–4
- economic self-interest 44–5
- economic theory 105–6
- economistic models 234
- education 72, 152, 183–4
- educational lag 184
- elite capture 164
- employment 40, 185
- empowerment 49
- entitlement 24–5
- entitlement theory 97
- environmental entitlements 113n7
- epistemology 2–3, 12, 193
- epistemological: assumptions 11, 21,
 21–2, 262; and poverty
 measurement 9, 10, 46; clashes 2;
 divides 3; perspective 2; starting
 points 270; tensions 48, 270;
 understandings 2, value 2
- equality, of status 34
- ethical issues 82, 137
- Ethiopia: cash transfer
 programme 161; data analysis
 126–7, 127, 129; data collection
 methods 209–10, 211–12; drivers of
 change 130, 132–3; Eritrean and
 Somali refugees 197, 209–10, 213,
 215–16, 217, 220; gender relations
 158–9; generalisability 137–40;
 methodology 123–4; qualitative
 impact protocols evaluation 119–43;
 respondent selection 138–9
- ethnographic tradition 189–90
- European Union, material
 deprivation 13
- evaluation designs 200
- evaluation management
 structures 208
- Evaluation Quality Assurance System
 (EQAS) 198–9, 200, 268–9;
 accessibility 219; assessment 225–6;
 construct validity 222–3;
 contribution analysis 217–18, 221;
 data collection methods 209–10,
 211–12, 214–16; dissemination
 phase 207, 218–20, 224–5;
 evaluation management structures
 208; evaluation phase including
 fieldwork 204–5, 214–16, 221;
 evaluation phases 202, 203–7,
 225–6; evaluation reports 218;
 external evaluators 209; external
 validity 220, 223; inception phase
 204, 209–10, 211–12, 213–14;
 inconsistencies 217; internal design
 validity 221; internal validity 220,
 220–3; management response and

EQAS – *continued*

- follow-up action 219–20; objectivity 220–1; preparatory phase 202, 203, 208–9; qualitative methods 215–16; Quality Assurance 208; quantitative data collection methods 214–15; reliability 221; reporting phase 205–7, 216–18; sampling strategies 210, 213–14; statistical conclusion validity 222; strengths 225–6; threats to validity 220–5; transparency 225; triangulation 217, 226; utilisation 223–5; validity 216
- evaluation studies 3–4
- Evidence for Development (EFD) 145n12
- exclusion 34, 42–3, 44, 45, 81
- exclusion errors 153
- exploratory research approaches 81
- external evaluators 209
- external validity 3, 220, 223

- Fan, S. 88
- feedback 19, 27, 140, 201, 208, 209, 214, 269
- feedback loops 249
- Fielding, J. 11
- Fielding, N. 11
- fieldwork methods 168, 208, 215, 250–1
- Finland 13
- Fitoussi, J.-P. 232
- focus groups 17–19, 26, 71, 182–3, 251, 261; cognition issues 22–4; decision-making 23–4, 24–5; methodological issues 21–2; need for contextual information 24–5; participant interactions 19, 21–2; participant recruitment 18; resilience 92; task understanding 22–3; truth claims 21
- food security 87, 89–90; and access to labour 106–7; and autonomy 108–9; and decision-making 103, 104, 105, 112; and flexibility 105; level assessment 92; levels of 95, 97; livelihood diversification 101–2; quantitative analyses 97–8; refugee situations 198, 215–16; and
 - resilience 91–2, 97–8, 110, 264–5; and savings 107, 107; strategies 98; variations in 101
- Foster, J. 179, 239
- Frey, B. 242

- GDP 234
- gender relations 158–9
- generalisability 80, 137–40, 141, 188, 266
- generational dynamics 159
- geographic isolation 186, 193
- geographical bias 139
- Ghana 160
- Giddens, A. 43
- Gough, I. 36, 244
- governance problem, the 238
- Graham, Carol 257n12

- Haiti 91
- Hammersley, M. 10, 122
- happiness economists 242
- happiness scores 242
- Harrison, G. 165
- Harriss, J. 49
- health and healthcare 39, 184–5
- hierarchical cluster analysis 73, 73–4, 74
- holistic understanding 109–10
- Honduras 91
- household surveys 9, 67, 71, 79, 211, 213, 214, 233, 266
- Howe, G. 66, 74
- 'How's Life' Framework (OECD) 244–5, 245, 255
- Hulme, D. 64
- Human Development Index 243
- Human Kind Index 247

- identification problem 12
- impact evaluation 2, 3, 130, 265–6;
 - aims 119, 149; application of mixed methods 198; bias 150; comparison groups 201, 213; concurrent design 201; confirmation bias 134–7; confounding factors 123, 123; credible 121–2; data analysis 125, 125–7, 126, 127, 128, 129–30; data collection 124; definition 121;

- drivers of change 129–30, 131, 132–3;
 evaluation designs 200; evaluation
 management structures 208;
 exclusion errors 153; external
 evaluators 209; fieldwork methods
 168; generalisability 137–40; impact
 indicators 123, 123, 125; inclusion
 errors 153–4; indicators 149;
 intended impacts 165–7, 166;
 interventions 123, 123; intra-
 community dynamics 161–2;
 Intra-household dynamics 158–61;
 limitations 149–50; literature review
 200–2; methodology 122–5, 123;
 monitoring surveys 138; other
 impacts 149–68; outcomes
 attribution 119–20, 121, 123, 125,
 133–4, 201–2; possible divergences
 167–8; programme actors 162–5;
 programme processes 150, 150–7,
 151, 152; pro-project bias 129;
 protracted refugee situations
 197–228; qualitative impact
 protocols 119–43; quality assurance
 208, 227; quantitative approaches
 120–1; quantitative monitoring 134;
 quantitative surveys 150;
 respondent selection 138–9;
 scientific rigour 121–2; sequential
 designs 201; short-term impacts
 155–6; social dynamics 150, 157–67,
 158, 166, 266–7; social impacts
 165–7, 166; stakeholder
 participation 200; strengthening
 227–8; substitution effects 154, 155,
 166–7; success 141; unintended
 impacts 165–7, 166; validity 120;
 wellbeing 270; Wellbeing-Focused
 Evaluation (WFE) 248–53, 250, 254,
 255–6. *See also* Evaluation Quality
 Assurance System (EQAS)
- inclusion 34, 40, 45
 inclusion errors 153–4
 inconsistent data, reconciling 139
 Index of Multiple Deprivation 40
 India 154
 inequality 42, 44–5, 75–6, 81
 integrated analyses 70
 intended impacts 165–7, 166
 interface analysis 168n2
 internal validity 220, 220–3
 International Advisory Group 14
 inter-personal level conflicts 238
 interviews 41, 51, 69–70, 95, 99–100,
 124–5, 126, 126; frequency 139–40;
 multiple 139; researcher-participant
 relationship 79; schedule 142–3;
 semi-structured 183, 251; themes
 215–16; timing 139–40
 intra-community dynamics 161–2
 intra-household dynamics 158;
 beneficiary vis-à-vis recipient 160–1;
 biological children vis-à-vis
 non-biological children 160; female
 vis-à-vis male 158–9; older
 generation vis-à-vis younger
 generation 159
 intra-personal level conflicts 237–8
 Ireland 13
- Japan 13
 Jo, Y. N. 45
 Jordan, B. 33, 43
- Kenya 154, 155
 Kovacs, P. 176, 177
 Kraus, M. W. 44
 Kristjansson, E. 154
- Lansley, S. 20, 26, 36
 Lawson, D. 174
 Layard, R. 242
 leakages 153–4
 Leeuw, F. 197, 201
 Lister, R. 40
 literature 200–2
 lived experiences 34–5, 37, 42–3,
 182–3, 262–3; and citizenship 38,
 43–5, 47; credibility 48–52, 53;
 deprivation 48–52; interviews 41;
 mixed methods approaches 48–52;
 objective significance of 45–8, 52;
 and policy recommendations 52;
 selection criteria 41–2; subjective
 perceptions and 39, 53; validity 53
 livelihood assets 233
 livelihood systems 3
 living conditions 185–6

- living standards 9–10, 13, 19
 local adaptive capacity (LAC)
 framework 90, 98, 265
 longitudinal mixed methods 3
- Mabry, L. 200
 MacDonald, R. 37
 McGilchrist, I. 144n3
 McGregor, J. A. 232
 Mack, J. 20, 26, 36
 McKay, A. 64, 66, 74, 174
 McKay, S. 21
 macro-level data 47–8
 McWey, L. 188
 majoritarian approach 26
 Malawi: data analysis 125–6, 126, 129;
 drivers of change 130, 131;
 generalisability 137–40;
 methodology 123–4; qualitative
 impact protocols evaluation 119–43;
 respondent selection 138–9; school
 feeding scheme 154
 malnutrition. *See* school feeding
 scheme
 Mani, A. 46
 Manzano-Santaella, A. 120
 marginality 42
 Marshall, T. H. 33–4
 Marti, T. S. 52
 material resources, access to 81
 Mayne, J. 201, 217, 227
 Mertens, D. M. 4–5, 52
 methodological bilingualism 3,
 189–90, 192, 193
 methodological divides 10
 methodological opportunities 3
 metrics 234
 Mettler, S. 45
 Mexico 157
 micro-level innovation 99
 Millennium Development Goals
 (MDGs) 64
 Millennium Villages Project 61
 minimally adequate living
 standards 9–10
 Minimum Income Standard
 project 37
 mixed methods: advantages 173–4,
 176, 186–90; application 67–78, 73,
 74, 109–11, 198; approaches 174–7;
 benefits 112; challenge of 82,
 112–13; contribution 2, 66–7;
 definition 176–7; effectiveness 174,
 193; evaluation designs 200;
 formative stage 175; integration
 51–2, 80, 173–93; limitations 100–1;
 literature 200–2; methodology
 178–9, 180–1, 182–3; positionality
 78–9; research design 112–13; role
 of 4, 82; strengths 226–7; use of 1
 monitoring surveys 138
 mono-method approaches 10
 Morgan, D. 176, 188
 Mphale, M. 159
 multidimensional poverty 177,
 180–1; 239–40, 245, 267–8
 Multidimensional Poverty Index
 (MPI) 64, 65, 81, 194, 268
 Multidimensional Poverty Index
 Colombia (MPI-Colombia) 178,
 179, 180–1, 182, 191
- narrative data 125, 127, 128,
 129–30, 141
 Natsios, A. 119
 necessities of life 26, 233, 244, 261;
 cognition issues 22–4; consensual
 approaches 2, 12–14, 17, 17–19,
 36–7; decision-making on 23, 24–5;
 definition 9–10, 26; desirables 19;
 hypothetical family scenario 24–5;
 initial findings 17; interpretation 23;
 item selection 14; luxuries 19, 23;
 methodological issues 21–2;
 necessities 19; perspectives on 9–28;
 public attitudes towards 10, 13–14,
 20, 27; public perceptions of 16, 17;
 qualitative perspectives
 contribution 19–21; qualitative
 perspectives on 17–26; quantifying
 14–15, 15, 16, 17
 need and needs 24–5, 36
 Nepal, Karnali Employment
 Programme (KEP) 164–5
 Netherlands, The 13
 New Zealand 13, 35
 Northern Ireland Statistics and Research
 Agency Omnibus Survey 14

- Nussbaum, M. 241, 243
 nutrition surveys 214–15
- objective poverty 239
 objectivity 45, 220–1, 256n4
 O'Connor, A. 50
 OECD 244–5, 245, 255
 Office for National Statistics Opinions Survey in Britain 14
 Oswald, A. 242
 outcome evaluation 3–4
 outcomes, attribution 119–20, 121, 130, 133–4
- pairwise ranking 113n9
 panel studies 156–7
 Participant Observation 179
 participation 40, 235–6, 253
 participatory methods 235–6
 Participatory Rural Appraisal (PRA) 2
 Pawson, R. 120
 performance management
 culture 119
 Peru 154
 Plano Clark, V. 175
 policy assessments 236–7
 political actors 164–5
 Poor Britain study 12, 13, 26
 positionality 78–9, 82
 poverty: absolute 23, 35–6; affective states 46; approaches to 62–6; and citizenship 43, 262–3; complexity 175; consumption 63–4; definition 12, 14, 27, 38, 45, 52; denial of 38; dimensions of 39–40, 243–4; indicators 48, 62, 63, 64, 72, 80, 180–1, 216–17, 263, 268; intergenerational transmission of 152; lived experiences 34–5, 37, 42–3, 262–3; local definitions 81; macro-level data 47–8; mixed methods diagnosis 173–93; moral distinctions 25; multidimensional 177, 180–1; 239–40, 245, 267–8; new dimensions 186, 193, 268; objective 239; and participation 40; polarised approaches to 62–6; policy recommendations 52; psychosocial dimensions 40; psychosocial studies 46; public conceptions of 13–14; qualitative research 48–9; quantitative research 48–9, 173; relations of 53, 262–3, 267–8; relative 27, 36, 40; relative deprivation theory of 13, 23; risk markers 39; secondary 54n2; social relations of 40; structural causes 157; subjective perceptions 45–6, 47, 263
 poverty alleviation 37, 53, 80, 175; measurement 61; Rwanda 73
 Poverty and Social Exclusion Survey (PSE-UK) 9–28; 1999 13, 17, 21, 26; 2002 13, 17; consensual approach 9–10, 12–14; focus groups 14, 17–19, 28; initial findings 17; item selection 14; quantifying 14–15, 15, 16, 17; response rates 14; shuffle card exercise 15, 15
 poverty denial 47
 poverty measurement 5, 261–2; absolute 23; accuracy 48; analytical framework 35–43, 53; assumptions 33; challenges 188; cognition issues 22–4; Colombia 177, 178, 179, 182; consensual approaches 9–10, 12, 12–14, 36–7; consistency 48, 63, 64, 263; consumption approach 63, 65; and context 33, 263–4; context-specific approaches 65; focus group methods 17–19; hypothetical family scenario 24–5; identification problem 12; and lived experiences 37; methodological challenges 9; methodological issues 21–2; Minimum Income Standard 37; mixed methods approaches 10–12, 40–3; multidimensional 175; objective 35–6, 39, 42, 80; participation standard 36; PSE-UK survey 9–28; qualitative methods 17–26, 19–21, 62; quantitative methods 62; reference groups 38; subjective 37–9, 81, 264
 poverty prevention 53
 Poverty Reduction Strategy Papers 235
 power relations 68
 Pradel et al. 3
 preliminary visits 179

- private sector, impact
 evaluation 163–4
- program theory 202
- programme actors: economic 163–4;
 political 164–5; programme staff
 162–3
- programme impacts 4
- programme processes 150; cost-
 effectiveness 154; demand-side
 effects 156; exclusion errors 153;
 inclusion errors 153–4; inputs
 153–4; leakages 153–4; long-term
 impacts 156–7; outcomes 154–5;
 short-term impacts 155–6;
 substitution effects 154, 155; theory
 of change 150–2, 151, 152
- programme staff 162–3
- PSE-UK survey. *See* Poverty and Social
 Exclusion Survey (PSE-UK)
- psychosocial studies 46
- public policy, and wellbeing 235
- qualitative data and methods 9–28,
 48, 62–3, 175, 176; advantages 18;
 aims 18, 138; claims for 48;
 cognition issues 22–4; contribution
 19–21, 25–7; credibility 27, 266;
 data analysis 100; data collection
 methods 215–16; data collection
 tools 93; depth 40; evaluation
 designs 201; interpretation of
 quantitative 216; interviews 41;
 methodological issues 21–2; mixed
 methods approaches 10–12; on
 necessities of life 17–26; poverty
 48–9; poverty measurement 9–28;
 requirements 78–9; resilience 103,
 105, 110, 112; strengths 140, 226;
 triangulation 111, 188; utility 48;
 validity 51–2; wellbeing 70, 71–3,
 76–8, 237, 238–9
- qualitative impact protocols 129;
 attribution 119–20; confirmation
 bias 134–7; confounding factors
 123, 123; cost 140; data analysis 125,
 125–7, 126, 127, 128, 129–30; data
 collection 124; drivers of change
 129–30, 131, 132–3; generalisability
 137–40; household interview
 schedule 142–3; impact evaluation
 119–43; impact indicators 123, 123,
 125; interventions 123, 123;
 methodology 122–5, 123;
 monitoring surveys 138; outcomes
 attribution 119–20, 121, 123, 125,
 130, 133–4; quantitative approaches
 120–1; quantitative monitoring 134;
 respondent selection 138–9
- quality assurance 124, 208, 227.
See also Evaluation Quality
 Assurance System (EQAS)
- quality of life 232, 233, 236, 239–40,
 244, 245, 246, 252; literature 239
- quantitative data and methods 10,
 11–12, 48, 62–3, 176; collection
 methods 214–15; data analysis 100;
 data collection tools 93; evaluation
 designs 201; impact evaluation
 120–1; main features 174–5; poverty
 48–9, 173; qualitative analyses 216;
 qualitative impact protocols 120–1;
 requirements 78–9; resilience
 109–10, 111–12; secondary 35;
 strengths 226; surveys 150;
 triangulation 111; validity 51, 120;
 wellbeing 70, 73–6, 73, 74, 236–7
- quantitative monitoring 134
- questionnaires 92
- Rainwater, L. 47
- randomisation 144n1
- random control designs 120
- randomised control trials (RCTs)
 149–50, 167, 248, 252, 265, 270
- randomistas 2
- realism, pursuit of 4
- realist evaluation 120
- Reams, P. 176
- reference groups 38
- reflection 79
- refugee situations, protracted 268–9;
 aims 202; application of mixed
 methods 198; camp population
 databases 222; comparison groups
 201, 213, 222; conceptual
 frameworks 202; contextual factors
 218; contribution analysis 217–18,
 221; data collection methods

- 209–10, 211–12, 214–16;
 dissemination phase 218–20, 224–5;
 EQAS utilisation 223–5; evaluation
 management structures 208;
 evaluation phase including
 fieldwork 204–5, 214–16; follow-up
 actions 224, 225; food assistance
 199–200; food security 198, 215–16;
 host population groups 210; impact
 evaluation 197–228; inception
 phase 204; inception reports 210;
 institutional context 199–200;
 management response and
 follow-up action 219–20; outcomes
 attribution 201–2; preparatory phase
 202, 203, 208–9; qualitative analyses
 216–17; qualitative methods
 215–16; quantitative analyses
 216–17; quantitative data collection
 methods 214–15; reference groups
 208; refugee groups 197; refugee
 populations 210; reporting phase
 205–7, 216–18; sampling strategies
 210, 213–14; survey design 213;
 synthesis report 219; theory of
 change 198, 217, 217–18, 222–3
- relationships, importance of 4
- relative deprivation 2, 13, 23, 44
- relative deprivation theory of
 poverty 13, 23
- relative poverty 27, 36, 40
- reliability 221
- replicability 136, 266
- research design 112–13
- research ethics 79
- Research Group on Wellbeing in
 Developing Countries 254
- research methods, polarised 10
- researcher–participant relationship,
 the 66, 79
- researchers: dealing with confirmation
 bias 135–7; positionality 3
- resilience 87–113, 264–5; application
 of mixed methods to 90–1, 93,
 109–11; asset base 98–9, 105–7, 107,
 112; and asset entitlements 99, 103,
 104; and autonomy 108–9, 110;
 children 89–90; coping strategies
 88–9; data analysis 100; data
 collection tools 92, 93; and
 decision-making 99–100, 100–1,
 103, 104, 105, 110–11, 112;
 definition 88–9; enabling factors 92,
 98–100, 103, 104, 105–9, 107; focus
 groups 92; and food security 87–8,
 89–90, 91–2, 97–8, 110, 264–5;
 holistic understanding 109–10;
 indices 91; individual 89–90; and
 innovation capacity 99; knowledge
 base 99, 103, 105, 112; livelihood
 87–8; and livelihood diversification
 101–2, 102; local adaptive capacity
 (LAC) framework 90; measurement
 90–1; methodological approach
 91–5, 97–101; outcome indicators
 92, 97; process 89; process indicators
 92, 97–8, 101–2, 102; qualitative
 perspectives 103, 105, 110, 112;
 quantitative analyses 109–10,
 111–12; threats to EQAS validity
 220–5; variations in 101–2, 102, 104;
 and vulnerability 103
- resilient development 87
- results-based culture 119
- rigour 2, 66, 121–2
- risk assessment 99
- risk markers, poverty 39
- risk-pooling 95
- risk-taking 100
- Robb, C. M. 48, 53
- Roelen, K. 151
- Rohingya refugees, Bangladesh 197,
 210, 213, 220, 221, 222
- Roll, J. 45
- Royce, D. 187
- Rugh, J. 200
- Runciman, W. G., *Relative Deprivation
 and Social Justice* 38
- Russia 13
- Rwanda 61–82, 160; agricultural
 intensification 77–8; application of
 mixed methods 67–78; Congolese
 refugees 197, 213, 213–14, 220;
 consumption poverty 63–4; data
 collection methods 211–12; denial
 of politics 67; deprivation 74;
 development policies 63, 66, 76–8, 80;
 drivers of change 76; earnings 72;

Rwanda – *continued*

- economic growth 63–4; education 72; household sampling 69, 70; inequality 75–6, 81; interviews 69–70; land and land use 71–2, 78; national indicators 61; participatory poverty assessment (PPA) 66–7, 264; poverty alleviation 73; poverty assessment 66–7, 264; poverty dynamics 74; poverty levels 61, 63–4, 75, 81, 263–4; qualitative analyses 70, 71–3, 76–8; quantitative analyses 70, 73–6, 73, 74; rural wellbeing 67–78, 69, 73, 74; social relations 72–3; socio-ethnic diversity 69, 70; study areas 68–9; villagisation 76–7, 263; wellbeing in 65, 67–78, 76, 81; wellbeing priorities 71–2
- sample size 100
- sampling strategies 210, 213–14
- scarcity, psychology of 46
- school feeding scheme: aims 152; demand-side effect 156; inputs 153–4; long-term impacts 156–7; outcomes 154–5; short-term impacts 155–6; theory of change 151–2, 152, 155–6
- Scotland 247
- Scudder, Thayer 145n17
- Seers, D. 233–4
- selection criteria 40–1, 70
- Sen, A. 36, 97, 113n7, 232, 243, 247
- sequential designs 201
- Shaffer, P. 11, 12, 40, 48, 188
- Shulha, L. M. 190
- single resilience index 91
- Slater, R. 159
- social citizenship. *See* citizenship
- social dynamics 150, 266–7; intra-household 158–61; types of 157–8, 158; unintended impacts 167
- social facts 38
- social impacts 165–7, 166
- social network analysis 167
- social networks 109, 157–8, 158
- social policy, prioritisation 192, 193
- social protection 6, 123, 160; Wellbeing-Focused Evaluation (WFE) 248–53, 250, 255–6, 270
- social psychology 244
- social relations, and wellbeing 72–3, 233
- social scientists, role of 4–5
- social status, subjective perceptions 39
- sociocultural analysis 167–8
- sociocultural context. *See* social dynamics
- Solomon, C. 163, 168n2
- Somalia 91
- Sorde Marti, Teresa 4–5
- Soss, J. 45
- South Africa 13, 163–4; Child Support Grant 159, 160–1
- specificity 66
- stakeholder participation 224
- statistical conclusion validity 222
- status, equality of 34
- status anxiety 38, 262–3
- Stern, E. 201
- Stiglitz, J. 232, 234
- Stutzer, A. 242
- substitution effects 154, 155, 166–7
- survey methods 20, 213, 261–2
- Sustainable Development Goals 87
- Sweden 13
- Tashakkori, A. 15, 176, 177, 189–90, 190
- Taylor, D. 46
- Teddle, C. 15, 189–90
- teenage pregnancy 186, 193
- themes 1
- thinking 144n3
- timing 188–9
- Tincani, L. 97
- Townsend, P. 13, 23, 36
- transparency 225, 235
- triangulation 10–11, 21–2, 75, 92, 112, 173–4, 192, 193, 266, 270; Colombia 187–8; Evaluation Quality Assurance System (EQAS) 217, 226; qualitative data and methods 111, 188; quantitative data and methods 111
- trustworthiness 66
- truth claims 11, 21
- Twale, D. 176

- undernutrition 89–90, 90–1
 undeserving poor 25
 unemployment 40
 UNHCR 199–200
 unintended impacts 165–7, 166
 United Kingdom, poverty measurement 9–28
 United Nations General Assembly 234
 usability 1, 4–5, 52, 269
 USAID 157–8
- Vaessen, J. 197, 201
 validity 11, 80, 216, 266, 267;
 construct 121, 222–3; external 3, 223;
 internal 220, 220–3; internal design
 221; lived experiences 53; objectivity
 220–1; qualitative research 51–2;
 quantitative research 51; reliability
 221; statistical conclusions 222;
 threats to 197, 220–5
 Van Praag, B. 37
 variation, coverage of 139
 Vietnam 13
 ‘Voices of the Poor’ study 243–4
 vulnerability, and resilience 103
- Walker, R. 20, 26–7, 38
 welfare policy design 37
 welfare regimes, effectiveness
 assessment 53
 wellbeing 62, 65, 67, 191, 231,
 269–70; and agricultural practices
 72; application of mixed methods
 67–78; and bodily health 241;
 capability approach 68; community
 profiles 251; conceptual approach
 68–9; conceptualisation 253–4;
 conflicts 237–8; data requirements
 231; definition 68, 232;
 development impacts on 231–2,
 233–4; dimensions 232–3, 254;
 drivers of change 76; and education
 72; evaluation framework 248–53,
 250; frameworks 239–41, 242–5,
 245, 254; holistic framework 80, 81;
 impact evaluation 248–53, 254,
 255–6, 270; indicators 241, 243–4,
 246, 251; and inequality 75;
 inter-personal level conflicts 238;
 interpretation 231–56; interviews
 69–70; intra-personal level conflicts
 237–8; local definitions 81; material
 68, 75–6, 78, 81, 232–3, 238, 246,
 254; as a measure of development 4;
 measurement 231–2, 234–9;
 methodology 250;
 multidimensional 239–40, 242;
 multidimensional frameworks
 243–5, 246; needs assessment 236;
 objective 242, 247, 252; policy
 assessments 236–7; priorities 71–2;
 and public policy 235; qualitative
 analyses 70, 71–3, 76–8, 237, 238–9;
 quality of life 232, 233, 236;
 quantitative analyses 70, 73–6, 73,
 74, 236–7, 251–2; relational 68,
 76–8, 81, 232–3, 238, 246–7, 253,
 254; research 46; rural Rwanda
 67–78, 69, 73, 74; in Rwanda 65;
 self-reported satisfaction 247; single
 measures of 242; and social
 relations 72–3, 233; societal context
 246–8, 255; subjective 68, 76, 81,
 232–3, 242, 244, 246–7, 252, 254;
 subjective experiences of 69;
 three-dimensional framework
 245–53, 250; trade-offs 238;
 universal framework 236
 Wellbeing-Focused Evaluation
 (WFE) 248–53, 250, 254,
 255–6, 270
 White, H. 121
 White, S. C. 68
 women: autonomy 108–9;
 employment 185; savings 107;
 social dynamics 158–9
 World Bank 65–6
 World Food Programme 155
 World Health Organisation Quality of
 Life 240
- Zambia 231–2; Child Grant Cash
 Transfer Programme (CGP) 249;
 Child Support Grant 252; Social
 Cash Transfer (SCT) schemes 249–53;
 social protection programme
 248–53, 256, 257n13, 270
 Zimbabwe Emergency Cash Transfer
 (ZECT) pilot project 161–2, 238