Independent Review of the Reality Check Approach

Final Report

October 2018

Calder Consultants Inc.

Australian Aid

Contents

Acronyms

ACFID	Australian Council for International Development	**SBR**	**Socio-Behavioral Research**
ADB	Asian Development Bank	**SCG**	Senior Citizen Grant (Uganda)
BAPPENAS	Indonesian Ministry of National Development Planning	**Sida**	Swedish International Development Cooperation Agency
BEP	Basic Education Program (Indonesia)	**SSI**	Semi-structured Interview
BPD	Badan Permusyawaratan Desa	**TNP2K**	The National Team for the Acceleration of Poverty Reduction
BRG	Badan Restorasi Gambut	**ToC**	Theory of Change
DDD	Doing Development Differently	**TWP**	Thinking and Working Politically
DFAT	Australian Department of Foreign Affairs and Trade	**UNESCAP**	United Nations Economic and Social Commission for Asia and the Pacific
DFID	UK Department for International Develompent	**UNICEF**	United Nations Children Fund
DHS	Demographic and Health Survey	**UNOPS**	United Nations Office for Project Services
DST	Digital Story Telling	**VECO**	Vredeseilanden/VECO, now Rikolto
ERIC	Ethical Research Involving Children	**VfM**	Value for Money
FES	Focused Ethnographic Study	**WASH**	Water, sanitation and hygiene
FGD	Focus Group Discussion	**WHO**	World Health Organization
GoI	Government of Indonesia	**WFP**	World Food Program
HCD	Human Centered Design		
HOM	Head of Mission		
IP-SSJ	Integrated Programme for Strengthening Security and Justice		
KI	Key Informant		
KII	Key Informant Interview		
KSI	Knowledge Sector Initiative		
MEL	Monitoring, Evaluation and Learning		
MOFA	Ministry of Foreign Affairs		
ODI	Overseas Development Institute		
OECD	Organisation for Economic Cooperation and Development		
PDIA	Problem Driven Iterative Adaptation		
PEER	Participatory Ethnographic Evaluation and Research		
PPA	World Bank Participatory Poverty Assessments		
PRA	Participatory Rural Appraisal		
PSA	Poverty and Social Analysis		
QE	Quick Ethnography		
RAC	Review Advisory Committee		
RAP	Rural Access Programme		
RCA	Reality Check Approach		
RDI	Research for Development Impact		
REA	Rapid Ethnographic Assessment		
RCA	Reality Check Approach		
RDI	Research for Development Impact		
RRA	Rapid Rural Appraisal		
RTI	Research Triangle Institute		

Acknowledgements

This independent review of the Reality Check Approach (RCA) was commissioned and supported by the Poverty and Social Transfers Section in the Department of Foreign Affairs and Trade.

We would like to express our heartfelt appreciation for the time and effort of the RCA core team and large number of RCA researchers in engaging with this review. We are grateful to the DFAT officers in Canberra, but particularly in Jakarta post, for hosting us during our field visit and consultations. Thank you also to those who participated so enthusiastically in interviews in Jakarta, Canberra, and remotely.

Triashtra Lakshmi was an excellent support both during the field visit, and in the earlier stages of the review. Paulina von Lowis of Menar stepped in at the last hour to provide invaluable assistance with referencing and general copy editing.

The Review Advisory Committee has been extremely generous with their time and provided excellent guidance and advice at key moments during this review. We are extremely grateful. We would like to thank DFAT's Poverty and Social Transfers Section for being true thought partners in this endeavor – in particular Francesca Lawe-Davies who initiated this review, and Tess Connolly, who ably guided it from inception, forward, including participating in the field visit.

Any insights in this report that help those who commission, consume and conduct research to do more rigorously and ethically surely came from those we acknowledge above. Any shortcomings in this report are those of the review team alone.

Rebecca Calder
10 October, 2018
Vancouver, Canada

Emily Wylde
10 October, 2018
London, England

1 Executive Summary

1.1 Introduction

1.1.1 Background and purpose

The Australian Department of Foreign Affairs and Trade (DFAT) has an interest in innovative research and evaluation methods to improve the quality and impact of its aid investments. The Department employs a range of methods to undertake poverty and social analysis, including a variety of qualitative research methods. One qualitative research method that has attracted particular attention in recent years, but has not been thoroughly assessed by DFAT, is the Reality Check Approach (RCA). For this reason, the Poverty and Social Transfers team in the Department of Foreign Affairs and Trade seeks to review the experience with the RCA in the context of other comparable research methodologies in order to assess the strengths and weaknesses of RCA vis-à-vis other qualitative approaches to poverty and social assessment, the overall effectiveness in informing development programming, and its potential relevance for Australia's broader aid program and bilateral interests.

1.1.2 Review design and overall approach

This review has the following overarching features:

1. It is theory-based, meaning the review was informed by what is supposed to happen, i.e. a Theory of Change (ToC).
2. It is a process review, meaning that is generated evidence and insights from the unfolding 'story' of the RCA in Indonesia and elsewhere.
3. It employed comparative analysis both through a range of data collection and analytical methods and establishing a number of comparators that enabled us to assess the relative effectiveness of the RCA vis-à-vis other approaches to qualitative research.

Our overall review approach was designed to explore a set of Review Questions (RQs), set out in the table below.

Overarching RQ	Framing RQs
To what extent, how and why has RCA been effective in generating and using poverty and social analysis in program and policy work?	What are the strengths and weaknesses of RCA relative to its objectives?
	How does RCA compare to other approaches to qualitative poverty and social assessment?
	What is the potential relevance of RCA and similar poverty and social analysis approaches to the broader Australian aid program?

A range of sub questions were developed during inception to guide the literature review and key informant interviews.

1.2 Methodology

1.2.1 Data collection

Our three principle data collection tools were:

1. Collection of secondary data through a literature review: In total we reviewed 230 unique sources of information, of which 96 were produced by the RCA team. As this review was unable to delve deeply into all of the different RCAs that have been implemented to date (over two dozen, across multiple countries), we selected a sample of cases for which to conduct a "deep dive".
2. Key informant interviews: These included informants who were (a) internal, those who have been directly involved in implementing the RCA process or comparator approaches; (b)

connected, those who have commissioned or consumed RCA and/or comparator approaches but have not been involved directly in implementing these[1]; and (c) external, those not involved in commissioning or consuming RCA or comparator approaches, but with an expert view on the uses of mixed method research and evaluation research.

3. A field visit to Jakarta: We observed an RCA post-fieldwork debrief and a sensemaking workshop, conducted 2 focus group discussions with RCA researchers, and held a number of key informant interviews with commissioners and consumers of RCA studies; and a short visit to Canberra to hold discussions with key informants from within and outside of DFAT.

1.2.2 Data analysis

We employed 4 main analytical methods to explore the research questions; none of these were discretely applied but instead were used iteratively and to enable us to explore different angles of the question.

1. Theory of Change (ToC) analysis, which facilitated a comparison between how the RCA is intended to work in theory and how it has been applied in practice in various contexts.
2. Analysis of rigor and ethics using a framework arrived at through a literature review of international good practice in qualitative research, key informant interviews with qualitative research experts, examining how other qualitative research approaches (including comparators) interpret and apply rigor, and reviewing the RCA's own documentation on rigor.
3. Value for Money Assessment, which enabled us to explore the extent to which RCA achieves the '4Es' that are standard in VfM (economy, efficiency, effectiveness, and equity), using the expected outputs, outcomes, and impacts articulated in the ToC, as well as an understanding of comparator approaches (Human Centred Design Research, SenseMaker™, PEER, and "good standard" mixed method qualitative research).
4. Assessing utility in 2 main ways: a. through specific case studies of how RCA has been applied; and b. through discussions with DFAT staff and others commissioners and consumer who have used poverty and social analysis evidence.

1.2.3 Limitations and mitigating strategies

During Phase 1, the review team identified a number of important factors that we believed needed to be considered in how we framed and conducted the review. We reflected these factors in our review design, and have reviewed them at key points in order to better understand the implications of limitations and challenges as they played out in practice, and ensure that we were mitigating these as much as possible. The below is a summary of 6 main limitations and how we have sought to mitigate these.

1. The RCA's evolution over time presented a challenge in that we had to assess a moving target in terms of both theory and practice, but it also provided us with opportunities for inquiry, for example on why adaptations have been necessary and implications for the future design and implementation of RCA.
2. Many different stakeholders have been involved in the process of applying the RCA. This gave us a good pool from which to draw our key informants; however it also created a challenge in terms of selecting key informants. We nearly doubled our original target of 40 KIs, which was considerably more time and labour intensive, but has yielded rich and robust results.
3. A limitation with our original 40 internal: 40 connected: 20 external split of key informants was that it had the potential to unfairly weight KII evidence towards positive findings, in that "internal" KIIs were conducted with RCA core team members and RCA researchers. We therefore reduced the proportion of "internal" key informants and expanded the number of

[1] There are a few exceptions to this. Several of the commissioners we interviewed had taken part in the RCAs they commissioned as researchers. They provided a unique and valuable perspective. We have classified them as "connected" but they could have been classified as "internal". This would effectively increase the "internal" sample size.

interviews with external experts who were familiar with RCA and a broader range of research and evaluation approaches.

4. There were limited opportunities to see RCA "in action". This created challenges in terms of understanding how RCA researchers are trained, and how debriefing, synthesis and analysis are conducted. We mitigated this to some extent by observing a day of debriefing and a day long sensemaking workshop. We also reviewed a large amount of internal RCA documentation, as well as exploring these issues in depth during KIIs with RCA core team members and RCA researchers.
5. Selecting comparators was challenging due to the relatively unbounded nature of some approaches and their varied application in different contexts. We managed this by developing a two-phase process to screen comparators. We also bolstered a literature review of comparator approaches with KIIs with practitioners and evaluators of these approaches, and have applied the same methodological frameworks for assessing them as we do for the RCA.
6. Without a full evaluation of these other approaches we are aware that there could be a slight bias towards other approaches, given the inevitable divergence in the thoroughness of the assessments for comparators vs the RCA. Recognizing this potential imbalance, we were careful to contain our findings, so that the comparators are used to situate RCA within the range of potential options that could be used in similar research contexts, and the range of likely outcomes that might be achieved.

1.3 About RCA

1.3.1 What problem is it trying to solve?

The RCA was developed to respond to the problem that the voices and perspectives of the poor are not heard by policy-makers or those involved in the development, monitoring and evaluation of programs. This problem, according to RCA practitioners, is rooted both within the wider context, as well as within methodological failings of other poverty and social analysis methodologies. Assessing the accuracy of this analysis, and the effectiveness of RCA as a "solution" to these problems, is therefore central to this review.

In terms of the wider context, the RCA is proposed as part of a response to the following root problems:

1. Policy and program officers lack the knowledge and capacity to commission and consume qualitative data and risk aversion towards non-expert data and new forms of knowledge. The result is that policy and program officers rely disproportionately on quantitative research which fails to pick up key insights into how and why change happens.[2]
2. People are unwilling or unable to voice their perspectives and experiences because there are few spaces in which to do this, and a lack of trust and self-confidence, manifested in a belief that their views will not be respected or acted upon. This results in limited opportunities for intended beneficiaries to influence policies and programs.
3. There are weak on-going feedback loops during implementation, reducing the speed with which research findings enter the public domain and opportunities for program adaptation.

RCA has also emerged and evolved as a response to perceived problems in other qualitative approaches to poverty and social analysis. These perceived root problems are:

1. Single-sector, narrow lens, linear research and evaluation approaches predominate over systems-based, open-ended, contextual approaches.
2. Inherent researcher bias at design, data gathering, and interpretation of design, monitoring and evaluation research often remains unacknowledged and/or unaddressed.

The RCA critique suggests that there is a tendency to ignore the everyday experiences in which the poor live, which misses key insights into their rationale for behavior, and/or leads to a focus on the

[2] see also Shah, 2018; and Bell and Aggleton, 2016.

wrong areas, with serious consequences. Approaches that are focused on “finding out” rather than “learning” are seen to be conducted in a way that limits trust-building and fails to address power imbalances or encourage open sharing of issues both directly and indirectly relevant to the issue being researched. Similarly, approaches that use etic (“outsider”) over emic (“insider”) interpretations are seen to fail to present the insights and implications that could have the most profound influence on policy and programming.

1.3.2 How does RCA present a solution?

The Reality Check Approach (RCA) is a 4 day and 4 night immersive ethnographic approach in which teams of researchers live with families in study communities[3]. Over the course of the immersion, the researchers typically engage with hundreds, and sometimes thousands, of people; while some of these conversations are brief and cursory, others (particularly those with host household members and a small number of neighboring or “focal” households), are more in-depth.

RCA combines elements of rapid ethnography (living with people, usually those who are directly experiencing poverty or any other phenomenon being researched) with 'light touch' participant observation. As with other approaches to “people centred research”, such as listening studies and beneficiary assessment, the focus of RCA is on engaging with, listening to, observing, and documenting the voices, opinions, and experiences of people, and asking questions as a curious learner as part of a relaxed conversation. Due to the shorter timeframe than most ethnographic investigations, there is necessarily a stronger emphasis on conversations than on observing behavior and the complexities of relationships.

Lewis (2012) and Masset et al (2016) write about the 4 key principles that underpin the Reality Check Approach and that they suggest sets it apart from both quantitative approaches to research, and many other qualitative approaches. These are: i. depth of findings, ii. respect for the voice of participants, iii. flexibility of fieldworkers to pursue conversations rather than follow a set question format, and iv. simplicity of the less complex (‘light touch’) compared to large-scale quantitative surveys or full long-term ethnographic studies.

There is a strong emphasis on researchers as conduit rather than intermediaries, understanding people’s lives in context and on relaxed, informal, participant-led interactions. Researchers try to pose questions in an indirect manner so that participants do not feel that researchers are trying to seek a specific answer or are being judgmental. RCA practitioners argue that by asking indirect questions, and not being overly “hooked” on a particular agenda, the biases found in some other qualitative research can be reduced. The RCA approach promotional materials make much of RCA’s ability to reduce, or even eliminate, bias altogether. They claim that, while “Others have project bias or agenda bias. We don’t. Our only allowable bias is to take the position of people”.[4]

1.3.3 Emergence

The first RCA study was established by the Swedish Embassy in Dhaka and by Sida headquarters in Stockholm. It was conducted was in Bangladesh starting in 2007 as a 5 year project, with the same researchers returning to the same households every year over the course of 5 years. Following, this RCA was introduced in Indonesia in 2009/10 to provide insights into how activities under the Australian Government-funded Indonesia Basic Education Program (BEP) had been experienced by people living in poverty. In April 2014, the DFAT-funded RCA+ plus project was launched. The project sought to build the capacity of Indonesian researchers and research organizations to undertake RCA studies and develop a tradition of quality, people-centred qualitative research. The purpose of this review is not to assess the outcomes achieved by the RCA+ Project, but rather to look at the value of the RCA itself, as a qualitative research method

[3] This is generally – though not always - the case. It was true for all of the case studies for this review.
[4] http://www.reality-check-approach.com/related-resources.html

1.3.4 Evolution and application

RCA is now far more ambitious than its first, relatively modest, beginnings as a supplementary approach, aimed at "getting an inkling" about ground realities ((KII, Qualitative Research Expert, quoting Greene, 2009). Over the past 4 years the RCA has been applied in a number of ways, including as a:

- diagnostic tool, used to inform policy formulation or the design phase or early implementation of a program or activity
- "pulse taking" situational assessment, used to gather quick feedback on the roll out of a program or policy, or to better understand the impact of policy changes or large events; and
- longitudinal evaluation method, primarily aimed at informing programming rather than policy (although these are rarer).

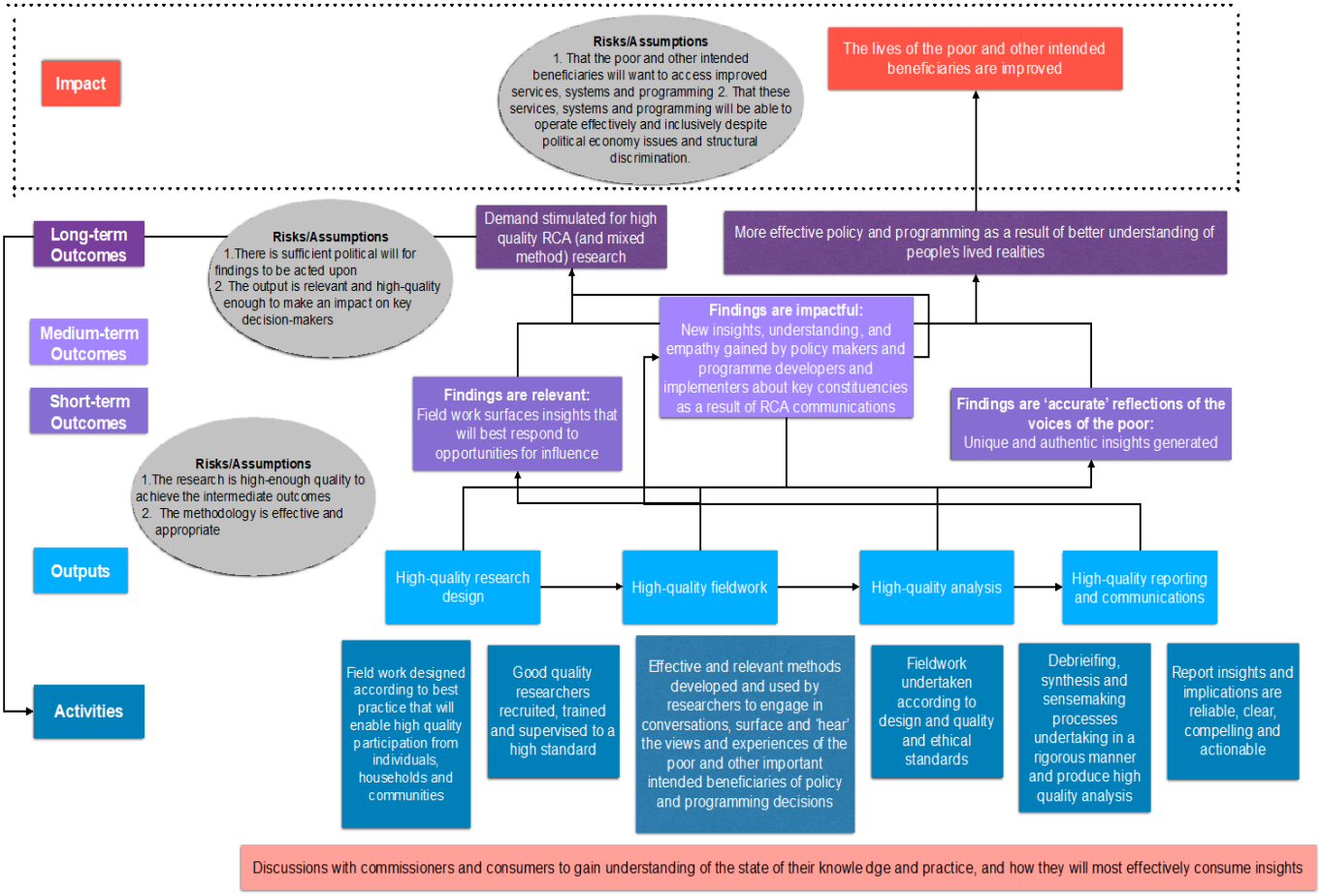
In addition to the above expansion in application, there have been 3 other major shifts in how the RCA is being used. First, there has been a stronger attempt at using the RCA alongside other methods – primarily quantitative approaches to poverty and social analysis, but also the use of visual methods such as participatory video and, more frequently, Digital Story telling (DST). Second, the RCA+ project has also enabled a stronger focus on policy in Indonesia, working with partners to identify policy moments where an RCA study could add value, and building relationships with key policy makers. Finally, there have been a few "experiments" in using the RCA approach to understand the lived realities of other constituencies, such as the urban poor, and university lecturers, rather than to better understand the lived realities of the rural poor (and those providing services to them).

The focus here is on RCA's effectiveness as a pulse-taking, diagnostic and evaluative research approach, as well as looking at how effectively it has been combined with other research methodologies.

1.3.5 Theory of change

In the context of this review, the overarching objective of RCA is to surface accurate, unique and authentic insights that will best respond to opportunities for influence and to communicate these to policy makers in a way that generates understanding, and empathy. It is theorised that this deeper understanding by policy and program stakeholders of whether and how their policies and actions translate into effective change on the ground, and how these efforts and changes are perceived, will not only inform but will influence future policy and practice so that it is better geared to local needs and context (see Masset et al., 2016 and Lewis et al., 2012). These changes (situated within the purple short-term outcome boxes in the Figure, below) we consider to be fully within the RCA's sphere of control. The medium- and longer-term outcomes are within the RCA's sphere of influence (they depend also on the actions of policy-makers), and the impact level change is within the RCA's sphere of interest (the ultimate goal, which also depends on many other external factors).

RCA Theory of Change



As we see from the above Figure, RCA attempts to achieve short and medium-term outcomes through 4 main outputs – which correspond to 4 main stages of research - and their contributing activities:

1. High quality research design - Field work designed according to best practice that will enable high quality participation from individuals, households and communities.
2. High quality field work - Good quality researchers recruited, trained and supervised to a high standard; effective and relevant methods developed and used by researchers to engage in conversations, surface and 'hear' the views and experiences of the poor and other important intended beneficiaries of policy and programming decisions; fieldwork undertaken according to design and quality and ethical standards.
3. High quality analysis - Debriefing, synthesis and sensemaking processes undertaken in a rigorous manner and produce high quality analysis.
4. High quality reporting - Reliable, clear, compelling and actionable report insights and implications generated.

1.4 Our Assessment

We orient our assessment around four key questions: Is it rigorous? Is it ethical? Is it effective? Does it provide value for money?

1.4.1 Is it rigorous?

As a concept, rigor is perhaps best thought of in terms of the quality of the research process; a more rigorous research process will result in findings that have more integrity, and that are more trustworthy, valid, plausible and credible (Given, 2008). Contrary to what RCA practitioners claim, we suggest that no research methodology has "intrinsic rigor" (RCA, 2017a). There are a number of features that are thought to define rigor in qualitative research (which differ from those found in quantitative research). We present our top-line assessment of the extent to which these are achieved by the RCA here.

Rigor in research design and preparation

There are four aspects of rigor related to research design and preparation that we assess the RCA against: i. whether there are experienced, reflexive, and well-trained researchers; ii. contextual understanding; iii. a framework to guide inquiry; and iv. a multi-disciplinary approach.

Experienced, reflexive, and well-trained researchers

There is a lack of experience and training amongst RCA researchers, and this is problematic for a number of reasons. Primary among them is that a lack of understanding of social and power dynamics limits the ability to know how to interpret what is being said, and inexperienced researchers are likely to be ill-equipped to know when and how to effectively probe issues (to 'know what they don't know' and to know what is important). Similarly, while inexperienced researchers may have *different* assumptions than more seasoned researchers, there is no reason to believe that they might have *fewer* assumptions. Indeed, more experienced researchers have spent years having any assumptions they might have had being thoroughly challenged in the course of research, years developing reflexivity, and years honing deep listening and probing skills. Level 1 training materials do provide a good foundation for growth, but it is questionable whether further levels of training support this adequately. While awareness of own positionality and biases appears to be a key component of RCA Level 1 training, it does not go deep enough in relation to gender, or local contextual issues.

Contextual understanding

There is inadequate research and orientation on key sectoral and other relevant issues prior to studies being undertaken. This lack of good contextual understanding hampers the overall design of the research (from the articulation of specific research questions to research instruments to sample selection) but it also makes it difficult for individual researchers on the ground to be fully reflexive and aware of how one is interpreting things and how they should be interpreted within a particular context. This can lead to bias, especially when amplified by the lack of training and experience.

Listening to people in their own context, as RCA core team members admit, is "surprisingly hard to do well" and needs not only a keen attention to detail, curiosity, and a good memory, but also an understanding of the context in which people are sharing their thoughts and experiences, and a deep knowledge of the ways that issues such as social norms and power dynamics influence attitudes and behaviours. While RCA claims that "people who are really knowledgeable go into studies with these pre-determined biases", we suggest that properly trained and experienced researchers are well aware of this possibility, and address it through rigorous reflective practice rather than avoidance of important contextual information.

A framework to guide inquiry

RCA's method of using an "Areas of Conversation" checklist to remind researchers of the kinds of areas that might be relevant to the overall research theme is to be welcomed. It enables researchers to conduct open and exploratory conversations, led by research participants.

However, the complete lack of focus in a vast majority of these checklists leads to major gaps in the relevance of inquiry. Further, there is no reason to believe that areas of conversation contain any less

bias than actual research questions or lead to richer and more meaningful conversations. It also reduces the comparability of data collected across multiple interactions and multiple sites.

Without a good conceptual framework and a set of research questions that have emerged from this, a concern is that researchers having little to no background in the context or issues being explored often fail to identify what lines of inquiry are worth pursuing and what are not; what statements might need to be challenged or probed further; and when attitudes, beliefs and behaviors conform with or challenge norms. One need not develop a rigid theoretical frame to guide research; a conceptual framework also allows linkages and relationships between issues to be explored in a more open manner (for example, exploring the relationship between formal and informal institutions related to land inheritance, or between social capital and human capital). A conceptual framework can also help to reveal and be explicit about assumptions, through a deeper pre-field discussion of how, and how well linkages and relationships within the framework are evidenced.

Multi-disciplinary approach

Simply including researchers with different academic backgrounds, does not make RCA research multi-disciplinary. RCA researchers are not recruited for the disciplines they practice: indeed, this appears to be irrelevant. And, they are not expected to utilise frameworks or experience from their professional training.

Rigor in fieldwork

At the fieldwork stage, rigor can be assessed in terms of: i. the unobtrusiveness of the researchers; ii. whether triangulation across people, methods, and time is done; iii. whether respondents are involved in validation; and iv. whether there is faithful and accurate recording of data.

Unobtrusive researchers

Considerable time is spent preparing researchers to "fit in" while staying in local communities. This is a good practice and is to be applauded.

According to researchers, they often arrive late in the day, and sometimes in inclement weather, and have to walk some distance to reach a suitable research location far enough away from their colleagues. Most researchers actively avoid local protocols for entering communities in order to try to remain unobtrusive and to avoid being afforded "respected guest status" with village notables[5]. While this is a very real challenge that ethnographers face, we suggest that the "solution" of a covert one-size fits all approach is disrespectful, and ineffective and, as we discuss in the report, can do harm to households and researchers.

Indeed, part of doing ethnographic fieldwork entails time and care taken to explain to local authorities the purpose of the research, and the importance of living with basic households. Flouting these conventions is not only disrespectful, it reinforces, rather than challenges, power dynamics by the researcher assuming that "normal" conventions don't apply to them. The RCA core team acknowledge that there are some trade-offs between complete openness of the presence of researchers on the field with the effort to gather genuine and candid responses from study participants (RCA+ transitional Report). We suggest that the trade-off is unacceptable, and fails to meet minimum standards for informed consent, including making people aware of any potential implications stemming from how their views on a policy or program are positioned in a public document, and the implications of a commissioner's response (or non-response) .

Suspicion, and lack of clarity concerning the purpose of the research and, potentially, host households who feel that they had to consent to researchers staying with them, all erode good feelings and trust. RCA practitioners claim the RCA enables a level of trust to be built between researcher and research participants that is absent in other approaches. While this is no doubt the case for some researchers

[5] In some cases courtesy calls to power brokers are make after households have agreed to host researchers.

and some RCA studies, difficulties in building trust raised in RCA documentation, and by researchers themselves, suggest that there is some over-claiming of a close and trusting relationship between researchers and household members.

Triangulation across people, methods and time

The RCA approach of using conversation, observation and participation in daily life is a much-needed approach to social research. It enables researchers to have more natural and free flowing discussions, gain empathy for research participants, and use observation to compare what people say with what they do. It also enables researchers to access types of information that cannot be gained through surveys or through other qualitative and participatory approaches.

However, despite the fact that the RCA enables researchers to speak to a large number of varied individuals, it is weak in terms of triangulation of people as the distinct voices of different research segments are often lost due to not faithfully recording conversations and aggregation in analysis.

A failure to fully integrate other qualitative and participatory approaches renders it's triangulative power in relation to methods weak. RCA cannot claim triangulation of method through merely practicing something akin to anthropological participant observation.

Respondent validation

While clearly some of the more experienced RCA researchers do have skills in eliciting respondent validation as a natural part of conversations, respondent validation is not actively and purposively sought by RCA research teams. This has implications for both rigor and ethics. Respondent validation is an issue of rigor because when a researcher is relying heavily on emic (insider) analyses of complex issues, the researcher needs to ensure that they have fully understood the research participants. This becomes even more important when the researcher themselves know little about the issue under discussion, as they will need to check both "factual" and "interpretive" data with participants in order to ensure that they insights that they deduce and present as implication's or recommendations to policy makers are valid. Respondent validation is also an issue of ethics because it helps to ensure that participants, and their views and experiences, are not misunderstood and misrepresented, and that the research process is empowering and not extractive. An important ethical foundation for this is that research participants understand how the information will be used and with whom as this affects the context for individual decision-making on consent; people might not mind being represented in one way to one audience, but very much mind being presented in another way to another audience.

Good practice suggests that validation should be sought at a number of different levels: during the course of conversations, at the end of conversations or research exercises, and at the end of the research process or afterwards, by holding a group or community debriefing, or returning with draft analysis and interpretation to check this with respondents. But this would require that RCA researchers present themselves much more openly as researchers than is currently the practice.

Faithful and accurate recording of data

Ad hoc and inconsistent note taking practices significantly undermine RCA's claims to accurately report data collected in conversations, in particular, but also observations and experiences. Memory recall is deemed to be unreliable and unnecessarily introduces researcher bias and unsupported interpretation.

Rigor in synthesis, analysis, and reporting

The final stage is that of synthesis, analysis and reporting, and here we assess rigor in terms of i. the quality of iterative debriefing, synthesis, and analysis; ii. whether reports faithfully provide evidence in a way to demonstrate the credibility of findings; and iii. the transparency of results.

Iterative debriefing, synthesis and analysis

Iterative analysis processes when researchers are in the field are absent and this is a missed opportunity. Researchers take few notes (so are not processing information in the field as part of early analysis) and do not meet up while in the field to discuss emerging findings and other key issues.

Collective and participatory analysis processes are weak, and Team Leaders, who may not have been in the field, conduct the bulk of analysis.

Debriefing, while enabling an impressive amount of information to be "downloaded", suffers from a number of flaws. It is relatively extractive, which leads to limited and arbitrary probing, clarifying, and nuancing of information; little attention to reflection on possible biases; and overcategorization and limited exploration of different respondents' experiences and viewpoints.

The RCA's sole reliance on emic views, rather than judiciously balancing this with etic interpretation, yields descriptive data but largely fails to surface meaningful ethnographic insights. The RCA approach generates a collection of narratives, interesting, at times informative, but not analytical, and prone to bias due to a very weak understanding of power, positionality and context.

The RCA does not analyse issues of gender and social difference, rendering it unable to achieve interpretive depth. Not only that, RCA uncritically collects people's views without deeper interrogation of their positions, biases, blind spots, which can magnify researcher bias and be erroneous and misleading to policy makers who may assume, for example, that the lack of explicit reference to gender issues means that these are absent and therefore do not need to be considered.

Though RCA practitioners claim that grounded theory can emerge from RCA analysis, it is not clear how this is possible without the analytical process supporting the building, testing and evidencing of hypotheses. RCA's use of "framework analysis" appears to be useful for the purposes of organising data (particularly in relation to report writing), but does not appear to be fully utilised as an analytical tool.

Faithful and credible reporting

RCA reports do capture the interest of their readers, and provide a great deal of detailed information. Without interpretation and clear recommendations, however, RCA reports do not live up to their full potential.

It is clear that recent improvements in the documentation process have been made in the last several years. With the addition of more formal archiving processes greater transparency has been achieved.

However this review found that there is still considerable room for improvement. In particular, a lack of systematic note taking and rigorous debriefing reduces transparency.

Reports do a fair job of acknowledging limitations, but there is scope for improvement in this area, particularly concerning what they methods they use can and cannot achieve, and how their approach, and their sampling, affected findings, and how far they were able to mitigate limitations.

1.4.2 Is it ethical?

As with rigor, we can look at RCA's ethics throughout the research process.

Ethics in research design and preparation

While RCA's own ethical guidance is taken from the American Anthropological Association Code of Ethics (RCA, 2017), we do not consider these to be appropriate or sufficient ethical guidelines for conducting short-term research for development with large teams of researchers, many of whom have not conducted field research previously, or who's only exposure to qualitative research is through the

RCA. Nor do we accept that the RCA is an inherently ethical approach, in comparison to other qualitative poverty and social analysis methods used in development.

Training and research preparation have gone some way to addressing ethical issues, but there are considerable gaps and shortcomings that render RCA design and preparation unable meet high ethical standards. The training materials and new procedures we reviewed – including several risk assessments generated by these new procedures - are inadequate to mitigate some of the risks experienced by researched populations and researchers that we discuss in this report.

The responsibility for building ethical foundations rests not only with the RCA core team, but also with the entities within which RCA is "housed", and commissioners. If ethical foundations are not laid during researcher training and research design and preparation, there is a much greater likelihood of ethical problems emerging during research execution, putting both researchers and research participants at risk.

Fieldwork ethics

The RCA pays insufficient attention to critical ethical considerations, including consent, power and positionality. The RCA's attention to ethics is neither broad enough nor deep enough, particularly considering that researchers are living in the homes of study participants, not merely meeting them in tea shops, field, or schools, or community meeting places. Nor is it nuanced enough for the wide range of different contexts in which research is conducted.

New guidelines and procedures on entering communities and obtaining informed consent are welcome. However, without prior knowledge of local dynamics and proper introductions, researchers and the household in which they stay can be exposed to risk, during, and after the researchers have left. A less-than well thought out and executed community and host household selection process means that researchers are not familiar with local context within which the data is to be collected.

A lack of completely informed consent is not only unethical, but can result in suspicion, which reduces openness required for research participants to engage fully in research processes, as well as placing research participants at risk.

At the time of this Review, we identified training and orientation on child protection policies and procedures is extremely limited and did not equip RCA researchers to reconcile their responsibilities for child protection with RCA's quasi-anthropological non-intervention policies. RCA researchers are confused about what their responsibilities are in relation to responding to child safeguarding incidents, or indeed even what would constitute a child safeguarding incident.

Post facto note: in the latter part of this review, new policies and practices regarding child protection have been instituted. These include a full 1 day training in child protection issues, and identification of local partners on the ground who can support in such cases. These have not been assessed as part of this review as these changes were made outside of the Review period.

Ethical considerations in synthesis, analysis, and reporting

Archiving processes and secure data storage are adequate to ensure privacy and confidentiality. RCA core team members are aware that preserving anonymity requires that there be no link between the data and the source, and work hard to prevent the identification of communities, households or individuals in reporting.

There is value in listening to a range of voices and presenting these authentically, but the RCA is plagued with a range of shortcomings that undermine respect for voice. An approach whereby researchers "just listen" to what people say about their situation and provide what is perceived to be a relatively unfiltered perspective is valued by many. However, there are weaknesses in terms of faithful documentation and reporting of these voices, that continue to undermine RCA in terms of

respect for voice. Further, we argue that it is not ethically acceptable for an approach that claims to give primacy to local people's views and interpretations to have no process or methods for actively engaging participants with the findings and their implications.

Internal peer review processes are inadequate as they do not involve researchers with the experience and expertise necessary to challenge findings and recommendations (insights and implications).

1.4.3 Is it effective?

We assess effectiveness against the outcomes in the theory of change, namely whether there is a high-quality research output, and whether it leads to policy impact.

Does it produce high-quality research outputs?

Assessing the 'quality' of a research output is inherently somewhat subjective, but we can make the review as transparent as possible by setting out some objective criteria that should serve as reasonable benchmarks for determining whether the research meets a high standard of quality. These are: inclusiveness, validity, and relevance.

Inclusiveness

RCA fails fairly comprehensively to deliver on one of its main claims about amplifying the voices that might otherwise be missed; not only are they systematically under-represented in the sample as a result of the (lack of) rigor in the design process, their voices are then systematically under-represented in the synthesis and analysis that appears in the report.

The sampling selection criteria are inadequate, and reflect a lack of understanding about poverty and vulnerability in general. It also doubtful whether researchers can really consistently identify the poorest, most vulnerable, and marginalised households in the field given the way RCA is currently practiced, and whether these households are always in a position to host an out-of-town guest for several nights.

There are also major concerns with the way in which the voices of different groups are synthesised and reported. Data and interpretation are jumbled and tend to be highly generalised, with little meaningful analysis across different groups.

At all stages in the research, we found there was a tendency to view 'the poor' as an undifferentiated group, with no nuanced understanding of class, caste, ethnicity, gender or livelihoods.

Validity

RCA reports provide significant descriptive detail; however true depth, both descriptive and analytical, is largely absent. There is very little 'thick' description, instead quite a lot of aggregated or generalized data is provided, in a fairly unstructured way; there is no systematic presentation of findings against the key research questions. Evidence is not produced to convincingly back up the findings that are made. Instead, there is an impression of very selective use of examples, with no real analysis by group or context.

Claims to be able to collect and analyse data on sensitive topics are also not supported by the evidence.

Relevance

Relevance is severely compromised by the lack of contextual research – which would help to situate findings within what is already known and therefore what value the research adds to the knowledge base – as well as the lack of a clear research framework. The RCA claims to present findings that are surprising and counter-intuitive, challenging commissioner and consumer assumptions. While it is sometimes the case that research findings are counterintuitive and surprising, and could only have

been gained through immersive participant observation, some findings are simply resurfacing old knowledge.

If stronger background research were conducted in order to inform research design, these 'findings' should instead have been the *starting point* for the research, and then much more time and effort could have been put into the what, why, when, and how questions where the studies could have actually added value.

In which contexts is it more and less effective?

This review finds that if RCA findings are used to complement data collected in other ways, and are seen to provide a view onto a broader landscape, then RCA – with the ethics and rigor caveats discussed in Chapter 6 – has the potential to provide an important contribution to conventional research and evaluation.

RCA's particular value is in its ability to provide a glimpse into people's everyday lives. In contexts where little is known and ethnographic insights are important for filling an evidence gap, there is a stronger justification for more immersive research. For example, framing security and justice issues in ways consistent with how they are understood by communities was seen to be an important contribution of the IP-SSJ RCA study.

The RCA's weakness is its inability to provide real interpretive depth, which comes from a judicious combination of emic and etic.

Shah suggests that, "RCA findings should be presented as limited, though able to provide considerable insights and triangulation when interpreted within a wider evidence-base" (2018: 23). We similarly conclude that RCA, rather than offering greater local insights than other qualitative methods, can offer different and complementary insights, in certain contexts and if implemented to a high standard.

RCA has potential in contributing to theories of change that reflect more closely local people's views and aspirations for change. However changes are need in terms of how these are used, and by whom.

Does it influence policy and programming?

Policy influencing is a long and convoluted process; the policy landscape is complex and determining attribution - or even contribution - is difficult. The ability to translate RCA findings into policy-relevant insights is not a simple technical linear process, but depends all three actors in the system – policy makers, commissioners, and researchers (Lewis et al., 2012).

Some commissioners reported that RCA reports can provide them with information that is not surfaced by normal monitoring and can lead to greater understanding of ground realities. While RCA is not always able to achieve depth, commissioners certainly appreciate the breadth of detail provided in RCA reports. Others – primarily those with a strong sectoral or research background - are more sceptical of the RCA's ability to provide high quality analysis of ground realities that is useful for policy makers.

Where RCA has been able to "punch above its weight" in relation to policy influence (and these instances have been relatively rare) it has been due less to RCA as a method that is able to generate knowledge and translate this into actionable evidence, only very partially to connections that RCA has been able to make with policy actors, and more to the spaces into which RCA has been invited, and the opportunity that these have afforded. There is good evidence that RCA has been most successful when i. it has strong commissioners and an institutional "home"; ii. policy-makers are engaged throughout; iii. there is a significant gap between policy makers and poor people's lived realities.

The potential for robust qualitative research to contribute to more relevant, adaptive and politically astute development is profound and should not be overlooked. In particular, qualitative research can:

- help to support the identification and engagement of local conveners, as well as foster the relationships necessary for enable design and implementation processes that draw on local knowledge, feedback and energy.
- provide understanding of risk levels associated with different solutions, so that "small bets" indeed have small local risks (i.e. avoiding a situation where what outsiders perceive to be low risk might not be perceived so by insiders).
- provide timely insights into whether "small bets" are paying off, feeding in local views on which activities have promise and which should be dropped.
- Help to build trust and empower local people through respectful participatory research that listens to local voices, including the marginalized, thus promoting sustainability.

There is reason to believe that there are significant opportunities for rigorous, ethical qualitive research to support Australia's Indonesia program, and other development programming. If there are central incentives around doing development differently, then our arguments that qualitative research can contribute significantly to the doing development differently agenda should start to provide the justification needed to Heads of Mission to open up space for research approaches that enable a deeper understanding of local level dynamics, and provides timely insights to enable programs to learn about and adapt to what is working on the ground and what is not.

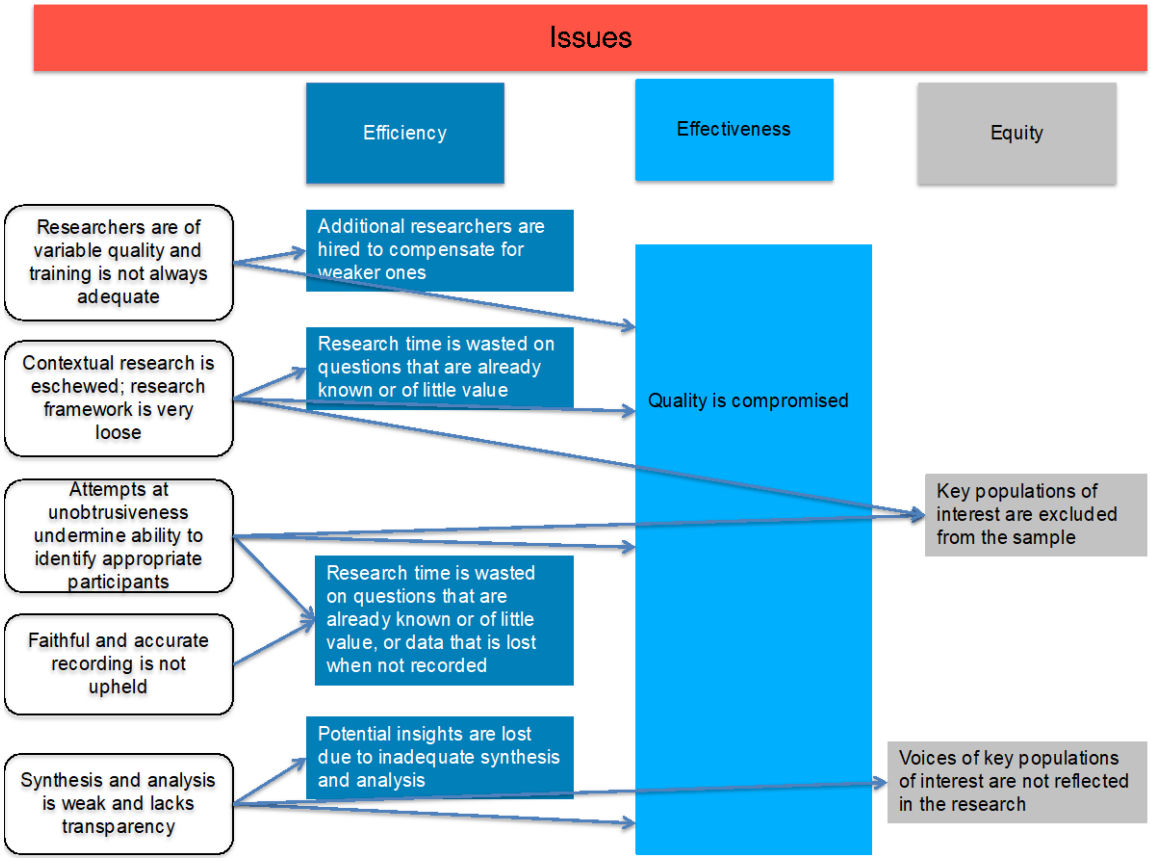
The elephant in the room, however, is the capacity and incentives in commissioning institutions such as DFAT to absorb and respond to the outcomes of research. Whilst there is no doubt that RCA practitioners do need to get better at promoting the accessibility of their research, for this to result in policy outcomes, however, attending simply to this supply side will not be sufficient. The demand side absorptive capacity of policy makers also needs to be addressed.

1.4.4 Does it provide value for money?

In theory, the approach that was able to produce research for the lowest cost would be the most efficient. Across the comparators here, PEER would be the most efficient, as it is implemented at an overall very low cost, and general qualitative work can also be similarly efficient. RCA and Sensemaker are both more expensive and therefore less efficient, and HCD is very expensive indeed (over 10-20 times the cost of PEER, for example) and therefore fairly inefficient from the perspective of research alone, although this is part of a wider design phase that yields additional outputs beyond the research itself.

The key to assessing VfM in terms of effectiveness ultimately rests on the quality dimension, which is encapsulated in terms of the rigor of the exercise and, relatedly, the strength, validity and relevance of the findings. From a VfM perspective, the ideal scenario is an approach that is high quality and low cost. We find that on balance, it is the 'general' qualitative and PEER approaches that achieve both higher quality and lower cost. By contrast, HCD is very high cost and tends to be lower quality. Both Sensemaker and RCA are in the middle, with a range of quality in terms of research output, and also reasonably expensive.

The figure below provides a summary of how issues with rigor contribute to challenges with respect to efficiency, effectiveness, and equity. Our overall assessment with respect to VfM is that these challenges, combined with relatively high costs, undermine RCA's ability to provide good value for money across the board.



1.5 Discussion

An ethnographic approach to research whereby ground realities are more deeply understood and contextualised is much needed and appreciated across a wide range of stakeholders. RCA's strength is in providing unfiltered and uninterpreted "snippits" (KII, Qualitative Research Expert) into the lives of the people (many in remote rural areas) with whom researchers live and interact over a four-day period. RCA has had a positive effect in that it has been utilised across a range of geographies on a range of issues with different donors, as one off and longitudinal exercises, more or less embedded within project MEL. It has piqued the interest of donors in the role of qualitative data as an important part of the evidence base. The fact that they are doing immersion, and that they live in remote communities and in challenging circumstances, is a well-valued contribution..

However, while RCA is able to generate detailed descriptive (though largely socially unsegmented) data, it is much less able to generate the rich explanatory data needed for greater understanding of poverty and social dynamics. Nor, in its current construction, can it surface, or analyse, the complex, heterogeneous experiences and power hierarchies that deeply shape different people's experiences, and the ways in which research must be sensitive to the gendered nature of poverty.

This is due to a number of factors related to the RCA methodology itself discussed above. Some factors are by design, as it intentionally:

- works with largely unseasoned researchers (who have little understanding of the local political economy or social norms);
- neither seeks researcher interpretation nor supports the generation of local interpretation (through, for example, testing and validating insights, and related recommendations, with participants themselves);
- does not draw on any analytical frameworks within which to make sense of what has been seen, heard or experienced by researchers.

Much of how RCA has been designed and executed focuses on the mitigation of researcher bias. While this is a crucial issue with which all good qualitative work must grapple, it is only one of many. All of the key methodological aspects of rigor need to be assessed holistically (because there may be trade-offs between them). The RCA's concerted focus on researcher bias actually undermines its ability to meet its own objectives of authentically representing the voices of the poor and marginalized.

Robust qualitative research has a long tradition (extensively documented in the literature) of grappling with these issues, and the key message from this is that bias and other inherent challenges to rigor need to be dealt with through conscious, reflexive, and careful practice, based on a solid understanding of the context. This includes the research framework (what questions to ask), the sampling strategy (who to ask), the instruments/tools to use (how to ask them), and how to then synthesize, analyze and interpret the responses.

Other factors that reduce the RCA's efficacy as an approach to poverty and social research are not by design, but relate more to implementation, such as the lack of triangulation across methods, and gaps in researcher training. There are easier to address, and we make suggestions for this in summary below (Section 1.6), and in detail in Chapter 9.

Where RCA generate findings that are useful to policy makers, it is generally through highlighting the difference between what people say, and what people actually do in the course of their daily lives - but it cannot, using the current methodological approach, credibly go much beyond this. This appears to not always be understood by commissioners, who make choices between using RCA or another qualitative research approach, or are unaware that there are other qualitative and ethnographic approaches that have a much stronger focus on analysis and interpretation to generate usable findings and actionable recommendations.

In the context of many consumers having little understanding or experience of qualitative approaches, and having little opportunity to experience ground reality themselves, they are intrigued by RCA and find that it brings a new perspective. For some donors, the RCA helps them to "tick a box" in terms of feedback from direct beneficiaries. However much a more grounded ethnographic approach is needed, the RCA needs to be viewed within a much larger field of possible qualitative research options than it currently is. Commissioners should be aware of the trade-offs between the RCA and other approaches, in terms of both rigor and ethics, as well as overall value for money. While RCA has grown under the primogeniture of some big names in qualitative research for development, there are a substantial number of other seasoned development researchers and evaluators who are extremely concerned with the method.

1.6 Key Recommendations

Our key recommendations in regard to the RCA approach are as follows. We have 13 key recommendations for practitioners and a further 7 key recommendations for commissioners. A full set of conclusions and recommendations against review findings can be found in Chapter 9.

1.6.1 Practitioners

Recommendation 1 - Place a much stronger emphasis on recruiting RCA researchers who have some previous qualitative research experience as well as knowledge of the issues that RCA studies will explore and the geographies within which research will take place; ensure that training is of sufficient length; and revamp training materials to improve content, and ensure a stronger progression of learning and competencies is established.

Recommendation 2 - Conduct significantly more background research to inform study design (including robust research and sampling frameworks) and execution. This should include both desk-based research and discussions with expert practitioners.

Recommendation 3 –Immediately develop and implement a vulnerable population safeguarding plan based on accepted ethical standards of research with vulnerable populations and in particular with children; and revamp the approach for entering communities ensuring that proper protocols are followed and permission is sought prior to research

Recommendation 4 – Develop and roll out a comprehensive standalone researcher training on child safeguarding, adequate orientation to child protection policies and procedures that guide the research[6], and discussion of ethical issues related to other vulnerable populations such as women experiencing domestic violence.

Recommendation 5 – Continue the good practice of training and orientation on how to be less obtrusive in the field, but use much deeper contextual knowledge to support this, and completely re-think the strategy for entering the community and obtaining consent

Recommendations 6 - Integrate more standard qualitative tools such as interviews, discussions, and PRA to aid triangulation, comparison and aggregation of data collected across multiple sites by multiple researchers. Present this data alongside more immersive ethnographic data.

Recommendation 7 - Ensure that respondent validation be more actively sought by researchers, and that it is an embedded practice with the approach.

Recommendation 8 – Incorporate much more rigorous note taking, and sufficient training on when, and how, to do this well.

Recommendation 9 – Reflect, rethink, and adapt to increase the ability of the RCA to do robust iterative analysis, including: experimenting with longer, punctuated immersions; drawing on detailed and comprehensive fieldwork notes in debriefings; seeking a judicious balance between emic and etic analysis; incorporating much stronger gender, power and social analysis; longer, more structured and more collaborative debriefing and sensemaking sessions.

Recommendation 10 - Unless the RCA evolves to become more analytically rigorous and to include expert interpretation – whether etic or emic but preferably a combination of both – it should steer away from providing conclusions or recommendations. We suggest, however, that successful policy influencing requires not only the generation of evidence, but the provision of advice, and that the RCA should adopt measure to ensure that it can provide sound advice.

Recommendation 11 - Stand-alone RCA studies are not recommended, with the exception of high level "journalistic" pulse-taking exercises that can fill information gaps in regard to how policies are affecting local people, and highly exploratory landscaping exercises followed up by more rigorous research methods.

Recommendation 12 - RCA studies should only ever be implemented when combined with other qualitative methods, and preferably embedded within a mixed qualitative quantitative exercise.

Recommendation 13 – Increase political saliency and policy relevance of RCA study topics is good, this could be significantly increased by:

[6] As per the Post Facto Note in 1.4.2, there have been updates to policy and practice in this regard during the course of this review that we do not assess here.

- ensuring sufficient knowledge of previous research on the issues to be studied, increasing contextual understanding at the local level, and engaging researchers who are experienced in formulating questions to get to complex and highly nuanced information.
- Continue to work closely with "evidence translators" that have political savvy and credibility.
- Work with commissioners to plan for and dedicate significant time and effort to policymaker engagement, relationship building and co-creation.
- Reflect on and address the relationship challenges that have plagued the RCA throughout its history.
- Continue to increase the accessibility of research and at the same time address credibility of research issues through some of the recommendations provided in other sections.

1.6.2 Commissioners

Recommendation 1 - A more formal ethical review and approval process is recommended in light of the fact that RCA research – and much other qualitative research - addresses sensitive issues or topics, involves vulnerable groups, uses considerable participant time, and is largely exploratory.

Recommendation 2 – Allocate sufficient time and budget to enable researchers to returning to the field to check draft analysis and interpretation in order to increase rigor –and adherence to ethical good practice.

Recommendation 3 - Commissioners should demand high quality gender, social and power analysis in all of the qualitative research they commission, including evaluations. If internal capacity to assess this is insufficient, then robust quality assurance should be commissioned.

Recommendation 4 - Make a concerted effort to address demand side absorptive capacity issues, as recommended in DFAT's ODE Review of Research (2015).

Recommendation 5 - Quality assurance of final report products should be carried out be qualified individuals, either as part of an internal review process or an external, commissioned, review.

Recommendation 6 - Be clear from the outset what the whole Theory of Change (ToC) is for a research output, and commission some follow-up/monitoring of impacts to be able to assess Value for Money against the full ToC.

Recommendation 7 - Increase understanding of the potential approaches that could be used for a particular research objective, and value for money considerations should enter at the commissioning stage. For example, rather than specifying that an exercise should necessarily use RCA, research could be tendered based on the research requirements, and then the approach offering the greatest VfM could be the one that is ultimately selected.

2 Introduction

This introduction provides a brief background to the review (2.1) and presents the review design and overall approach (2.2). Section 2.3 discussed engagement and use of the review, and 2.4 lays out the structure of the remainder of the report.

2.1 Background to the Review

The Australian Department of Foreign Affairs and Trade (DFAT) has an interest in innovative research and evaluation methods to improve the quality and impact of its aid investments. DFAT's Poverty and Social Transfers Section provides training and guidance to staff on analytical tools for assessing poverty and social issues as they relate to country, regional and sector strategies and the design, monitoring and evaluation of aid investments.[7] The Department employs a range of methods to undertake poverty and social analysis, including a variety of qualitative research methods.[8]

One qualitative research method that has attracted particular attention in recent years, but has not been thoroughly assessed by DFAT, is the Reality Check Approach (RCA). Originally, the RCA+ project was open tendered through the Australian Government's Poverty Reduction Support Facility (PRSF). As PRSF came to an end, the continuing phase 2 financial support was subsequently channelled through DFAT's Knowledge Sector Initiative (KSI)[9] in Indonesia. RCA has also been used by a range of other development actors in Indonesia, including the Government of Indonesia, and in other countries, most notably Nepal and Bangladesh. DFAT seeks to review the experience with the RCA in the context of other comparable research methodologies in order to assess the strengths and weaknesses of RCA vis-à-vis other qualitative approaches to poverty and social assessment, the overall effectiveness in informing development programming, and its potential relevance for Australia's broader aid program and bilateral interests.

2.2 Objectives of the review

The overall objective of the review was to explore a set of Review Questions (RQs). The overarching RQ and three Framing RQs are presented in Table 1, below. A range of Sub-EQs (see Annex 1) for each Framing RQ were developed during inception to guide the literature review and key informant interviews. These evolved during the course of the review to take account of new information, new avenues of inquiry, and DFAT's evolving appetite for information in different areas, for example an increasing desire to understand the ethics of the RCA, and less interest in understanding how the RCA might contribute to adaptive management approaches, and thinking and working politically.

Table 1 – Main Review Questions

[7] The Poverty and Social Transfers Team sits within the Department's Development Policy Division. It is the primary internal source of advice and training on poverty and social analysis. It works closely with the Development Economics Group, which provides economic analysis to inform strategies and designs, and the Development Policy Section, which provides guidance on development research. The Aid Management and Performance Branch provides operational guidance on design and evaluation issues.

[8] DFAT's research investment is highly decentralized, with 97 per cent of DFAT-funded research commissioned directly by individual country and thematic programs (in the period 2005-2013). Office of Development Effectiveness (2015). Research for better aid: an evaluation of DFAT's investments. Available at: http://dfat.gov.au/aid/how-we-measure-performance/ode/odepublications/Pages/research-for-better-aid-an-evaluation-of-dfats-investments.aspx. p.1.

[9] The Knowledge Sector Initiative (KSI) is a joint program between the governments of Indonesia and Australia that contributes to more inclusive and equitable growth in Indonesia by supporting the production of high-quality public policy grounded in rigorous research, analysis and evidence.

Overarching RQ	Framing RQs
To what extent, how and why has RCA been effective in generating and using poverty and social analysis in program and policy work?	What are the strengths and weaknesses of RCA relative to its objectives?
	How does RCA compare to other approaches to qualitative poverty and social assessment?
	What is the potential relevance of RCA and similar poverty and social analysis approaches to the broader Australian aid program?

This review sought to assess how rigorously and ethically RCA was implemented, and to what effect, with the support of the RCA Secretariat in Indonesia (at that time implemented by Palladium through the RCA+ Project). But we also looked more widely in terms of our case studies, as there were a number of studies implemented by the RCA core team outside of Indonesia, for example in Nepal. Our assumptions was that this would enable is to see RCA as it was meant to be, and at its best.

What this review is not is an evaluation of RCA+ as a project; this would have required us to use the logframe as the framework against which to assess rigor, ethics and effectiveness. We were also not evaluating RCA an open source research approach, as we would have then assessed how RCA study quality was influenced by using researchers who had not been trained in the RCA approach by senior members of the RCA core team and how study quality was influenced by other implementers.

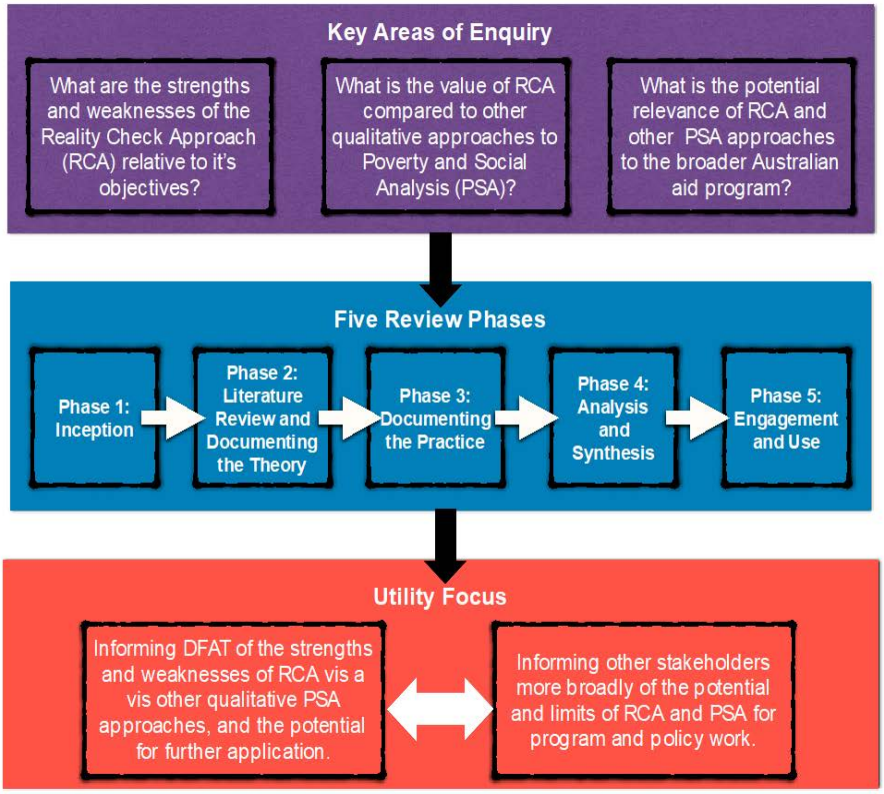
2.3 Review design and overall approach

The review has the following overarching features:

4. It is theory-based, meaning the review was informed by what is supposed to happen i.e. a Theory of Change (ToC). It began by establishing what it is that RCA seeks to achieve (particularly DFAT's objectives for RCA in Indonesia). It then tested whether the identified objectives have been achieved in practice. This approach has enabled us to investigate the 'why' and 'how' of change, helping us to address the key review questions (see Table 1 below). This ex post facto/ex ante RCA ToC was articulated together with the RCA core team during Phase 2 of the review.
5. It is a process review, meaning that is generated evidence and insights from the unfolding 'story' of the RCA in Indonesia and elsewhere. Documenting RCA in practice is central to a comparison with how the RCA was expected, in theory, to achieve the desired changes.
6. It employed comparative analysis in two ways. First, we used a range of data collection and analytical methods that enabled us to articulate the theory behind the RCA, document the practice, and compare the two. Second, we established a number of comparators that enabled us to assess the relative effectiveness of the RCA vis-à-vis other approaches to qualitative research. Our methodology is outlined in more detail in Chapter 3, below.

Our approach to the review is illustrated in Figure 1, which highlights the key areas of inquiry that we explored through the five phases of the review.

Figure 1: Review design and overall approach



2.4 Engagement and use

The review is utilization focused. Our primary audience for the review is our review partner, DFAT. Our secondary audience includes those who commission the RCA and other qualitative approaches poverty and social analysis and consume the data generated by them in order to inform more authentic people-centred policy and programming. A third audience is the RCA core team and others who conduct qualitative and mixed method research and evaluation.

There is clear demand for the review to document the application of RCA in Indonesia, and beyond, in order to demystify the approach, assess strengths and weaknesses, and examine the broader relevance of RCA and other similar approaches. To increase the usability of the review, recommendations are presented in Chapter 9 for two main audiences: 1. DFAT and others who commission and consume the RCA and other approaches to qualitative research, and 2. the RCA core team and those who work with them to conduct research and evaluation.

We invested in engagement and use through establishing a Review Advisory Committee (RAC). The RAC was comprised of senior DFAT officers from the Development Policy and Aid Management Divisions and the Embassy in Jakarta, a representative from Palladium,[10] and an independent

[10] Palladium provides advisory and management services for international development, and has utilized the RCA extensively, both in Indonesia through the RCA+ project, and elsewhere. It was therefore seen to be relevant to have a Palladium member of the RAC.

Qualitative Research Specialist and trained anthropologist. We have engaged with them at key moments and around key deliverables (Inception Report, Field Visit Aide Memoire, and Final Report). We have also had regular (weekly) informal check-ins with the DFAT commissioning team in the Poverty and Social Transfers Section, as well as frequent and wide ranging discussions with the RCA core team. Through these formal and informal touch points, we have aspired to increase the usability, and therefore impact, of this review.

2.5 Structure of this report

This report is organised into 9 Chapters, each with a number of sections and sub-sections.

Chapter 1 is the Executive Summary

Chapter 2 is this Introduction.

Chapter 3 presents the Methodology for this review, and discusses potential limitations to this.

Chapter 4 presents the RCA Problem Analysis and Theory of Change (ToC); places RCA within the landscape of rapid ethnography for development research and evaluation; and provides background and an overview of the RCA, it's emergence, evolution and application.

Chapter 5 provide an overarching Conceptual Orientation to our assessment, looking at rigor and ethics as the two essential pillars of quality research. We use these as the framework for our analysis of RCA in Chapter 6 on RCA praxis, and again in Chapter 9, where we look at the Value for Money of RCA versus comparators.

Chapter 6 assesses RCA against the characteristics of rigor and research ethics that we outlined in Chapter 5. We do this through looking at each stage of the research process: research preparation and design, data collection, analysis, and reporting.

Chapter 7 seeks to answer the questions, "To what extent does RCA deliver on its claims to provide more policy and programme relevant evidence?" and "In what contexts and for what purposes does it work best?" through assessing RCA against a "quality of outcome" framework.

Chapter 8 situates RCA within the range of other comparators in terms of methodological rigor and effectiveness, in order to then assess the value for money of RCA compared to these other potential approaches.

Chapter 9 presents conclusions from Chapters 6 through 8, and provides recommendations for commissioners, and research and evaluation practitioners.

3 Methodology

This chapter sets out the principal methods that we used to collect and analyze data (3.1 and 3.2). We close with a discussion of limitations, challenges and mitigating strategies (3.3).

3.1 Data collection

Our three principle data collection tools were: 1. collection of secondary data through a literature review; 2. key informant interviews; and 3. a field visit to Jakarta where we observed an RCA post-fieldwork debrief and a sensemaking workshop, conducted 2 focus group discussions with RCA researchers, and held a number of key informant interviews with commissioners and consumers of RCA studies; and a short visit to Canberra to hold discussions with key informants from within and outside of DFAT.

3.1.1 Literature review – purpose and approach for sourcing data

Our review of key literature began in Phase 2 and continued until the end of the data collection phase, with some light review extending into the synthesis and analysis phase where we needed to bolster evidence in specific areas.

We used the literature review to gather data in order to:

- articulate of the problem and theory behind RCA (Chapter 4)
- establish a robust rigor and ethics framework for qualitative and ethnographic approaches to research and evaluation (Chapter 5)
- select RCA case studies, document the application of the RCA approach in practice for these case studies in order to better understand they extent to which they are rigorous and ethical, and have been effective in informing policy and programming (Chapters 6 and 7)
- identify possible and select final comparators; articulate and quantify costs and benefits of comparator approaches for VfM assessment (Chapter 8)

In total we reviewed 230 unique sources of information, of which 96 were produced by the RCA team. A full set of documents reviewed can be found in at Annex 2.

More than two dozen discreet RCA studies have been implemented to date in more than half a dozen countries. As this review was unable to delve deeply into all of these, we selected a sample of cases for which to conduct a "deep dive". We reviewed the literature, and then worked with the RCA core team to select case studies that would enable us to assess:

- how the RCA is presently implemented, but also provide an understanding of how the RCA has evolved over time;
- how the RCA performs as a diagnostic vs. evaluative tool, longitudinally or as a one-off study, stand-alone or part of a mixed method approach;
- the applicability and relevance of RCA in different geographic contexts and for understanding different issues.

Another consideration was the familiarity of the review team with the case study geographies and topics. We purposively selected 2 case studies in Nepal, where the Lead Reviewer has lived and worked for over a decade, and other case studies where the review team have deep issue-based knowledge. Our sample was skewed towards Indonesia, given the significant DFAT investment there, including in the RCA. Nepal is second only to Indonesia in terms of the number of RCA studies conducted. We discounted RCA studies that we felt were "outliers" (for example RCA studies with university researchers in Jakarta), as we wanted to understand the extent to which RCA delivers on main claims in regard to poor and marginalised populations.

Through applying the selection criteria above, we selected 6 "deep dive" case studies. We added two additional "light touch" case studies to ensure that we had a better understanding of the RCA's

origins and original intention ("Bangladesh RCA") and its potential use as part of a longitudinal mixed method evaluation (Millennium Villages RCA). See Annex 3 for a detailed description of the intended purpose of these RCA studies.

Table 2 – Case Studies

	Study	Country	Year	Mixed method	Longitudinal or one-off?	Design or evaluation?	Donor/ commissioner
Deep dive" case studies							
1	Rural Access Programme ("RAP3")	Nepal	2015 - present	Y	Longitudinal	Both (integrated into MEL)	DFID
2	The Integrated Programme for Strengthening Security & Justice ("IP-SSJ")	Nepal	2015	N	Longitudinal	Intended to be both (integrated into MEL), but only one study to date	DFID
3	Perspectives of People Affected by Haze from Peatland & Forest Fires ("Haze")	Indonesia	2017	N	One-off	Design - to feed into HCD prototyping of solutions	UNICEF
4	Perspectives and Experiences of the Village Law ("Village Law")	Indonesia	2016	N	One-off	Evaluation – quick feedback on roll out of village law	KOMPAK (DFAT-funded decentralisation program) with part funding from KSI
5	Perspectives, Observations & Experiences of People Living in Poverty on their Household Finance Management ("Household Finance Management")	Indonesia	2016	N	One-off	Design – to inform design of social assistance policy and programming	DFAT/RTI KSI for BAPPENAS (National Development Planning Agency)
6	Adolescents & their Families Perspectives and Experiences on Nutrition & Physical Activities ("Adolescent Nutrition and Physical Activity")	Indonesia	2017	N	One-off	Design - to inform quant survey and design of UNICEF pilot	UNICEF
"Light touch" case studies							
7	Bangladesh Reality Check: A Listening Study. Realities of people living in poverty concerning healthcare & primary education ("Bangladesh RCA")	Bangladesh	2007 - 2012	N	Longitudinal	Evaluation	SIDA
8	Millennium Villages Impact Evaluation, Baseline* Summary Report ("Millennium Villages")	Ghana	2013 - present	Y	Longitudinal	Evaluation (complex impact evaluation)	DFID

*other reports were not reviewed

3.1.2 Key Informant Interviews (KIIs)

During Phase 1 we worked with evaluation partners to identify a list of priority informants. To distinguish between informants' relationship to and knowledge of the RCA, we identified three categories of key informants:

1. **Internal:** Those who have been directly involved in implementing the RCA process or comparator approaches. In this report, we refer to RCA researchers as either "RCA core team

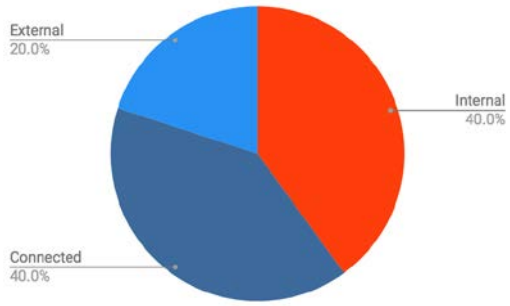
members" – those employed by the RCA Secretariat in Jakarta who set RCA policy and practice as well as conducting research and analysis – or "RCA researchers" – those who are employed on a contract basis to undertake specific RCA studies.

2. **Connected**: Those who have commissioned or consumed RCA and/or comparator approaches but have not been involved directly in implementing these. This includes: DFAT (post and Canberra), other users (e.g. Development programs such as KOMPAK, Mahkota, the Knowledge Sector Initiative, INOVASI, who have commissioned RCA studies), Government of Indonesia officials who have commissioned or been engaged with RCA studies, including from TNP2K, the Ministry of Villages, the Ministry of Health, the Ministry of Home Affairs, the Ministry of Education and Culture and Bappenas, some funders (e.g. DFID, UNICEF, WB), Palladium. These are referred to in the report as "Commissioners" and "Consumers".
3. **External**: Those not involved in commissioning or consuming RCA or comparator approaches, but with an expert view on the uses of mixed method research and evaluation research. These informants are referred to as "Quantitative research experts" and "Evaluation experts". Experts in comparator qualitative research approaches, as well as experienced qualitative researchers (including those who are part of the broader RCA "community of practice"), are referred to as "Qualitative research experts".

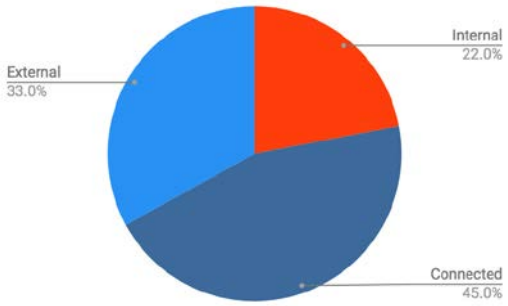
Our goal was to have an approximate split of 40% internal, 40% connected and 20% external. For reasons that we discuss in Section 3.3, below, we adjusted these targets slightly to 25:40:35. Our anticipated and achieved split by unique key informant, and our achieved split by total number of interviews is presented in Figure 2, below. In total, we held KIIs with 77 different people over the course of 88 different interviews. Multiple interviews were held with RCA core team members, as well as a few connected key informants. Annex 4 provides a full list of key informants.

Figure 2

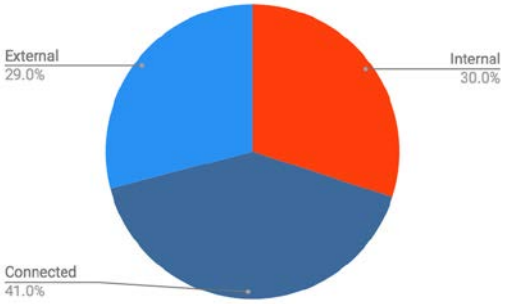
Anticipated split by unique KI



Achieved split by unique KI



Achieved split by total number of interviews



3.1.3 Case study field work

The Indonesia field visit took place between 30 November and 9 December 2017 and had two main aims. The first was to conduct a series of key informant interviews with commissioners and consumers of the RCA. These included Government of Indonesia, staff from DFAT supported projects, Australian Embassy staff in Jakarta, UNICEF and the World Bank. Some of these discussions were wide-ranging; others focused on the context in Indonesia for poverty and social analysis research; yet others focused on better understanding why a specific case study RCA was commissioned, how it had been received, and to what extent it had been used to support more effective policy and programming.

The second aim of the Indonesia case study visit was to better understand how RCA is actually carried out, with a focus on the synthesis and analysis stages of RCA. To this end, we observed an RCA team debriefing and a synthesis workshop, and engaged in a number of other interactions with the RCA team, including a small team workshop, two focus groups discussions with RCA researchers and a number of in-depth interviews with current and former RCA team members.

Case study fieldwork in Indonesia enabled us to deepen our understanding of the four Indonesia case study RCAs through face-to-face discussions and observations. We also sought to put these – and RCA more generally – into the wider DFAT context through a number of key informant discussions in Canberra between 10 and 13 December 2017.

3.1.4 Strength of evidence assessment

We assessed the strength of evidence collected for each Sub-RQs according to the following scale:

- **Strong**: a least 4 different sources, including from both KIIs and documentary sources. KII sources must include both connected and internal informants.
- **Moderate**: at least 4 different sources, but may be *only* from KIIs or *only* from documentary sources. KII sources must include either connected or internal sources.
- **Weak**: three or fewer sources of either type (KIIs or documentary).

At several points during the synthesis and analysis process we assessed strength of evidence, and, where evidence was seen to be weak, sought out additional sources of information to strengthen our assessment. While this considerably increased the time and effort needed to complete data analysis and interpretation, is has also considerably strengthened the reliability and validity of this review. Where we present findings in Chapter 6 onwards, all of these are supported by strong evidence according to the above scale.

3.2 Data analysis

This section sets out our approach to analyzing the data collected. Analysis focused on answering, to the extent possible, the Sub-RQs. We employed 4 main analytical methods to explore these questions; none of these were discretely applied but instead were used iteratively and to enable us to explore different angles of the question.

3.2.1 Theory of Change analysis

The Theory of Change provided our primary frame of analysis to answer the first framing Review Question: "What are the strengths and weaknesses of RCA?" The ToC analysis facilitated a comparison between how the RCA is intended to work in theory (Section 4.2.4) and how it has been applied in practice in various contexts (Chapters 6 and 7). The Theory of Change was also used to look at RCA's effectiveness in the Value for Money Assessment in Chapter 8.

3.2.2 Analysis of rigor and ethics

We developed a rigor and ethics framework in order to provide greater granularity (and comparability) to our analysis of the quality of the RCA in practice. The characteristics of rigor and ethics in this framework were arrived at through a literature review of international good practice in qualitative research, key informant interviews with qualitative research experts, examining how other qualitative research approaches (including comparators) interpret and apply rigor, and reviewing the RCA's own documentation on rigor. This framework is presented in Chapter 5, and applied in Chapter 6 to analyse RCA praxis, and again in Chapter 8, where we examine the Value for Money of RCA versus comparators.

3.2.3 Value for Money Assessment

The second framing RQ asks: "How does RCA compare to other approaches to qualitative poverty and social assessment?" We compared RCA to comparators using a value for money framework.

Comparator selection

The Inception Report presented a mapping of possible comparators that we discussed with the Review Advisory Committee. The report set out three different 'uses' for RCA and potential comparators: evaluative, diagnostic and empathy-building, with the intention of selecting one main comparator for each use. During Phase 2 of the review we used the below criteria in order to make a final selection:

- **Identifiability:** How easily identifiable is the approach as a discrete, bounded approach? Is it commonly understood as an approach to qualitative poverty and social analysis?
- **Comparability:** Has the approach been used for evaluative, diagnostic and/or immersive objectives, i.e. can it been situated within the RCA theory of change?
- **Data availability:** has the approach been well-documented and has its effectiveness been assessed? Can we get costings of the approach?

While we were keen to include a broad range of relevant comparators, we needed to keep the number of comparators small (ideally no more than 4) to ensure sufficient depth and address potential bias from taking a microscopic look at RCA but only a birds-eye view of other approaches. The approaches selected as final comparators based on the above were: Human Centred Design Research, SenseMaker™, PEER, and "good standard" mixed method qualitative research.

Comparator data gathering

We approached the data gathering for the comparators in a similar way to RCA, but on a much more limited scale. We gathered literature for each, including articulation of the approach, research outputs (with 3-5 case studies per approach), and both independent evaluations and costing data

where available. We also conducted KIIs with practitioners and commissioners, with between 1 and 4 KIIs for each comparator.

Comparator assessment

Although we initially envisaged that each compactor would map to one 'use' of RCA, in practice we found that the selected comparators were often comparable across 'uses', whether intentionally designed to do so or not, so a one-to-one mapping in that way would not necessarily capture the full range of either RCA or the comparators. Instead, we were able to apply the same approach to assessing the comparators as we employed for RCA, using the same Theory of Change that we introduce in Chapter 4, the framework for assessing rigor that we introduce in Chapter 5, and the criteria used for assessing the quality of outcomes in Chapter 7. The aim of this approach is to ensure transparency and objectivity in the comparison so that it is fair and balanced.

Value for Money methodology

Unfortunately there is not enough available data on either costs or quantifiable outcomes and impacts for RCA or the comparators to undertake a thorough and conclusive VfM assessment. The focus here is instead on using the VfM framework to draw out initial findings on the '4Es' of economy, efficiency, effectiveness and equity. What cost data is available comes from KIIs and, for some of the comparators, external evaluation reports.

3.2.4 Exploring utility

Key to answering the third framing RQ – "What is the potential relevance of RCA and similar poverty and social analysis approaches to the broader Australian aid program?" – is assessing utility. We assessed utility in 2 main ways: 1. through specific case studies of how RCA has been applied; and 2. through discussions with DFAT staff and others commissioners and consumer who have used poverty and social analysis evidence. Specific cases of where the RCA has attempted to influence policy and program investments were identified through the literature review, KIIs, and the Indonesia country visit. We use these in Chapter 7 to highlight the extent to which the RCA has been successful in using qualitative research evidence to influence policy and programming.

3.3 Limitations and mitigating strategies

During Phase 1, the review team identified a number of important factors that we believed needed to be considered in how we framed and conducted the review. We reflected these factors in our review design, and have reviewed them at key points in order to better understand the implications of limitations and challenges as they played out in practice, and ensure that we were mitigating these as much as possible.

How the RCA has been used has evolved over time. Most recently, the RCA+ Project – supported by DFAT and implemented by Palladium - has been set up in Jakarta, with core team members leading trainings and studies internationally. While this presents a challenge in having to assess a moving target in terms of both theory and practice, it also provides us with opportunities for inquiry, for example on why adaptations have been necessary and implications for the future design and implementation of RCA.

Many different stakeholders (internal and external stakeholders) and individual staff members have been involved in the process of applying the RCA. This gave us a good pool from which to draw our key informants; however it also created a challenge in terms of the number of key informants we could potentially interview. While we did conduct some careful prioritization, we found the data to be so rich from the KIIs that we increased the number of key informants from a target of 40 to a total number of 77. This was considerably more time and labour intensive, but has yielded rich and robust results.

A limitation with our original 40 internal: 40 connected: 20 external split of key informants was that it had the potential to unfairly weight KII evidence towards positive findings, in that "internal" KIIs were conducted with RCA core team members and RCA researchers. While this is important for understanding how the approach works in practice, it had the potential for generating highly subjective and biased data. We therefore, in discussion with DFAT, decided to reduce the proportion of "internal" key informants and expand the number of interviews with qualitative and mixed method research experts who were familiar with RCA and a broader range of research and evaluation approaches.

There were limited opportunities to see RCA "in action". This created challenges in terms of understanding how RCA researchers are trained, and how debriefing, synthesis and analysis are conducted. We mitigated this to some extent during the in-country case study in Indonesia by observing a day of debriefing and a day-long sensemaking workshop. We also reviewed a large amount of internal RCA documentation on these issue, as well as exploring these issues in depth during KIIs with RCA core team members and RCA researchers.

Further, we were unable, for ethnical reasons, to speak with host households. In order to mitigate this limitation, we conducted an extensive review of secondary data, including other reviews that did talk to hosts, and explored this through key informant interviews with RCA researchers who shared their views on the reactions and experiences of host households.

Selecting comparators was challenging. The relatively unbounded nature of some approaches and their varied application in different contexts makes it difficult to (a) define approaches neatly; and (b) find appropriate – and representative - comparators. We managed this by developing a two-phase process to screen comparators, where the first phase screened against general ethnographic qualitative research attributes, and the second screened against identifiability, comparability and data availability (see 3.2.2, above). We bolstered a literature review of comparator approaches with key informant interviews with practitioners and evaluators of these approaches, and have applied the same methodological frameworks for assessing them as we do for the RCA.

Without a full evaluation of these other approaches, however, we are aware that there could be a slight bias towards other approaches, given the inevitable divergence in the thoroughness of the assessments for comparators vs the RCA. While this might not be true of "general qualitative research" which we use as a comparator - due to the fact that both Reviewers have considerable field experience designing and implementing a range of different types of qualitative research - if we were able to do the same kind of field assessment for the other comparators as we were for the RCA, we might find additional issues with the comparators that do not emerge from KIIs or independent evaluations. Recognising this potential imbalance due to the inherent divergence in the depth of the assessments of the comparators, we are careful to contain the findings here, so that the comparators are used to situate RCA within the range of potential options that could be used in similar research contexts, and the range of likely outcomes that might be achieved.

4 RCA: the problem and the theory

4.1 What is the problem RCA seeks to address?

The RCA was developed to respond to the problem that the voices and perspectives of the poor are not heard by policy-makers or those involved in the development, monitoring and evaluation of programs. This problem, according to RCA practitioners, is rooted both within the wider context, as well as within other poverty and social analysis methodologies.

In terms of the wider context, the RCA is proposed as part of a response to the following root problems:[11]

4. Policy and program officers lack the knowledge and capacity to commission and consume qualitative data and risk aversion towards non-expert data and new forms of knowledge. The result is that policy and program officers rely disproportionately on quantitative research which fails to pick up key insights into how and why change happens.[12]
5. People are unwilling or unable to voice their perspectives and experiences because there are few spaces in which to do this, and a lack of trust and self-confidence, manifested in a belief that their views will not be respected or acted upon. This results in limited opportunities for intended beneficiaries to influence policies and programs.
6. There are weak on-going feedback loops during implementation, reducing the speed with which research findings enter the public domain and opportunities for program adaptation.

These problem are well-articulated in the development literature. Lewis writes,

> *A persistent theme in the history of international development – and within the analysis of policy making and implementation processes more widely – is the problem of the insulation of so called ''policy makers" (in the context of the RCA this means high-level government personnel, international donor staff, and senior program planners) from the realities faced by the people whose problems their policies are supposed to address (2018: 17).*

Lewis, and others (Chambers, pers. comm.; Eyben, 2013) suggest that this problem is getting worse, and that the past decade has seen development agencies become less interested in the human and social aspects of development and more concerned with approaching development as a technical and managerial process.

Many would agree that policy-makers and development practitioners far too often find themselves identifying and planning to solve problems which may or may not be perceived of as problems by those that programs are trying to benefit. They also fail to seek or support local solutions, which is both erroneous and dangerous: "*we understand far less about what those in the Global South are doing than we think, and…our assumptions about life in such places are a) mostly incorrect and b) potentially very dangerous in to the well-being of everyone on Earth*" (Carr, 2018 in Shah, 2018: 5).

There is good evidence that richly textured interpretations of why services are being delivered or not - or delivered *well* or not - matter for improving planning as much as the quantitative questions do.

[11] These were identified together with the RCA core team as part of a process of joint problem analysis for this Review. These three factors reflect those identified by Lewis (2018:17) as those that underpin policy makers detachment from "the people who experience their policies writ large in their own lives": 1. Geographical remoteness, 2. Socio-economic remoteness, and 3. A dominant ideology that privileges quantitative data and positivist knowledge over qualitative data and knowledge rooted in a recognition of "the messy complexity of ordinary people's experiences" (Greene, 2009:9 in Lewis 2018:17).

[12] see also Shah, 2018; and Bell and Aggleton, 2016.

Getting policy and program officers to recognize this is indeed the Achilles heel of qualitative research, and a problem that many qualitative approaches attempt to address, with varying success. RCA is in fact not alone in recognizing the lack of voice of beneficiaries, and approaches such as participatory poverty assessments, beneficiary assessments, participatory action research and listening studies have emerged in response to this, in addition to more traditional qualitative approaches.

RCA has also emerged and evolved as a response to perceived problems in other qualitative approaches to poverty and social analysis. These perceived root problems are:

3. Single-sector, narrow lens, linear research and evaluation approaches predominate over systems-based, open-ended, contextual approaches.
4. Inherent researcher bias at design, data gathering, and interpretation of design, monitoring and evaluation research often remains unacknowledged and/or unaddressed.

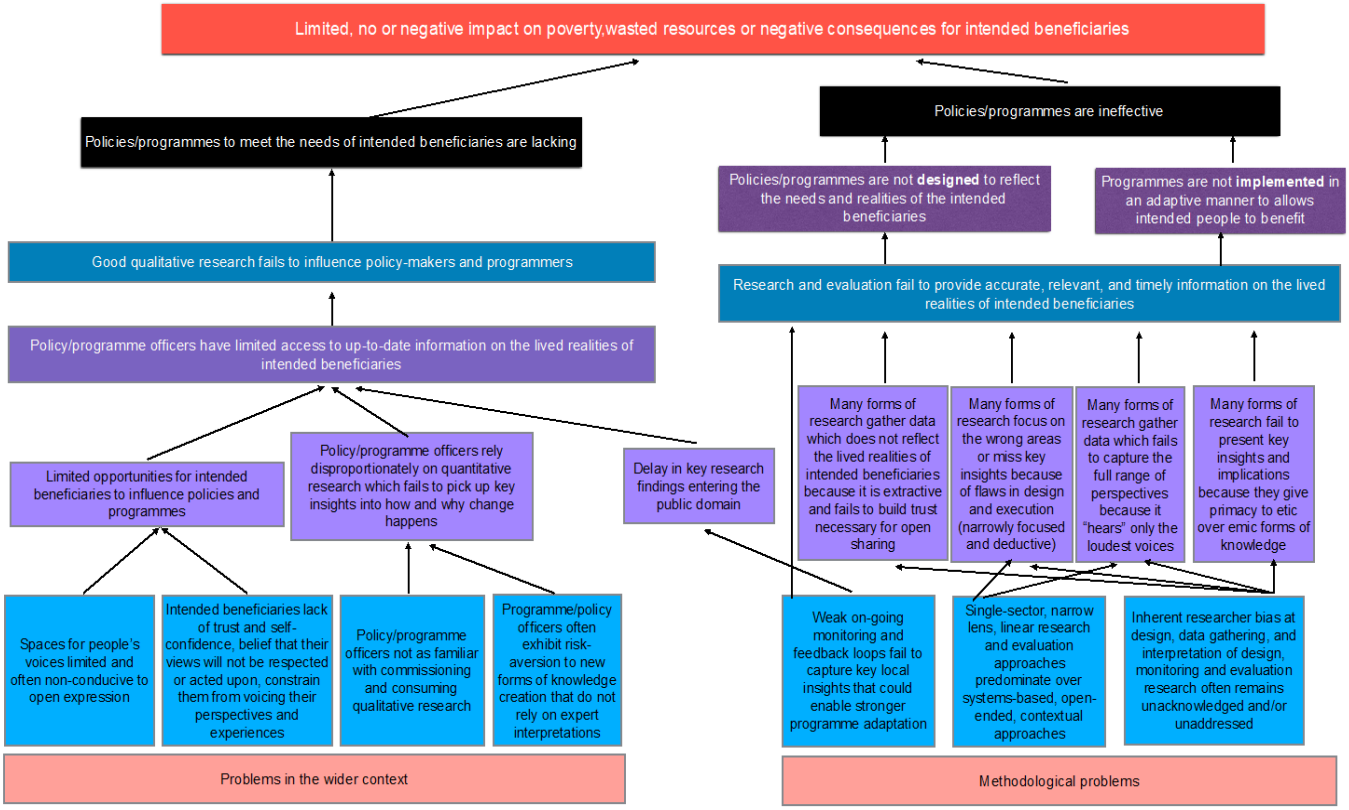
The RCA critique suggests that there is a tendency to ignore the everyday experiences in which the poor live, which misses key insights into their rationale for behavior, and/or leads to a focus on the wrong areas, with serious consequences. The RCA levels a particularly strong critique of theory-based (deductive rather than inductive) approaches, which impose researcher bias and fail to generate grounded theory, and of "etic" (outsider) interpretation over "emic" (insider) interpretation. Approaches that are focused on "finding out" rather than "learning" are seen to be conducted in a way that limits trust-building and fails to address power imbalances or encourage open sharing of issues both directly and indirectly relevant to the issue being researched. Similarly, approaches that use etic over emic interpretations are seen to fail to present the insights and implications that could have the most profound influence on policy and programming.

According to RCA's understanding of the problem, together these wider context and methodological challenges have contributed to a situation in which,

> *We've become more and more detached from understanding context, and it is important to increase understanding of how people live their lives, make choices, etc. We need to respond to that rather than our own normative views of what people need. We [RCA practitioners] are pushing back against received wisdom. People are so far removed from reality, and there are so many layers of reporting with so much vested interests embedded. This is particularly true for senior policy makers* (KII, RCA core team member).

This diagnosis suggests that the ultimate result of these failings is serious, in that either programs and policies to benefit the poor are lacking altogether, or they are ineffective in design or implementation. This, in turn, leads to limited - or even negative - impact on poverty, wasted resources or negative consequences for intended beneficiaries. This problem analysis is illustrated in Figure 3, below.

Figure 3 – RCA Problem Analysis



The crux of the rationale for using the RCA thus rests on both contextual and methodological failings. Assessing the accuracy of this analysis, and the effectiveness of RCA as a "solution" to these problems, is therefore central to this review.

4.2 How does RCA present a "solution"?

In this section we start with a brief discussion of the role of ethnographic research in development, in order to situate RCA within a broader context. We then provide a brief description the reality check approach and discuss its origins and evolution. We finish this section with a discussion of how RCA, in theory, presents itself as a credible "solution" to the problems outlined above.

4.2.1 The role of ethnography in development research

Few would disagree the part of addressing the above described problems requires development research, design and monitoring and evaluation to be much more engaged with qualitative research findings, including those generated by ethnographic approaches. Ethnography is a well-established approach to research known to be capable of producing exactly the kind of rigorous evidence which is currently lacking in development. As outlined by Shah (2018: 6)
ethnography can enable researchers to:

- gain the trust of hard-to-reach populations, thereby countering issues of exclusion in sampling
- understand and communicate research participants' perspectives and interpretations, even when they differ dramatically from those of the researcher or the researcher's intended audience
- analyse people's experiences on their own terms, rather than according to predetermined, biased categories or projections

- produce unexpected and counter-intuitive findings, rather than being constrained by predetermined assumptions
- analyse social phenomena – including development interventions and institutions – holistically and in context, thereby addressing issues of self-referencing research
- find out what people do as well as what they say they do in everyday life
- collect data about sensitive topics
- interpret behaviour that has otherwise made little sense to development practitioners and policy makers

The drawback of ethnography, of course, is that in its full, traditional, form it is too slow and expensive to do at a large enough scale to be feasible in most development policy and programming contexts. The use of "rapid" or "light" ethnographic approaches in development is, however, by no means a new endeavour, [13] and cannot claim to be an "innovation" (as is often claimed of the RCA by its core practitioners). The use of ethnographic approaches in development spans a number of disciplines, but all have in common attempts to adapt ethnography to get similar – but more accessible, applicable and cost-effective - results more quickly (Chambers, 1994; Shah, 2018).

Some purists argue that rigour cannot be maintained without the long-term immersion, observation, participation, reflection and relationship-building that enables ethnographers to rigorously analyse and interpret complex social phenomena. However, many development practitioners argue instead that there is indeed a place for rapid ethnography-inspired methods, even if they are recognised as being a necessary 'second-best' alternative to the gold standard of a full ethnographic inquiry, provided that the limits of these are recognised, and the risks mitigated using credible strategies (Cernea, 1992 and Bernard, 2011, in Shah, 2018).

This point is crucial to the evaluation, and worth emphasising further: ultimately it is methodological rigor that allows these 'light' ethnographic approaches to, firstly, acknowledge potential limitations throughout the research process that might interfere with the ability to get an 'accurate' view of (a) whose voices are/should be heard and (b) what their views/perspectives are. Risks to achieving these objectives include the ability to gain trust in a short time and the ability to interpret correctly verbal and non-verbal responses or the nuances of relationships and power dynamics without a deep understanding of the local dialect or socio-cultural context. Risk mitigation strategies are then the second aspect of methodological rigor, providing a way to transparently and reflexively address the potential risks. This careful approach can allow even 'light' approaches to understand the context in which behaviour change occurs; and the complexity of distinguishing between normative statements and actual lived experiences.[14]

4.2.2 The Reality Check Approach

As noted above, efforts within the participatory tradition to better link people's local experiences with policy processes are not new. During the late 1990s, World Bank Participatory Poverty Assessments (PPAs) and Beneficiary Assessments tried to co-construct participatory knowledge with poor people to inform project and policy design (Lewis, 2018). The World Bank's Voices of the Poor (Narayan et al., 1999) and Sida's Views of the Poor (Jupp, 2007) were both attempts to move more ''people-centered" research knowledge into ''upstream" development policy settings and agencies. However despite the growth of interest in ''people centered" approaches to development among many development agencies in the latter part of the last century - and more recent attempts such as the

[13] See, for example, Rapid Ethnographic Assessment (REA) (e.g. Taplin, Scheld and Low, 2002), Rapid Ethnography (e.g. Millen, 2000), Focused Ethnographic Study (FES) (e.g. Pelto et al., 2013), Rapid Rural Assessment (RRA) (e.g. Chambers, 1981), Participatory Rural Assessment (PRA) (e.g. Chambers, 1994a, 1994b, 1994c), Quick Ethnography (QE) (Handwerker, 2001), short-term ethnography (Pink and Morgan, 2013), micro-ethnography (Spradley, 1980) and mini-ethnography (Leininger, 1985).

[14] See, for example, Maina, 2017.

Portfolios of the Poor approach (Collins et al., 2009),[15] "listening studies"[16], and immersion research[17] - the last decade of development has seen a return to focusing on deploying resources effectively and securing measurable results (Lewis, 2018), rather than developing the deeper understanding of other people and their lived realities that is essential to the design of progressive programs and policies that meet the actual (as opposed to assumed) needs of those experiencing poverty (Shah, 2018). It is from this genealogy that RCA emerged.[18]

RCA combines elements of rapid ethnography (living with people, usually those who are directly experiencing poverty or any other phenomenon being researched) with 'light touch' participant observation. As with other approaches to "people centred research", such as listening studies and beneficiary assessment, the focus of RCA is on engaging with, listening to, observing, and documenting the voices, opinions, and experiences of people, and asking questions as a curious learner as part of a relaxed conversation. Due to the shorter timeframe than most ethnographic investigations, there is necessarily a stronger emphasis on conversations than on observing behavior and the complexities of relationships. There is a strong emphasis on researchers as conduit rather than intermediaries, understanding people's lives in context.

The Reality Check Approach in brief

The Reality Check Approach (RCA) is a 4 day and 4 night immersive ethnographic approach in which teams of researchers live with families in study communities. Team sizes vary between 9 people across 3 locations in the smallest RCAs and as many as 30 people across 10 locations in the biggest RCAs. The 3 – 4 members of a sub-team are dispersed within each location to enable them to both meet different people but also triangulate location data. Over the course of the immersion, the researchers typically engage with hundreds, and sometimes thousands, of people; while some of these conversations are brief and cursory, others (particularly those with host household members and a small number of neighboring or "focal" households), are more in-depth.

The emphasis is on relaxed, informal, participant-led interactions. Researchers try to pose questions in an indirect manner so that participants do not feel that researchers are trying to seek a specific answer or are being judgmental. RCA practitioners argue that by asking indirect questions, and not being overly "hooked" on a particular agenda, the biases found in some other qualitative research can be reduced.

Lewis et al. (2012) and Masset et al. (2016) write about the 4 key principles that underpin the Reality Check Approach and that they suggest sets it apart from both quantitative approaches to research, and many other qualitative approaches. These are:

1. **Depth:** the RCA aims to document the experiences and perceptions of poor people in fine-grained detail. Some RCA studies attempt to achieve greater depth through their longitudinal nature which, according to Masset et al. (2016), helps capture changes, including in social norms.
2. **Respect for voice:** the basic idea of the RCA is to listen to what people have to say about their situation, and attempt to document people's views in ways that allow their voices to be properly heard by those higher up in the policy system, in government, donors or NGOs.
3. **Flexibility:** team members do not need to stick to a set question format, sample or schedule, making it possible to follow up and cross-check what people say, and to respond flexibly to new and unexpected information.
4. **Simplicity:** the RCA is intended as a simple, direct and immediate type of 'pulse taking'. As with many other approaches to qualitative research, it aims to use a less complex, 'light touch' approach than those used in the large-scale surveys common in quantitative research or

[15] Collins, Morduch, Rutherford, and Ruthven, 2009
[16] See, for example, Anderson's, Brown, and Jean, 2012, Time to Listen, which reported on how people think about international development assistance, drawing on a thousand informants across a multi-country study.
[17] See Chambers, 2007 and Chambers, 2012.
[18] For fuller accounts of RCA's genealogy and history, see Lewis et al., 2012 and Lewis, 2018.

evaluation. Lewis et al. also claim that it uses less time if compared to the long duration required by most forms of qualitative anthropological ethnographic fieldwork.

RCA proponents and many practitioners claim that this informal and exploratory approach helps researchers to set aside their biases and agendas, access the perspectives of their hosts through listening deeply, and capture counter-intuitive insights which are both hidden from other forms of inquiry, and critical to the development of more relevant and effective development projects and policies.[19] The RCA approach promotional materials make much of RCA's ability to reduce, or even eliminate, bias altogether. They claim that, while "*Others have project bias or agenda bias. We don't. Our only allowable bias is to take the position of people*".[20]

4.2.3 Emergence, evolution and application of RCA

Emergence

The first RCA study was established by the Swedish Embassy in Dhaka and by Sida headquarters in Stockholm. It was conducted was in Bangladesh starting in 2007 as a 5 year project, with the same researchers returning to the same households every year over the course of 5 years. The RCA used informal conversations and observations with service users at community level to construct a form of participatory, ethnographic policy knowledge. The aim was to gain a "fleeting glimpse" into how people were accessing and experiencing health and education services, to supplement formal monitoring and evaluation data with "people-centred data" and thus to close the gap a little between people and policy makers by illuminating the lived experiences of poor people (Lewis, 2018; Pain, Nycander and Islam, 2013).

Following the initial RCA in Bangladesh, RCA was introduced in Indonesia in 2009/10 to provide insights into how activities under the Australian Government-funded Indonesia Basic Education Program (BEP) had been experienced by people living in poverty. In April 2014, the DFAT-funded RCA+ plus project was launched. The project sought to build the capacity of Indonesian researchers and research organizations to undertake RCA studies and develop a tradition of quality, people-centred qualitative research (RCA, 2017a).

The purpose of this review is not to assess the outcomes achieved by the RCA+ Project, but rather to look at the value of the RCA itself, as a qualitative research method. The RCA+ Project, however, has had an impact on the ability of the RCA team to further build, test and refine the approach. In addition to Sida and DFAT, RCA studies have been commissioned by DFID, Swiss Agency for Development, UNICEF, World Bank, Asian Development Bank, Monash University, Family for Every Child, and Voluntary Service Overseas.

Evolution

As a research method, as with many other qualitative methods, RCA has evolved since its initial application as a complement to other, more formal, longitudinal methods of monitoring and evaluation. While it remains the central intention of the RCA to rapidly learn and get feedback from people "on the ground" in order to test commonly held assumptions (to see if they are in fact "true") and provide new information that might question or challenge conventional wisdom (Lewis et al., 2012), RCA is now far more ambitious than its first, relatively modest, beginnings as a supplementary approach, aimed at "getting an inkling" about ground realities (KII, Qualitative Research Expert, quoting Greene, 2009). In keeping with greater ambition as an approach, over the past 4 years the RCA has been applied in a number of ways, including as a diagnostic tool, as a "pulse taking" situational assessment, and as an evaluation method. The methodology, however, has remained relatively constant despite evolution in application.

[19] See also Shah, 2018.
[20] http://www.reality-check-approach.com/related-resources.html

Application

Some RCA studies have been one off "pulse taking" studies, used to gather quick feedback on the roll out of a program or policy, or to better understand the impact of policy changes or large events (such as the earthquake in Nepal).[21] Two of the review case studies - Village Law and Household Financial Management – are considered to be pulse-taking studies, although it could also be argued that the purpose of the Village Law study was to inform governance interventions more broadly through KOMPAK, DFAT's Indonesia Governance for Growth Project.

Other one-off studies have been diagnostic, rather than pulse taking, used to inform policy formulation or the design phase or early implementation of a program or activity. This is true of two of our case study RCA studies: the Haze and Adolescent Nutrition and Physical Activity RCA studies in Indonesia and the IP-SSJ study in Nepal. DFAT has tended to use RCA for one-off diagnostic studies, to (as one RCA practitioner put it) "*prevent them from making early mistakes, for example not using mobile phones as a feedback mechanisms if people don't have mobile phones.*"

Evaluative longitudinal studies – where researchers go back to live with the same families every year or two (such as with the first RCA in Bangladesh) - are rarer, and are increasingly integrated within mixed method Monitoring, Evaluation and Learning (MEL) components of projects. These longitudinal studies are primarily aimed at informing programming rather than policy – although these programs themselves often have links to broader policy in country. Our case study examples include several RAP 3 RCA studies, as well as the "light touch" Millennium Villages Evaluation.

In addition to the above expansion in application, there have been 3 other major shifts in how the RCA is being used. First, there has been a stronger attempt at using the RCA alongside other methods – primarily quantitative approaches to poverty and social analysis, but also the use of visual methods such as participatory video and, more frequently, Digital Story telling (DST). Second, the RCA+ project has also enabled a stronger focus on policy in Indonesia, working with partners (such as UNICEF and TNP2K) to identifying policy moments where an RCA study could add value, and building relationships with key policy makers. Finally, there have been a few "experiments" in using the RCA approach to understand the lived realities of other constituencies, such as the urban poor, and university lecturers, rather than to better understand the lived realities of the rural poor (and those providing services to them).

This review will comment on RCA's effectiveness as a pulse-taking, diagnostic and evaluative research approach, as well as looking at how effectively it has been combined with other research methodologies.

4.2.4 RCA Theory of Change

> *RCA is really trying to put people's voice to the forefront…We are trying to communicate their voices as they are, unfiltered. We are the bridge between people and those who are making programs and policies for them. If we are able to do that, better policies and programming will result*" (KII, RCA core team member).

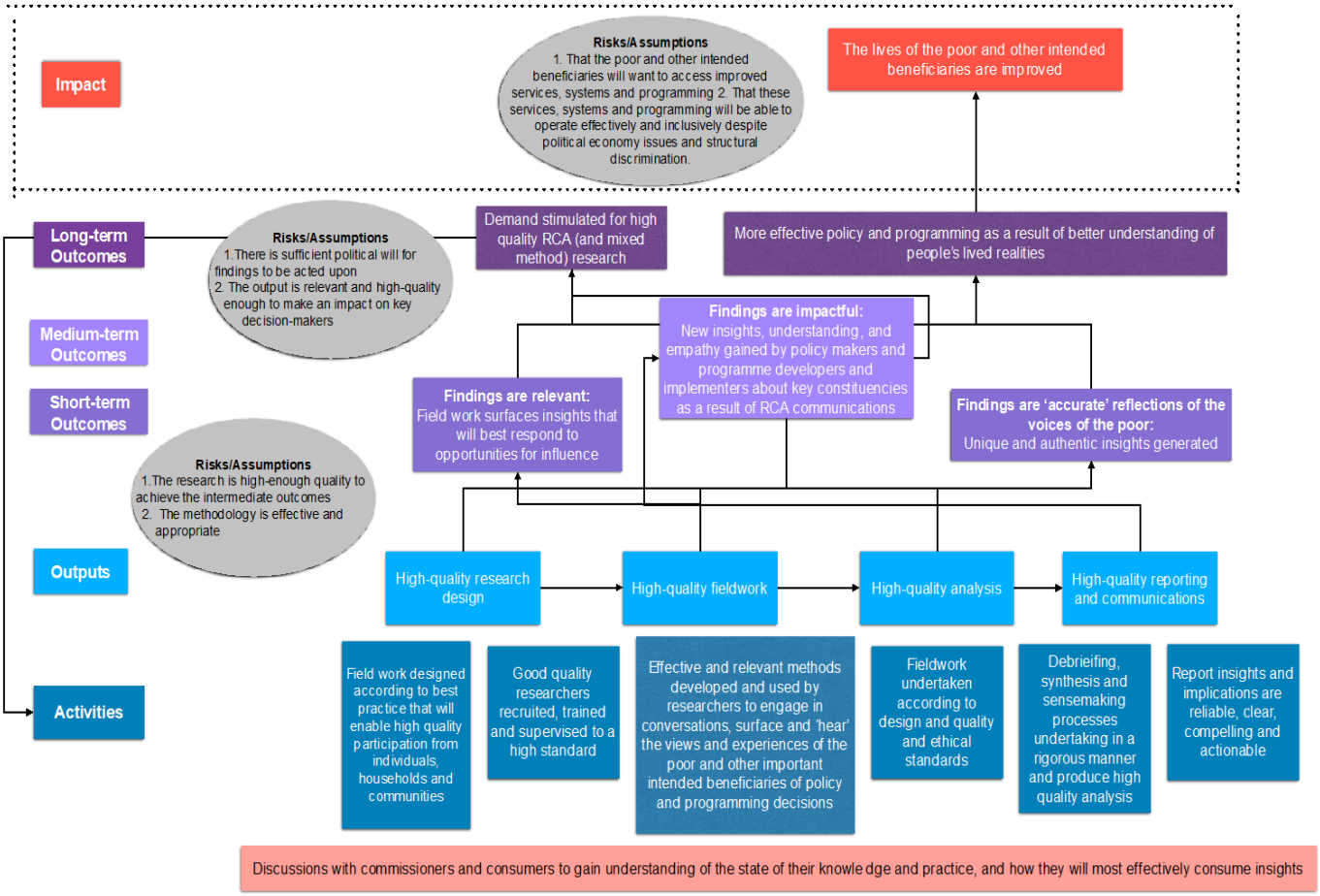
In Section 4.1, we described the problem that the RCA seeks to address. In sections 4.2.1 to 4.2.3, we described the approach in brief, situated it within a broader ethnographic research tradition; and provided an overview of how the RCA started and evolved, and the range of ways it has been applied. In this section, we explore in more depth RCA's theory of change – not only what it is trying to do, but why and how.

In the context of this review, the overarching objective of RCA is to surface accurate, unique and authentic insights that will best respond to opportunities for influence and to communicate these to

[21] Some examples of these studies are the Social Assistance studies conducted with TNP2K in Indonesia. For these studies, TNP2K wanted quick feedback on people's perceptions of the roll out of changes to some of the social assistance programs).

policy makers in a way that generates understanding, and empathy. It is theorised that this deeper understanding by policy and program stakeholders of whether and how their policies and actions translate into effective change on the ground, and how these efforts and changes are perceived, will not only inform but will influence future policy and practice so that it is better geared to local needs and context (see Masset et al., 2016 and Lewis et al., 2012). These changes (situated within the purple short-term outcome boxes in Figure 4, below) we consider to be fully within the RCA's sphere of control. The medium- and longer-term outcomes are within the RCA's sphere of influence (they depend also on the actions of policy-makers), and the impact level change is within the RCA's sphere of interest (the ultimate goal, which also depends on many other external factors).

Figure 4 – RCA Theory of Change[22]



As we see from the above Figure 4, RCA attempts to achieve short and medium-term outcomes through 4 main outputs – which correspond to 4 main stages of research - and their contributing activities:

5. **High quality research design** - Field work designed according to best practice that will enable high quality participation from individuals, households and communities.
6. **High quality field work** - Good quality researchers recruited, trained and supervised to a high standard; effective and relevant methods developed and used by researchers to engage in conversations, surface and 'hear' the views and experiences of the poor and other important intended beneficiaries of policy and programming decisions; fieldwork undertaken according to design and quality and ethical standards.

[22] The RCA Theory of Change was developed together with the RCA core team.

7. **High quality analysis** - Debriefing, synthesis and sensemaking processes undertaken in a rigorous manner and produce high quality analysis.
8. **High quality reporting** - Reliable, clear, compelling and actionable report insights and implications generated.

We will be exploring the extent to which these outputs are achieved, how and why in Chapter 6. First, we outline our conceptual approach to assessing rigor and ethics in Chapter 5.

5 Conceptual approach to assessment

This section will provide an overarching conceptual orientation to our assessment, looking at rigor and ethics as the two essential pillars of quality research.[23]

5.1 Methodological rigor in qualitative research

As a concept, rigor is perhaps best thought of in terms of the quality of the research process; a more rigorous research process will result in findings that have more integrity, and that are more trustworthy, valid, plausible and credible (Given, 2008). Contrary to what RCA practitioners claim, we suggest that no research methodology has "intrinsic rigor" (RCA, 2017a). There are a number of features that are thought to define rigor in qualitative research (which differ from those found in quantitative research). In this chapter, we outline characteristics, and some of the ways in which qualitative researchers attempt to maximize these features in their work. In Chapter 6, we assess RCA against these.

5.1.1 Rigor in research preparation and design

Experienced, reflexive and well-trained researchers

In putting together a research team, those leading qualitative research exercises generally look for researchers with the following skills and experience:

1. Previous experience conducting qualitative research (and sometimes using the research methods that will be deployed, for example participatory approaches, or focus group discussions);
2. Good knowledge of the issues to be researched, for example nutrition, or security and justice;
3. Previous work and/or living experience in the geographies that the research will take place (preferably within the country or specific area of the country where the research will take place);
4. Deep professional or personal experience with the populations who are being researched (this might be "women" or "fishing communities" but becomes even more important with vulnerable groups, such as "children" or "sex workers"); and finally,
5. Language proficiency and cultural fluency, which usually – but not always – suggests the inclusion of strong local researchers on the team.

While it is unlikely that any one team member has all of these different characteristics, at a minimum a team should cover them all, and each individual member should have a number of these attributes.

Rigorous qualitative studies are built on the notion of reflexivity, the process of examining both oneself as researcher, and the research relationship. Researchers should be trained to interrogate their own and other team members biases and blind spots, and research design should explicitly take account of their perceived role, motives, and power, and how all of this might influence the research process and results. In analysis and report writing, researchers need to account for the fact that their presence has some influence on the research findings, and should attempt to report how they, as the primary research instrument, may have influenced the study's results (Given, 2008). Reflexivity is thus woven throughout the entire research process and will be addressed in several different sections of this report.

Contextual understanding

Before proceeding with data collection, those designing research need to ensure that they have a good understanding of the issues, people, and geographies where the research will take place. Any gaps in knowledge should be filled in by conducting secondary research (including interrogating this

[23] See also Shah (2018), who also used the twin pillars of rigor and ethics to assess the Papua Education RCA Study.

"knowledge" and the assumptions underpinning it). Contextual understanding is critical for identifying what is "known" (and how this is known, in order to check these assumptions) and what is not known, and in developing contextually appropriate approaches to data collection. Within qualitative research, it is always important to have a good understanding of local social relations and power dynamics relating to gender, caste, language, ethnicity, age, etc. Failure to understand this can lead to poor quality research design, execution and analysis.

Framework to guide inquiry

Building on a strong contextual understanding of the context and the issue, ethnographic research is guided by conceptual (rather than theoretical) frames,[24] to enable researchers to maximise the collection of relevant data, while still recognising that an open, exploratory orientation is required in order to enable participants in research to define "relevance" and to enable surprising data and connections to emerge.

Multi-disciplinarity

Teams that include experienced qualitative researchers from a range of disciplines - political scientists, agronomists, and sociologists; designers, behavioral scientists and anthropologists – can help to surface different insights and perspectives and challenge biases in analysis and interpretation, all of which strengthens research design and analysis.[25]

5.1.2 Rigor in fieldwork

Being unobtrusive

The less the researcher disturbs the scene, the longer spent in it, and the deeper the penetration of the research, the more it might be claimed to be authentic. It takes significant amount of time, however, for subjects to stop 'playing up' to the researcher, doing different things, or things differently because the researcher is there.

Triangulation

The principle of triangulation is to examine the same issue through different lenses, using different methods and multiple observers to increase confidence in the findings. It counters criticism of single method, single researcher and single theory bias (RCA, 2017a). The most common forms of triangulation in qualitative work are:

1. **Method** - The use of several methods to explore an issue increases the chances of depth and accuracy. One of the commonest forms of triangulation is to combine interviews and discussions with observation. Observation will test and fill out accounts given in interviews and discussions, and vice versa.
2. **Time** - This allows for the processual nature of events. For example, one might wish to better understand community governance, so while it would be prudent to observe a community meeting, a deeper understanding might be achieved if the researcher were to: a) discuss with a range of community members the content and process of these kind of meetings; b) observe the meeting as it happens; c) discuss with a range of community members afterwards what had happened and why. Further, reporting at one single moment in time can be normative. For example, people will report when asked what they usually do, but if one were to observe actual behavior over time, one might see that reality diverges significantly from the perceived norm.
3. **Persons** - This might involve consulting a range of people, perhaps in different roles or positions. Having multiple researchers conducting investigations also aids triangulation, as

[24] A conceptual framework is based on key concepts and the relationship between those concepts. A conceptual framework embodies the component parts and scope of research and suggests how an overarching research question will need to be explored. It provides a context for the research and helps to explain observations. A theoretical framework lies on a much broader scale of resolution and rests on time tested theories that embody the findings of numerous investigations on how phenomena occur. A theoretical framework systematically tests theories, seeking to explain the relationship between phenomena for purposes of explaining, predicting and controlling.

[25] See also Shah, 2018.

this can enable researchers to check each other's accounts for differences, lead to closer interrogation of data (even revisiting the site to collect more data) and enable more accurate development of theory.

Respondent validation

If we are aiming to understand the meanings and perspectives of those being studied, how better to judge if our understandings are accurate and full than by giving our accounts back to those involved and asking them to judge? This increases validity and credibility (Mac an Ghaill , 1994; Given, 2008). This is not only about 'capturing what it's like', it is about grasping the experiences, feelings, problems, and managing to convey them authentically. The importance of involving research participants in analysis is central to some qualitative approaches,[26] and indeed was the driving force behind the evolution of Rapid Rural Appraisal (RRA) to Participatory Rural Appraisal (PRA), which Chambers describes as moving from "extractive to empowering", and with researchers going from "investigator to facilitator" (1994: 958).

Faithful and accurate recording

Faithful and accurate recording of experiences, conversations and observations increases the reliability and validity of qualitative data. Recording data objectively and comprehensibly, including the use of note taking, audio, video, photographs, drawings, and different levels of detail in the transcription of data are the main ways in which this can be achieved.

5.1.3 Rigor in analysis and reporting

In addition to continued vigilance in regard to elements of rigor in relation to research preparation, design, and fieldwork, researchers strive for the following elements of rigor in analysis.

Building, testing and evidencing hypotheses in data analysis

There are several approaches for analysing qualitative data after fieldwork that ensure that generalizations are supported by adequate evidence – that they are reliable and dependable – and that enable analytic induction and grounded theory to emerge. These include the following:

1. **Iterative analysis in the field** - Typically with qualitative research, analysis is highly iterative begins in the field. Researchers, as they are writing and reviewing notes, reflecting on the day, and preparing for the following day, start identifying emerging patterns as well as outliers to these patterns, and developing early and tentative hypotheses to test in further research. Hypotheses often best emerge based on local analysis, through asking people "why do you think that is?", and through placing local emic interpretations within the researcher's own landscape of understanding.
2. **Rigorous charting and coding** - Rigorous charting and coding of data by more than one coder to see whether the same kinds of themes result from their analyses. Furthermore, regular discussions of coding results with colleagues can also be a means for improving dependability. This helps to assess whether researchers' interpretations are in line with what others are thinking (Given, 2008).
3. **Deviant case analysis** - A systematic coding scheme also makes it possible to conduct deviant case analysis. On the one hand, this can reassure about anecdotalism, but it can also help in developing more inclusive theories to account for the data (Seale et al., 1997). Citing negative or deviant cases also helps with credibility and validity: it illustrates that researchers are not just looking for cases that support their theory.
4. **Comparative analysis** - Researchers should be comparing the various cases with one another so that they can build a theory that represents all of the voices present in their findings, and uses understanding of local political economy, social relations and social difference to account for these findings. Furthermore, it is also of value to compare findings with the findings of other research scientists so as to relate what has been found back to the broader research context (Given, 2008).

[26] For example PEER, and SenseMaker, discussed in Chapter 8.

5. **Supporting generalizations by counts of events (quasi-statistics)**. This can address a common concern about the reporting of qualitative data - that anecdotes supporting the writer's hypothesis or argument have been selected, or that undue attention has been paid to rare events, at the expense of more common ones (Ibid).
6. **Ensuring representativeness of cases**. This can be achieved through use of combined qualitative and quantitative methods to support generalizations, and recognition of the merits of representative (often random) sampling as well as theoretical sampling (the selection of cases according to theoretical criteria).

Transparency

This refers to clarity in describing the research process, where researchers are providing their audience with a thorough description of the steps taken in conducting their research. Documentation of research data and the subsequent steps of synthesis, analysis and interpretation (including making primary research material available), should be made freely available for a number of reasons: first, if others want to replicate the research to see whether they achieve similar results, they can; second, it enables readers to assess whether the method chosen was the most appropriate for answering the chosen research question (Given, 2008); third, it enables commissioners and consumers to trace findings, insights, implications and recommendations back to source data; and fourth, it enables reflection on how the research process itself was limited, and what the implications of this are.

5.2 Research ethics

In Chapter 6, we assess RCA against ethical standards for qualitative research. This section introduces those standards.

Ethical challenges in research

Research conducted in developing countries and particularly in relation to development practice raises distinct ethical, moral and political issues and dilemmas. These arise due to current and historical disparities in wealth, power, access to information, political interest, and status. The potential for trust and power imbalances between researchers and participants is heightened (particularly when research is linked to aid policy and program decisions) and unintended negative consequences are a potential outcome. For instance, it is possible to reinforce existing unjust social relationships, to generate conflict or to put participants at risk. Beyond this, when researchers originate from countries other than that in which research takes place, complex cross-cultural issues arise. Differences in culture, norms and values create challenges for both researchers and participants that must be carefully negotiated.

- ACFID, 2017

Ethical standards should be considered by those who commission, manage, conduct or review research and evaluation, particularly in relation to poverty reduction, development and social justice. Ethical practice in research and evaluation relies on active self-reflection, discretion, judgement and appreciation of context (ACFID, 2017). It is critical that this is guided by a set of ethical principles.

While RCA's own ethical guidance is taken from the American Anthropological Association Code of Ethics (RCA, 2017a), we do not consider these to be appropriate or sufficient ethical guidelines for conducting short-term research for development with large teams of researchers, many of whom have not conducted field research previously, or who's only exposure to qualitative research is through the RCA. Nor do we accept that the RCA is an inherently ethical approach, in comparison to other qualitative poverty and social analysis methods used in development. Arvidson writes,

> *The future success of RCA and similar approaches cannot rely on rhetorical presentations that hold them as ethically right, framed as responses to what is seen as previously flawed evaluation techniques, policy makers with a preference for quantifiable data or research that favours objective reporting as opposed to analysis based on empathy and immersion* (2013: 291–293).

We therefore use more widely accepted ethical research guidance for this review.

The Australian Research for Development Impact (RDI) Network has an excellent set of principles and guidelines for ethnical research and evaluation that have been endorsed and adopted by the Australian Council for International Development (ACFID).[27] We have employed the ACFID (2017) Guidelines as a framework for the assessment of the RCA and, unless otherwise cited, the following is taken from these.

Privacy and confidentiality

Research participants should have the right to remain anonymous and to have their rights to privacy and confidentiality respected. ACFID suggests that "*privacy and confidentiality refers to how much information a participant may wish to share and entrust with the researcher, as well as how the information they share is obtained, protected and stored*" (2017: 12).

Privacy and confidentiality concerns also relate to how research participants are represented in the research. It is important to respectfully represent participants, and in some contexts ask participants how they would like to be represented in research products should be considered. For example, representing people who live in informal settlements as "slum dwellers" might not be preferred by them.

Researchers should be aware that preserving anonymity requires that there be no link between the data (responses) and the source (the participant). In some cases, this may require aggregation of responses or changed details to prevent identification of communities, households or individuals. There may be limits to confidentiality (for instance, mandatory disclosure of cases of abuse); these limits should be made clear to participants during the consent process. Privacy also requires that researchers take responsibility for data to be stored securely with access limited to designated, authorised people.

Informed consent

A necessary (though insufficient) prerequisite of any qualitative social science research is that research participants give informed consent.[28] Specific procedures and considerations should be observed in the case of particular groups such as children and young people, and people with disabilities. Informed consent is an ongoing process and must be renegotiable so that participant understanding and comfort is assured. ACFID Guidelines suggest that the information provided to participants (either verbally or written) should include:

- Research aims and objectives
- Details of information that is being sought
- How responses will be recorded and used
- The degree to which participants will be consulted prior to publication
- How findings will be communicated to participants - Communicating the findings of the research to participants needs to be ensured, and the ways in which this will happen also needs to be communicated to participants during the consent process
- Potential benefits and consequences of participation, including potential risks

[27] https://rdinetwork.org.au/effective-ethical-research-evaluation/principles-guidelines-ethical-research-evaluation/
[28] See also Shah, 2018.

- Reimbursements or incentives (if any) that will be provided for participating in the study - careful, culturally appropriate decisions need to be made on the nature of any recompense to participants or a community for participation, ensuring that any recompense is not perceived as an undue inducement
- The name of the organization that is funding the research
- Contact details for someone independent of the research process for inquiries and complaints
- An explanation of the voluntary nature of the participant's involvement
- The name of the investigator

In addition, consent for the use of images (photographs) should be sought. It is advisable that even where consent to participate is verbal, an informed consent form be completed by the researcher that confirms that all elements of the informed consent process have been adhered to.

In addition, much ethnographic-style research, including RCA, research, takes place amongst poorly educated rural populations, where participants are unlikely to have a framework within which to understand the research (Shah, 2018). Clear, simple and repeated communication about the purpose of the research is essential: "*Ongoing communication and confirmation of consent may be required during the research to ensure that the concept of consent is understood, particularly towards the end of the research process*" (ACFID, 2017: 14). This can be done in a manner that does not intrude on rapport-building and establishing a trusting relationship between researchers and participants, but is particularly important when researchers are not writing notes and are working hard to "fit in". Participants in research can easily forget that what they say and do may be recorded and used.

Working with vulnerable populations

Compounding the challenges of working with rural populations , who may have no or low literacy, are the challenges of working with vulnerable participants – such as adolescent girls - and on sensitive issues – such as violence against women and girls, or security and justice. Work with these populations and on these issues requires researchers to follow proper procedures for obtaining truly informed consent.

Informed consent raises particular challenges when research involves children and young people, arising from four main concerns:

1. Children's or young people's capacity to understand what the research involves, and therefore whether their consent to participate is sufficient
2. Possible coercion of children or young people by parents, peers, researchers or others to participate
3. Potential for conflicting values and interests of parents, guardians or primary caregivers and children
4. During a research process, possible disclosure by a child of information that raises child protection concerns (e.g. information indicating that they are currently at risk of or are experiencing violence, exploitation or abuse), obliging researchers to report such circumstances.

- ACFID, 2017

Working with vulnerable participants also requires researchers to have a sound understanding of the "risk context" for those participants. By this we are referring to the risks that participants face in their everyday social worlds, but also the risks that the research can exacerbate or even introduce. There may be additional risks, for example, where research is being conducted with ethnic minorities within a broader community or with groups who are otherwise stigmatized, or with participants who's mobility and contact with outsiders is highly mediated.

All research should have robust policies and procedures related to working with "common" vulnerable populations such as children. These protect not only research participants, but also researchers. ACFID (2017) suggests that a child's or young person's consent is sufficient only if he or she demonstrates sufficient maturity to understand the relevant information and to give consent. Commonly, additional consent is required from parents (or guardian or primary caregiver) when a

child under the age of 18 has agreed to participate, as an extra safeguard. At all times, children and young people have the right to cease participating in research activities if they choose and opportunity for this should always be provided (and any limits to this explained). This relates to the principle of respect in consideration of the dignity of the individual participant, their capability and right to make decisions about matters that affect them and the researchers' responsibility to uphold 'children's right to dissent, that is to refuse participation and to withdraw at any time and to prioritise this over their parents' or others' wish for them to participate.[29]

Researchers should also be well equipped to handle a disclosure of abuse (in terms of training and skills) and should have a reporting or referral plan in place to be able to respond. Researchers should be familiar with in-country child protection referral mechanisms and child protection focal points. Good practice would be for researchers to have made contact with local organisations – governmental or non-governmental – who's role it is to provide support to children at risk of abuse or exploitation.

Researchers should be cognisant to ensure that research processes are inclusive of people with disability and that their specific participation limitations are addressed. The impact of a disability is often dependent on environmental barriers; hence many people with disability will have full capacity to participate in research and should not be deemed to be of high risk solely due to their disability. Guidelines for ensuring fully informed consent can be found in ACFID (2017).

Culturally sensitive research design and consideration of context

Research design should reflect the context in which the research will take place; this is an issue of ethics, as well as of rigor. Research cannot be assumed to have beneficial outcomes for host communities or relevant research participants. What is considered appropriate in one context might not be in another; research design requires a firm grounding in the relevant local cultural values, norms and the local historical and political context.

> *For any given context, a first step is to identify key cultural values and customs and analyse how these impact on meaningful adherence and interpretation to ethical research principles. This process requires critical reflection on researchers' own cultural values, how these influence proposed questions and design, and challenging them with alternative perspectives at all stages of the research. Models to support ethical research should be explored, many of which increase reliance on participatory, collaborative processes and concepts of partnership and reciprocity* (ACFID, 2017: 17).

Examples of models and approaches to support ethical research

- Involvement of participants in framing research and research questions and/or other steps such as deciding data collection methods, analysis methods or validation, seeking alignment with cultural norms.
- Well-designed processes for how research findings are communicated to participants including methods to actively engage participants with the findings and their implications.
- Community/participant control over the research process itself, with the local people leading and implementing the research.
- Involvement of local co-researchers (not from the relevant community/locality) who carry relevant cultural values and insight. This approach may also provide other benefits through supporting local research capacity building. In such cases, researchers have the responsibility to ensure that local co-researchers undertake the research in an ethically appropriate manner.

\- ACFID (2017: 17)

[29] For specific guidelines regarding informed consent involving children and youth, see Ethical Research Involving Children (ERIC) (2013). Ethical Guidance: Informed Consent. Available at: http://childethics.com/wp-content/uploads/2013/10/ERIC-compendium-Ethical-Guidance-Informed-consent-section-only.pdf, accessed 23 April 2018.

Gender should be a key concern since gender norms and laws in many countries require specific research design considerations. This involves reflecting on the implications of the research for male and female participants in design, and being respectful of gender and sexual identities in research implementation. Consideration should also be given to the intersection of gender with other factors that shape a person's circumstances and interests, including age, ethnicity, (dis)ability or religion. Attention to such intersections is important because these can magnify disadvantage, risks and barriers to participation, as well as change the potential benefits of the research.

Researcher safety

Ethical research also includes ensuring that the research is designed and implemented in a way that does not unduly compromise researcher safety. At a minimum, a risk assessment should be done that takes into account not only risks to researchers in general, but also looks at possible risks to researchers from different ethnic, linguistic and religious backgrounds, as well as specific risks to researchers because of their sex or sexual identity.

A proper risk assessment can only be done when there is sufficient knowledge concerning the research site (as outlined above), including any local tensions that the researcher might be exposed to. Risk planning should include both common risk mitigation measures (such as ensuring that researchers have the means to communicate with research managers and have relevant numbers and locations for emergency services), as well as highlighting potential local specific risks, such as risks from malaria, risks due to religious tensions, or risks posed by research populations negative relationship with government or law enforcement. Risk planning needs to be carried out based on risk assessment, and regularly updated.

Review and approval processes

Increasingly, the quality of outputs are subject to review by groups of peers, ideally independent (and sometimes anonymous or at least anonymized) review processes. Good practice suggests that review processes be set up at research inception, to enable expert inputs into research design and preparation, as well as to comment on research products.

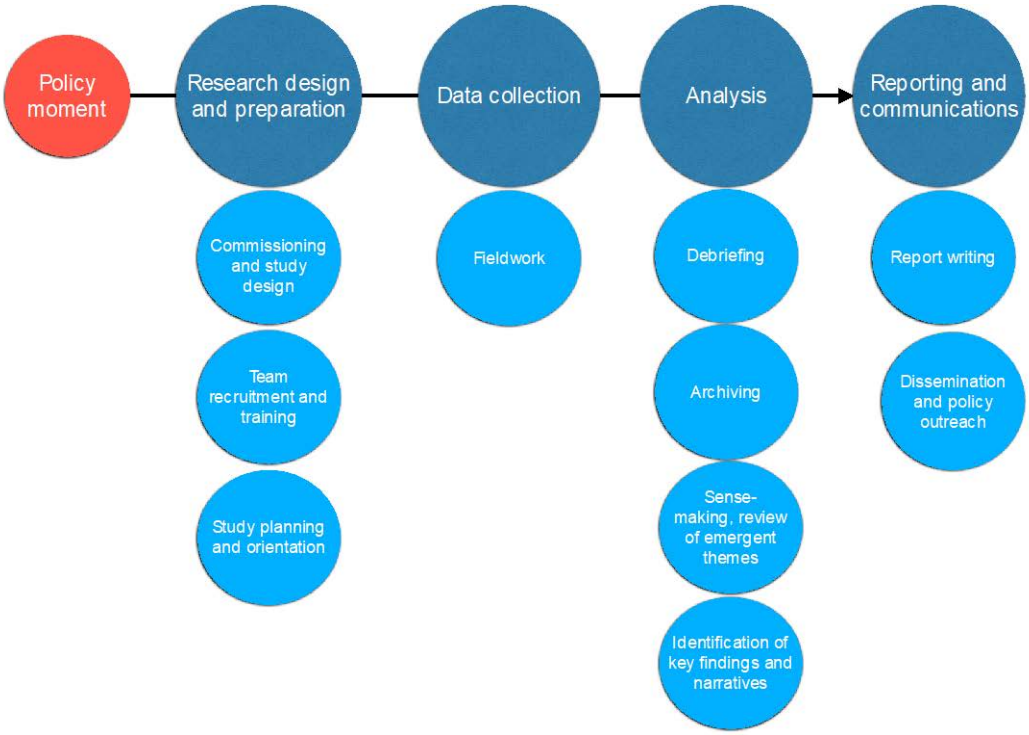
ACFID (2017) suggests a "rule of thumb" for determining whether an external ethical review process is necessary for approval. If the research is considered of 'negligible risk' (where any foreseeable risk is no more than inconvenience) or 'low risk' (where the only foreseeable risk is discomfort), then a reduced or internal assessment of ethical issues may be considered rather than a formal ethical review and approval. Such cases do not apply, however, where risks to participants are not considered "low" or "negligible", and where evaluation or research processes address sensitive issues or topics, involve vulnerable groups, or use significant participant time. Further, a more formal ethical review process should be considered for less well-established evaluation methods, where the aim or purpose of the research goes beyond improving the implementation of an established intervention or program (quality assurance), or where research is largely exploratory in nature.

In closing this section, two things are worth noting. The first is that ethics are not an area where it is advisable to pick and choose some considerations over others; all are equally important. The second is that certain approaches are not more inherently ethical than others, as we noted in the introduction to this section. Ethnographic approaches as used within development are just as duty bound to adhere to the highest ethical standards of qualitative research as are other approaches, and perhaps even more so due to the intimacy required by these approaches.

6 RCA praxis: methodological rigor and ethics

The RCA research process, as conceptualised by the RCA core team, is presented below in Figure 5. This Chapter looks at each main stage of the research process – i. research preparation and design, ii. data collection, iii. analysis, and iv. reporting and communications – and assesses the quality of these against the rigor and ethics framework described in the previous chapter.

Figure 5 – Typical RCA Study Process[30]



6.1 Research preparation and design

6.1.1 Experienced, reflexive and well-trained researchers

Key Findings

1. There is significant variability in experience of RCA researchers – some have no qualitative research experience prior to joining RCA studies - and training is insufficient for those with little to no background conducting qualitative research.
2. While training materials for Level 1 are of reasonable quality (and would serve as a good refresher and RCA orientation for experienced qualitative researchers), training materials for other levels are insufficient to support progression from novice to more experienced researcher levels.

[30] Adapted from: RCA, 2017b

RCA's approach to recruitment

Various RCA study reports and documents, as well as RCA practitioners, suggest that the skills and experience required to conduct RCA research are somewhat different from those required for other qualitative research. RCA researchers should be "*people persons, who are friendly, warm, open and able to put others at ease*" (Masset et al., 2016: 174). They also need to be good listeners, and open minded enough to be able to put aside their own assumptions and suspend judgement (Ibid). Those familiar with RCA claim that "*Research skills are not necessarily the skills you need to do a good reality check. You just need good conversational and listening skills*" (KII, Qualitative research expert).[31]

A number of RCA studies state that a good understanding of community dynamics gained through working experience at local community level is very valuable in a researcher, more valuable than extensive experience or formal qualifications. Indeed, several internal RCA key informants and other proponents of RCA suggested that researchers' lack of experience as well as their lack of knowledge of the issue being investigated are key strengths within the RCA. Younger researchers, it was explained by several RCA core team members, have fewer biases, are less stuck within an "*emic expert paradigm*", and are more willing to spend several nights and days staying in what are often quite challenging circumstances. According to an RCA core team member, with inexperienced researchers it is also "*easier to spot weaknesses and mould them the right way*" (KII, RCA core team member). Some proponents of the RCA suggested that researchers coming in fresh, without previous research "baggage", and learning on the job, is a significant strength of RCA.

RCA's approach to training

In any qualitative research exercise, it is often extremely challenging to find sufficient numbers of good qualitative researchers who are available for short term research assignments. To compensate for any lack in skills and experience, good quality training of researchers is therefore almost always needed. RCA promotional materials claim that researchers are both experienced and "well-trained". The RCA+ project has enabled a huge investment in researcher training that has helped RCA researcher training to evolved based on learning and the need to be more rigorous. There are currently 4 Levels of researcher training.

Level 1 - RCA literature and discussions with RCA core team members suggest that all new researchers go through a 5-day Level 1 training, including 2 days spent in the field with a more experienced RCA researcher. Level 1 training concentrates on what the Reality Approach is and how it differs from other approaches to research; practical sessions on, for example, how to enter the community and find a host household; and researcher mind-sets and behavior. Training materials include a number of interactive exercises that enable trainees to deepen their understanding of qualitative research and practice key elements of the RCA, such as probing, asking open ended questions, conversation skills, reflexivity, and offsetting bias. There are also a number of training resources on power differences between researchers and research "subjects" and the potential for bias as power dynamics play out in the field.

There are a number of other Levels of training that RCA researchers can progress through:

[31] Claims about the uniqueness of these stated attributes required for RCA are debatable, as these are also important characteristics of any good qualitative work, and certainly do not preclude specific research experience.

Level 2[32] **-** Once researchers: i.) have participated as a team member in at least two different studies and participated as a translator (for non-local RCA research team members) in at least 1 study; ii.) are comfortable with uncertainty, able to provide sufficient reassurance to sub team members without interference/over protection, serve as a' critical friend', level headed; and iii.) are "*fully aware of the principles and ideology of RCA*"(RCA, 2016a: 3) and are able to communicate this to others, they are eligible for Level 2 training. This training enables them to take on the role of sub-team leader. It is conducted over 2 days and focuses during the first day on "advanced field work skills", going deeper into issues such as triangulation, power, bias, and different questioning approaches, and on the second day on how to support sub-teams in the field, ensure that all materials are archived when the team returns from the field, and how to contribute to certain aspects of final reports, such as boxed case studies and study limitations.

Level 3[33] **-** Training consists primarily of instructions and exercises related to report writing, explaining style and content requirements, with a significant emphasis on what language should be used, for example using "my mother" rather than "the female head of household" in order to personalise findings. Once completed, this short training allows the researcher to be a study Team Leader, with responsibilities for negotiating with clients and commissioners of research, undertaking secondary data review, conceptualising and designing the RCA study, undertaking risk assessment, developing areas for conversation, training researchers, leading team debriefings, undertaking analysis and report writing, and engaging in policy dialogue.

Level 4 - Level 4 does not require additional training, but being recognised at this level empowers a Team Leader to contribute to proposal writing and bid preparation.

What we found: recruitment and training in practice[34]

In terms of recruitment, interviews with senior RCA researchers (Level 2 and above) suggested that many of them do have sound knowledge and experience of how to skilfully guide conversations and probe into areas of interest. Unfortunately, they are a generally a small minority on research teams. We also spoke with a number of relatively new RCA researchers who lacked relevant experience. These researchers had backgrounds in business, law, science, media and advertising, English teaching, and human resources, and many were young urbanites who had not spent time in rural settings.[35]

Even where claims were explicitly made about the high level of experience, such as the Adolescent Nutrition and Physical Activity Report, as far as this review can ascertain, these researchers were not substantively better trained nor more experienced than researchers on any of the other RCA case studies, and none had specific expertise in adolescents or nutrition. Indeed the report later corroborates this by explaining that they have backgrounds in, for example, political science, sociology and law, and that research teams were "*purposely convened to include younger*

[32] At the time of writing, Level 2 and 3 training materials were unfinished, with modules and exercises missing, so it is not clear whether they have just recently been developed or are under revision.
[33] RCA (n.d.b) Level 3 Training Materials.
[34] As we were unable to attend a training course, we are basing our description of levels of training on documentation and key informant interviews, rather than observation, and cannot draw any firm conclusions concerning how consistently training materials and protocols are applied.
[35] According to RCA records, all researchers have Bachelor's degrees and above. While some had previous qualitative field research training and experience prior to joining RCA studies, a significant number did not. This is true even of the Senior RCA team, which include a number of members who have had no prior experience of qualitative research before joining the RCA+ Project.

researchers who were closer in age to the study cohort and would find it easier to hang out and be accepted by them" (RCA, 2016b: 5).

In the words of one key informant, who has both participated in and consumed RCA studies, "*I think it is safe to say that only about 25% of the team are really good quality. The rest are very variable, with some being extremely poor quality*." There are likely to be better quality teams in countries such as Indonesia and Nepal, where a number of RCA studies have been conducted and where a local cadre of RCA researchers has been built up. In other countries, where only one or two RCA studies have been conducted, the team is likely to be far less experienced, with the more experienced sub-team leaders having to be brought in from other countries (primarily Indonesia, Nepal and Bangladesh), so with little contextual understanding. For example, an RCA study in Ethiopia utilised a number of quantitative "enumerators" who had very little research experience (according to informants, they were not even professional enumerators, but rather young professionals taking leave from their regular jobs), and no qualitative research experience.

A number of key informants from within the RCA core team agreed that researcher quality is a challenge, with one core team member explaining that this had been addressed through adding more non-RCA trained researchers to teams, to ensure that there are enough good quality researchers on each study, presumably to balance out researchers who are of poorer quality. As noted by one experienced researcher, "*RCA has many young researchers with a variety of experience, and takes a lot of experience to have a really good conversation with local people. It takes a lot of skill, especially talking to older people, rural people, when these researchers are so young and urban mostly. I'm not sure how researcher quality is really addressed*" (KII, RCA researcher).

The length and depth of training is particularly critical considering the lack of previous qualitative research experience amongst many recruits. We found that the Level 1 training times can be significantly shorter than the norm outlined above. For example, it appears that an RCA study conducted in Ethiopia in 2016 included 1.5 days of training, a short immersion (2 nights), and another ½ day debrief before the researchers (a majority of whom had never conducted any qualitative research previously) returned to the field to conduct the RCA study. Prior to the IP-SSJ study in Nepal, researchers: "*…participated in a 2-3 day briefing led by the Principal Investigator before going into the study locations. This included training on RCA methodology as well as a specific discussion to develop areas of inquiry*" (RCA, 2015a: 16).

In theory, the lack of experience should be addressed by the training, but the training is too short to allow this. One key informant noted that "*RCA training seems too brief to get recent graduates to suspend judgment and really listen. In my view this would that would take more than 2 - 3 weeks training" (KII, Research expert).* It is not entirely clear without having observed a training program how deep the investigations of personal bias go, though the training materials themselves appear to provide ample opportunities to explore potential bias and to develop non-judgemental research strategies.

Whether training can even be realistically expected compensate for a limited experience amongst researchers is debatable. Morse writes, *"From my own experience, I know that short "training" courses will not enable lay interviewers to conduct high quality unstructured interviews. These team members will not always know why, what or when to probe, and gems will pass them by in every interview"* (1994: 4). The use of researchers with little previous research or community level experience introduces the strong potential that a number of RCA researchers are simply "skimming the surface"; and there is significant potential for assumptions being made because investigative depth is not being achieved.

This could be exacerbated by the generic content of the training, which does not sufficiently explore a range of contextual issues that are relevant for research. Although researchers are provided with a short one day orientation prior to each RCA study, this does not reach the depth that is needed in order to achieve a good understanding of contextual issues (we discuss this further below in 6.1.2).

The above issues ultimately combine to raise questions about the validity of findings, as we found in our review of RCA reports. While we discuss this in some depth in Chapter 7, one salient example from the IP-SSJ final report illustrates this clearly. The report states, "*Though this topic rarely came up in conversation, people in Kapilvastu (2) and Dhanusha (1) mentioned that previously women would commonly be accused of witchcraft, though no incidences had occurred for over a decade. One person explained this change by saying, 'people are now more aware so they now know.'*" (RCA, 2015a: 29). One is left wondering not only 'what do they know now? and who are 'they'? and how did they come to know this?' but more generally whether the researchers had the skills to able to effectively explore social and gender issues in relation to security and justice.

6.1.2 Contextual Understanding

Key Findings

1. RCA actively excludes contextual understanding, claiming 'virtue in ignorance' as a way to eliminate researcher bias.
2. RCA researchers therefore, unsurprisingly, neither seek nor achieve good contextual understanding prior to conducting research.

RCA's "virtue in ignorance" approach

As discussed in 6.1.1 above, not only do many researchers potentially lack a good understanding of community dynamics based on having some working experience at local community level, and not only does training fail to explore these community dynamics in any detail, but there is inadequate research and orientation on key sectoral and other relevant issues prior to studies being undertaken.

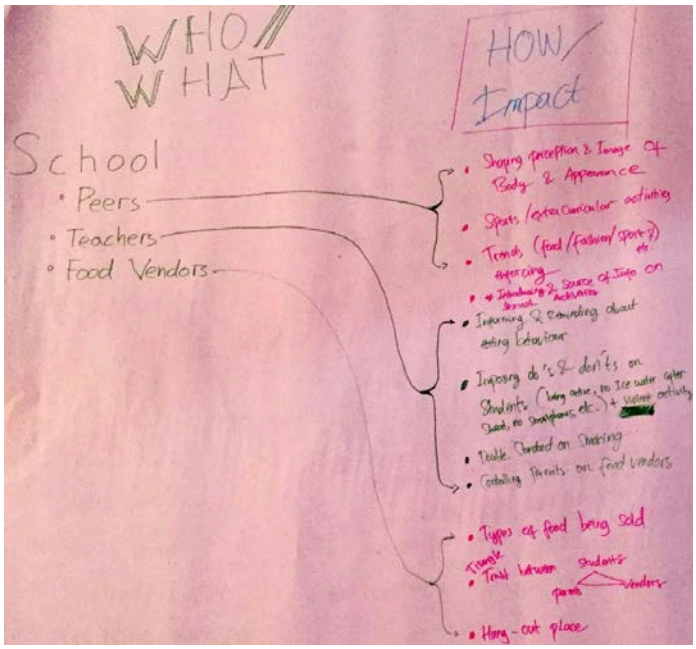
An RCA core team produced video entitled "Rigour in RCA"[36] attests to the fact that RCA underemphasises background research. The video claims that "*Others have project or agenda bias. We don't.*" This is accompanied by a cartoon person wearing a t-shirt that says "just curious" with a thought bubble above their head saying, "blank." Interviews with RCA team members confirm this "virtue in ignorance" approach, as do RCA reports:

> *We don't give out project information to researchers before, so that that if something does or does not come up about a project intervention, they won't either make assumptions of impact where there is none or make more of it than should be made of it, or try to dig to surface something if*

[36] http://www.reality-check-approach.com/related-resources.html

> *it does not come up. This reduces bias and* conjecture (KII, RCA core team member).
>
> *Another potential limitation to this study is researcher bias in recollection during the debriefing process…Though this limitation applies to all qualitative research, the research team mitigated this bias in a number of ways. First, by providing the study team with minimal information on IP-SSJ goals and activities which allowed them to explore the areas for conversation openly and without programme bias* (RCA, 2015a: 19).

As we can see from the above, the main reasons that contextual understanding is de-emphasized in RCA is to avoid exposure to previous research and "received wisdom" in an attempt to reduce bias and generate insights that are experiential rather than academic.



For most RCA researchers, the only exposure that they have to the issues under research is in a short one day briefing prior to fieldwork. Pre-field briefings include a power mapping exercise to help researchers to recognize the potential for bias related to dominant personalities, positions and views (RCA, 2017a). While this has the potential to fill a gap in previous training on this issue, the exercises themselves appear to be relatively "light" and not particularly analytical.[37] There is a distinct lack of gender analysis, and – we think somewhat strangely considering that researchers are not experts in the issues they are researching – exercises the rely on researchers developing what can only be called "assumptions" about what actors and factors influence the issues being researcher. An example is provided in the box above where we can see this lack of gender analysis, as well as the absence of potentially key stakeholder groups such, parents, siblings, and media personalities.

Pre-field briefings also focus on potential personal bias and engage research team members in reflecting on what their biases might be concerning the issue under investigation. For example, the urban migration study RCA team highlighted the possibility that women are the ones staying at home taking care of children and that young girls are disadvantaged by migration as a possible bias to be aware of during research. Also recognised was a personal bias that needed to be guarded against was slums being "horrible places" to live.

Finally, there have been some RCA studies where pre-field briefing includes a brief presentation and discussion session on sector-specific information, though not of the specific group to be

[37] We examined these in the Adolescent Nutrition and Physical Activity and Village Law case studies.

researched. For example, in the markets for the poor research in Uganda, researchers spent a morning learning about "agricultural basics". RCA team leaders in Indonesia, though not nutrition specialists themselves, gave researchers an orientation on nutrition (based on some light touch secondary research) for the nutrition in UNICEF adolescent nutrition and physical activity study. According to an RCA team member, this briefing provided some information "*but not so much that we developed biases that we were looking to confirm in the research*" (KII, RCA core team member). In addition, consultations with a reference group, which can contain sector specialists, to develop "areas of conversation" that guide research (discussed below in 6.1.3) is seen as adding some rigor to the process, though our findings suggest that these some of these consultations appear to be rather cursory.

What we found: "ignorance does not ensure insight"

There is a real and legitimate concern that researchers are not attuned to issues of gender, power (beyond researcher/researched power dynamics), and exclusion, and therefore are simply unable to pick up nuances or understand data within context, or even appropriately identify certain issues as relevant for consideration in the first place. This is particularly a problem for studies that demand just this, for example the IP-SSJ study in Nepal. We ended section 6.1.1 with a brief discussion of researchers' ability to effectively explore gender issues in this particular RCA, which had as an objective, "*to improve the evidence base on how people experience and understand security and justice in Nepal, with particular reference to poor and marginalized groups (including women)*" (RCA, 2015a: 12). There is clearly an expectation this the study will explore issues of gender, religion, caste and ethnicity in relation to security and justice. Yet the team had little background on this and clearly had not filled this gap prior to research.

This highlights the twin problems of knowing what to ask about, or ask more about, and understanding what, if any, insights might be gleaned from the response.

> *It strikes me that there are two logistical questions when conducting research from the no-going-to-the-library mode. First, how do you know if a problem – a research question – is really a good question, if you do not search to know if anyone already knows the answer, or perhaps has worked on some portion of the problem, and may be able to shed some light on the matter…Ignorance does not ensure insight. Nor does ignorance ensure theoretical sensitivity. Rather, ignorance is more likely to limit one's theoretical vision and narrow one's mind by restricting conceivable theoretical options* (Morse 1994: 3).

Using knowledge to make sense

A range of academic sources suggest that researchers need to have an intimate knowledge of the issues and research context in order both to collect relevant information, and to make meaning of this. Shah writes that there is "*near unanimous insistence across disciplines and approaches that background research is essential to contextualise short-term ethnographic studies and inform study design*" (2018: 13).

Morse and others (cf. Bell and Aggleton 2016) urge for a middle ground, where the researcher neither goes with an agenda, using some theoretical model and collecting and sorting data accordingly, nor goes in blind, without a vast compendium of knowledge. "*Ideally the researcher should be an active astute observer and a wise listener, so that when indicators appear, they may be appropriately recognized as having some significance or meaning that links the data with what is known. Rather the researcher must be a knowledgeable instrument: listening, observing,*

thinking and making sense" (Morse 1994: 4). In order to do this, it is critical that the researcher constantly asks informed questions and checks the fit of established theory to see if it holds in the new situation, new context, and with the new question. RCA researchers, who know little of the context, issue, or theory of that which they are investigating, are ill-equipped to do this as they purposely avoid *"stand[ing] on the shoulders of giants*" (Morse, 1994: 3).

The cursory way in which RCA collects and analyses secondary data is also of concern to some RCA commissioners and consumers:

> *In my view the RCA team didn't demand enough information. It felt like the team did not understand what they were after. I had to request meetings and suggest they read certain things* (KII, Commissioner).

It is of concern to several more experienced RCA researchers with whom we spoke. One researcher reported, "*RCA also did not accept if we found secondary data which can support our analysis and avoid bias. There is one of our team members who secretly copied data from the head of village's laptop*" (KII, RCA researcher). The copying of data reportedly occurred so that s/he could present these figures as things s/he had remembered or jotted down, as secondary data would not be accepted by the Team Leader.

And finally, largely dismissing secondary data it is also of concern to all of the other qualitative researchers with whom we spoke:

> *Displacing accumulated knowledge and expertise that anthropologists have brought to development studies for decades seems like folly to me. It's like the RCA top brass think "We're not going to burden them with knowledge or information, or frameworks" no matter that many of these have been developed by intelligent and reflexive practitioners"* (KII, Qualitative research expert).

Using knowledge to counter bias: an accepted practice in qualitative research traditions

Recent academic sources challenge the notion that if researchers read extensively about a topic or participant group before commencing research, this newly acquired knowledge would introduce bias into one's data collection or analysis. Morse (1994) argues that researchers should not ignore accumulated wisdom, but should use it, question it, challenge it, and add to it. She is not alone. Shah (2018) argues that without access to both high quality background information and good methodological training, researchers tend to make ethnocentric assumptions leading to poor study design, unethical behaviour in the field, and inaccurate interpretations.

> *With more experienced researchers, and in long-term ethnography, such assumptions are both less likely to be held, and more likely to be challenged and corrected over the course of the research, but in short-term immersions there is insufficient time for researchers to be exposed to sufficient data, and to build sufficiently close relationships, for this to occur* (Shah 2018: 14).

Similarly, Bell and Aggleton write that, "*Rapid approaches are best deployed by researchers who have long-standing experience in a particular field site, within a specific culture, or with the population groups under study*" (2016: 7). Morse too suggests that insights that do not fall on "fertile ground" may be missed:

> *…for an insight to occur, the researchers' minds must be ready. They must have much knowledge about what they are observing or hearing, know and be able to link it to relevant literature, and be able to think conceptually and to link seemingly unconnected events, representations, and ideas...we must use the library, read everything possible before going into the field, so that we will be able to recognize a "wheel," for instance, if we see one. We must become immersed in our data, have good-quality data, well-scoped data, and seek saturation (2005: 4).*

6.1.3 Framework to guide inquiry

Key Findings

1. RCA uses a series of questions, captured in "areas for conversation" that enable an open, exploratory approach to research; however these are often very vague, leading to low-quality data in terms of relevance and depth.
2. At the same time, there is little evidence that this approach either reduces potential bias or leads to more meaningful conversations.

RCA's approach to inquiry

RCA studies do not use a preconceived research framework or research questions, which they argue can suffer from normative bias. Instead, they work with commissioners to identify a number of "areas of conversation" that researchers familiar themselves with prior to research, and can return to as an aide memoire during field research. While the areas of conversation for many studies appear to be vague and ill-defined, others (for example for the IP-SSJ RCA Study in Nepal) contain a vast number of highly detailed and complex questions, making them difficult to explore with an approach that relies largely on memory.

RCA practitioners claim that "*These topic areas enable open-ended, assumption and value judgement-free conversations*" (RCA, 2017a: 50). The Haze RCA study report explains:

> *…asking whether their children wear a mask during haze period might be interpreted as an outsider passing judgement on the parents, who might tend to simply answer 'yes' as they do not want to be seen as 'bad parents'. Additionally, the family might have a different opinion regarding the use of masks or other coping mechanisms, which might not emerge following the prescribed questions. Instead the RCA researchers will have detailed conversations and chat about people's reaction to worsening haze* (RCA, 2016c: 5).

We suggest, however, that asking about "worsening haze" makes an assumption that people perceive the haze to be worsening. Clearly, having "Areas of Conversation" rather than research questions does not entirely avoid bias or assumptions.

RCA practitioners defend their lack of a theoretical base as a strength, claiming that theory-driven research is arguably inherently biased in that it has an onus to prove or disprove a logical connection. Further, they suggest that research questions derived from a theoretical frame also

contain normative bias and may contain value judgement and can limit exploration of a topic. By contrast, they suggest, RCA researchers are free to explore, gather new insights and discover unexpected connections, which provides important insights into negative as well as positive, indirect as well as direct, and unintended as well as intended effects of interventions. A lack of a guiding frame, RCA practitioners argue, contributes to its ability to extend inquiry beyond sectoral lenses and project/logic/theory of change boundaries, to accommodate multiple realities and embrace unexpected, non-normative experience.

The practical benefits of having a framework for inquiry, and the pitfalls of not having one

One of the hallmarks of qualitative research is that it is inductive, rather than deductive - enabling theory to emerge from evidence rather than evidence being collected to prove or disprove a theory. However, most qualitative research approaches do use contextual understanding and issue-based knowledge to develop a conceptual research framework that helps to guide investigation into different areas and issues, enabling researchers to take account of accumulated knowledge around what issues are well-evidenced, and where there are evidence gaps, as well as what is well-known about certain research segments, such as adolescent girls, or sex workers, for example, then to identify research questions to explore through the research and analysis process.

Indeed, not having this guide can make bias less explicit (and therefore harder to address) before and during research:

> *Everybody comes in with biases, and you can't report from the perspective of local people without it being filtered through your biases. Biases are checked if your frame of reference includes wide body of knowledge and debate on an issue and if it includes local people's own analyses and challenges of your preliminary findings. You need to actually ask the right questions. If you don't know the issues, the area, and have questions to probe on this, then you can totally miss important issues* (KII, Qualitative research expert).

The ability to extend inquiry beyond sectoral lenses and project/logic/theory of change boundaries, to accommodate multiple realities and embrace unexpected, non-normative experience is clearly a potential strength of ethnographic research in general. However, not having a framework for inquiry potentially lead to more superficial conversations that fail to capture this richness. Shah writes of RCA that if a researcher, "*has numerous, ill-defined questions or themes, she is more likely to have superficial conversations about a wide range of related topics without ever understanding the connections between them*" (2018: 15).

Another disadvantage of not having an agreed research framework is that data collected by one researcher might not be comparable to data collected by other researchers because modes and areas of inquiry are so divergent, and also because researchers only infrequently, if ever, return to their "areas of conversation". This makes both aggregation and comparison more difficult and less rigorous, an issue we discuss in section 6.3.3 below.

6.1.4 Multi-disciplinarity

Key Finding

1. RCA seeks to achieve multi-disciplinarity through putting together large diverse teams. However the RCA fails to achieve multi-disciplinarity, as researchers' disciplinary backgrounds are largely irrelevant to specific RCA studies.

In this short section, we comment on multi-disciplinarity in both research design and analysis in order to avoid unnecessary repetition.

RCA practice

Wherever possible, RCA sub-teams (the 3 – 4 researchers who conduct research in one location) have a diverse composition. There is an explicit attempt to include men and women, older and younger researchers and, where important, researchers with different religious/ethnic backgrounds in order to off-set potential biases. However researchers are not recruited for their disciplinary background or experience in the expectation that they can bring this to bear on the research in any way.

What we found

Despite the diverse backgrounds of these teams, RCA does not seek a multi-disciplinary lens on research design or analysis. Not understanding the Reality Check Approach to research, some commissioners and consumers assume that multi-disciplinarity is a potential strength of RCA. As one commissioner commented, "*The team had a varied background, which I thought would be a strength*" (KII, Commissioner). However, in our view it is not. There can be little point in having an engineer on a team studying a health issue, for example.

Different researcher perspectives are not sought on research design or during the analysis process, and indeed are actively discouraged as potentially introducing "bias". A few RCA researchers shared their views that this was a weakness of RCA, with one experienced qualitative researcher commenting, "*I had a background in this area, so knew that some of what people reported and was being taken as 'the people's views" was actually quite normative, what people think that outsiders want to hear, what they should be saying. I tried to raise this, but I was told that I was being biased, and that the people are the experts…*" (KII, RCA researcher).

6.1.5 Ethical considerations[38]

Key Findings

1. Training and orientation do not inadequately prepare field researchers, sub-team leaders and team leaders to act in the best interests of children and other vulnerable research participants in the field.
2. Ethical review prior to fieldwork is not standard, and it is clear that commissioner approval processes are insufficient in this regard.

[38] This review reports what we found to be the situation up to March 2018. As per 1.4.2, above, there have reportedly been significant changes to training materials, policies and protocols in relation to child protection. Assessing these new policies and practices is beyond the scope of this review. We leave it to commissioners to demand the highest standards of ethics from RCA and other qualitative research approaches.

Child protection

The RCA's policy

In response to previous criticisms regarding ethical standards, the RCA+ project developed a more comprehensive set of guidelines and training materials that aimed to address ethical issues; the covert/overt nature of disclosure to study participants; issues of consent, confidentiality, compensation for host households; and compliance with protection measures (RCA, 2017a). RCA documentation suggests that policy and practice on child protection is now a strength of the approach:

> *RCA+ Project took advice from the Palladium Child Protection Specialist and developed a Child Protection Policy which exceeds the compliance with DFAT regulations and is consistent with Palladium's Global Child Protection Policy. RCA+ Project has a designated Child Protection Officer who attended DFAT Child Protection training…and has developed a training module for Level 1 training. Researchers are required to re-read the Child Protection Policy and Data Protection Policy before every RCA study they participate in and are required to re-sign declarations with each contract. Training is also provided on 'intervention' and what action to take in the event of witnessing child abuse* (RCA, 2017a: 49).

What we found in practice

While increased attention to child protection in training is welcome, our review of training materials and other internal documentation suggests that this is not yet adequate. There is an extremely short and non-comprehensive orientation to child safeguarding during RCA Level 1 training that focuses on delivering key messages around not doing harm to children and ensuring that researchers obtain permission to take photographs. We did not find any consideration of protection for children and other vulnerable populations in other levels of training. A very brief mention of child protection is made in the pre-field briefing.

There is an absence of formal reporting protocols for researchers, and guidance for sub-team leaders and team leaders on how to respond to incidents of abuse or suspected abuse is not evident. There appear to be no attempts to identify local partners near research sites that could provide support to children or vulnerable people who are identified during the course of research, for example NGOs or community groups who work on child protection or provide services for those who have experienced VAWG. There are no special guidelines or modules in training that address ethical issues related to other vulnerable populations such as women experiencing domestic violence.

Site selection and researcher safety protocols

RCA guidelines

Criteria for the selection of study locations are worked through with research commissioners and Reference Group members. There are explicit attempts to minimize the potential for bias and to protect the anonymity of research informants. Only the research team knows the location of their own study sites, and then not until they are actually going to the field.

The RCA core team has also developed a new set of comprehensive guidelines and protocols related to researcher safety that include location-specific assessments of security and current issues; development of a risk management plan and communication of this to RCA team member

part of the briefing process; emergency response procedure; and implementation of a set of standard security mitigating principles that include the following:

- The need to ensure access to safe drinking water, through provision of bottled water and /or water purification tablets), clear explanations of other water-related risks and sound advice on mitigation
- Emergency arrangements agreed in advance, including a single focal contact in each sub team and with all participants having mobile phones for use in an emergency
- Contact details for nearest functioning health facility
- Contact details for drivers who can collect the researcher safely and at short notice
- Contact details for nearest police station
- Contact details for the sub team leader
- Where feasible, contact details of other development projects/organisations in the area to help with emergency situations
- Advice on the need for and appropriateness of prophylaxis (e.g. insect repellents, oral prophylaxis and nets), although team members will ultimately be responsible for their own decisions on the use of prophylaxis
- Advice on behaviours which maintain a low profile.

What we found in practice

New guidelines and protocols are comprehensive, and a welcome and necessary step to increase researcher safety. They go beyond what many qualitative research exercises put in place. However, it is unlikely that they will provide the contextual knowledge needed to fully address risk. We heard of a number of examples where local communities' suspicion of outsiders – foreigners (in Ethiopia) and outsiders from different ethnic groups (Ethiopia and Indonesia) - lead to situations in which researchers were detained by police, arrested, and chased from villages. These are not isolated events, and speak to a shallow consideration of the importance of contextual understanding for researcher safety. As a result, RCA fieldwork is fraught with ethical challenges. We discuss some these below in Section 6.2.5.

Though RCA has undergone ethical review processes for some select studies, there is no process for ethical review of RCA studies that would consider *contextual issues* related to research design. We found oversight and approval processes to be inadequate in this regard.

6.2 RCA Fieldwork

6.2.1 Being unobtrusive

Key Findings

1. RCA prides itself in researcher unobtrusiveness, and on some counts it has good reason to do so. Considerable time is spent preparing researchers to "fit in" while staying in local communities.
2. We find, however, that circumventing local protocols is a poor strategy. A lack of proper preparation and careful selection of households prior to the full immersion not only undermines the achievement of selection according to poverty proxies, but can make researchers more obtrusive, not less.

RCA protocols for being unobtrusive

RCA documentation and interviews with RCA researchers and members of the RCA core team suggest that team members enter communities independently on foot in order to keep the process 'low key'. They then spend time in the communities – at gathering or high traffic places such as tea shops, outside mosques or temples, near playing fields – chatting with people and trying to gain some preliminary insights into the community in aid of household selection. Researchers take considerable care in ensuring that people understand the importance of researchers staying with ordinary or "simple" families, participating in ordinary life, and not being afforded guest status.

Time in training and orientation is spent on exploring how researchers should conduct themselves in the community and with their host household, assisting in household chores, not passing judgement on local practices or reacting strongly to different ways of doing things, being accepting and not demanding any treatment different than that accorded to another family member of similar status. Researchers are schooled in not wearing or using or doing anything that would make them "stick out", for example not wearing expensive watches or sunglasses, not texting on their phones, not putting on sunscreen or using antibacterial hand-wash.

Many of these measures represent ethnographic good practice - though we consider some to be a little extreme (such as western researchers not wearing sunscreen) – and the RCA approach is a positive stand-out in this regard.

Are RCA researchers unobtrusive, and what are the implications of these efforts?

A visitor - whether they go in on foot or in a Land Rover - is going to be recognized as an outsider immediately, and treated as such. Local people, to their credit, might also wonder why a research from the capital is arriving on foot, when clearly they can afford other transportation. Power differentials are not solved by the choice of transportation; they need to be really seriously considered and acknowledged in a reflexive way.

Researchers often arrive late in the day, and sometimes in inclement weather, and have to walk some distance to reach a suitable research location far enough away from their colleagues. In the debriefing we observed in Jakarta, two researchers reported standing together in the rain outside the mosque as night fell trying to convince villagers leaving the mosque to host them; this is certainly not an unobtrusive way in which to find a host household, particularly for a female researcher.

Local protocols often require going through a community leader or official to obtain approval to stay in the community. Some researchers feel uncomfortable flouting these:

> *I never follow the RCA way of entering a community. I always go and explain my purpose to community leaders first. Once I have done this, I am almost always encouraged to go out and select my own household to stay with. Sometimes it takes me longer to explain, but I think it is important, respectful of local ways of doing things, to take this time and enter a community properly* (KII, RCA researcher).

However it appears that most researchers actively avoid local protocols for entering communities in order to try to remain unobtrusive and to avoid being afforded "respected guest status" with village notables. While this is a very real challenge that ethnographers face, we suggest that the "solution" of covert entry – as this is how it is sometimes perceived - is not an entirely adequate

one as it is culturally insensitive, can lead to suspicion rather than trust, and can re-enforce power imbalances. In both Ethiopia and Indonesia, covert entry has raised such suspicion that researchers were given no choice as to where they stayed, and were forced to stay with a village leaders who could "*keep an eye on me*" (KII, RCA researcher). In some cases, this led to seriously compromised researcher – and in some cases household - safety, as we discuss below in section 6.2.5. More positive experiences of trust-building have also been documented, for example in the longitudinal Bangladesh RCA. Reviews of the Bangladesh RCA studies[39] highlight four relevant findings concerning building trust. First, almost all host households felt uncomfortable initially to host unknown outsiders. Second, households were not exactly clear as to the purpose of the exercise or what would necessarily come out of it. Third, despite this, the experience of having an audience which was interested in hearing what they had to say was clearly a positive experience for many. Fourth, a relationship of trust and friendship was built up over time because, unlike other outsiders, RCA researchers returned. A more detailed discussion of these findings can be found at Annex 5.

6.2.2 Triangulation

Key Finding

1. We find that RCA achieves "triangulation over time" (particularly in longitudinal RCAs) and, to a very limited extent, "triangulation of people" (we assess this as very limited as the distinct voices of different research segments is often lost due to not faithfully recording conversations and aggregation in analysis).
2. Assessing RCA's achievement of "triangulation of methods" is somewhat more complex, but we find that this is generally weak and that this is a missed opportunity.

Triangulation of time

The RCA's claims

The RCA uses triangulation of time, conducting research over the course of several days and nights, which they suggests enables them to see how people actually live their lives and, with longitudinal RCA studies, returning over the course of several years to better understand trends. The RCA also suggests that the longer time spent in the field with households and communities, enables them to note differences between what people say and what people do, as well as experiencing more authentic interactions with research participants.

What we found

The RCA achieves good triangulation of time, as research is conducted over the course of several days and nights and, with longitudinal RCA studies, over the course of several years. The issue of seasonality is not much discussed; this is important to take into account in both one-off studies and longitudinal studies, in the latter particularly if the study is trying to explore trends and is required to attribute change to time, rather than seasonality.

A number of commissioners, consumers and qualitative research experts questioned whether it was reasonable to assume that, in 4 days, research participants would stop "performing" to such an extent that researchers were uncovering more valid data. One informant commented that "*there doesn't seem to be a recognition that the researchers were observing a*

[39] See Lewis et al., 2012 and 2018; Pain, Nycander and Islam, 2014; Arvidson, 2013.

'performance'…and this will have influenced what they observed and what people said. Indeed, it's well known that people will say very different things about the same topic, depending on who is listening" (KII, Qualitative research expert).

Triangulation of people

The RCA's claims

The RCA's claims to rigor rest heavily on having multiple conversations with different types of people, which enables them to triangulate people's stories and other information gathered during the research.

What we found

RCA achieves triangulation of people (multiple conversations with different types of people) to a very limited extent. Study teams are large and dispersed, which can help to reduce researcher bias though triangulating reports from a large number of team members and a very large number of research participants. Some studies claim more than a thousand research participants, though it is acknowledged that deeper discussions are held with a much smaller number. The researchers do not meet while in the community, which is seen as a strength of the approach as it maintains the uniqueness of experiences, observations and conversations, and increases confidence when findings from different researchers confirm each other (RCA, 2017a).

We suggest, however, that failure to record what different people say, or to surface similarities and differences in synthesis and analysis processes, result in the merging distinct voices, and the obscuring of differences on the basis of gender, age, ethnicity and other socio-economic and demographic characteristics. Further, it is also possible that, in the absence of field notes, less experienced researchers are more likely to agree with and bolster the findings of more experienced researchers.

Triangulation of methods

The RCA's claims

One of the hallmarks of RCA is that it does not employ formal data collection instruments. Researchers are instructed to not make use of formal discussion guides, questionnaires, or other specific research tools, as these are seen to create power imbalances between the researcher and research participants. Instead, the RCA claims to triangulate methods through using conversation, observation, and participation. In addition, participatory Rapid Appraisal (PRA) tools are sometimes used in RCA research to "*help to tell the story*" as does Digital Story Telling (DST). Increasingly, RCA is being combined with DST, a methodology that enables research participants to communicate stories that are meaningful to them through drawings and voice recordings.

What we found

We find the RCA's claims to triangulate methods through using conversation, observation, and participation to be somewhat misleading, as these are generally considered to be three elements of *one* method that is practiced by trained anthropologists: participant observation.[40] In most short-term ethnographic research, on the other hand, participant observation is complemented with a range of other qualitative tools – such as focus group discussions, pair interviews, multi-generational interviews, and case studies – and participatory tools – for example, mapping, ranking and sorting, role playing, and scenario response elicitation.

[40] Participant observation is not simply a case of participating and observing; it is a distinct method that anthropologists are trained in prior to fieldwork.

RCA reports refer to PRA as "*unconventional methods*",[41] though they are actually quite mainstream in mixed qualitative methods work. Shah notes that,

> *Using multiple methods is an aspect of triangulation that is particularly characteristic of most short-term ethnographic research, as the lack of time ethnographers have to develop reliable intuition and insight through their experience means that they have to ensure the insights they think they have gained through participant observation are validated by other methods* (2018: 16).

Some qualitative research exercises start with a brief period of immersion, before going into more structured research involving interviews, discussions, and PRA exercises, in order to both build trust and provide a more rounded contextual understanding within which to make sense of specific research findings (see, for example, Morse 1994: 5).

PRA tools are not easy to use well, and RCA researchers receive very little instruction on how to use them. How these tools have been used, and the results of the RCA exercises, are not generally presented in reports, so it is hard to judge their quality. The lack or presentation of PRA tools is a weakness, as good PRA should include documentation; the visuals generated are a joint product and not documenting them properly would seem to be dismissing their contribution to the research.

There are a number of places where the use of PRA tools would have enhanced research rigor. For example, commentators on the Uganda Senior Citizens Grant (SCG) RCA wondered why researchers did not generate kinship diagrams to help illustrate the networks of the core families in the study and show the relationship with other households (i.e. what kinds of interaction they had; who did they help and who did they receive help from; which households in the kinship network benefited from the SCG and how; etc). The review team is similarly surprised that the Haze Study did not include more pictures and diagrams created by and with children and adolescents. The Adolescent Nutrition and Physical Activity Study includes a much more interesting range of visuals illustrating PRA generated data. Shah notes for the RCA study that she assessed most closely[42] that,

> *Although the report says the team supplemented immersion and listening with other methods, data from these methods were not reported except in annexes where demographic host households is included, presumably garnered from an informal surve y and observation, though the information is not attributed, Data from the pile sorting is not reported at all, and an explanation for this absence is not given in the report (2018: 24).*

The case studies that we examined also suggest some weaknesses in the use of PRA methods. For example, the RAP 3 Midline study states that the way issues were ranked from most important to least was determined by what was discussed first and/or what was mentioned most. The graphic shows most talked about to least and explains: "*As these emerge naturally, it has made it possible to put these changes in order of significance to people, with the most important changes discussed first, as detailed below.*" Firstly, this is confusing: is priority based on what is

[41] see, for example, RCA, 2014
[42] This RCA study looked at education in Papua, Indonesia

mentioned first, or what is mentioned most? Secondly, we would question the rigor of this approach to determining priority issues, and ask why a standard pair-wise ranking – done by participants themselves - was not considered.

However, in the RAP Beneficiaries Study Report a good example of the use of PRA to enhance understanding is given (see Annex 6 for a description of this). This suggests the potential for PRA to enhance the RCA, provided there is sufficient training and guidance on this.

When RCA does use some supplementary tools, it appears that there are not always used rigorously to triangulate data collected through observation and conversation. This is true both of PRA and of Digital Story Telling. Findings from the review suggest that Digital Story Telling (DST) is geared primarily at advocacy, bringing people's stories alive in a way that engenders empathy in policy and program officers, and is not yet a truly integrated method and the data it generates, like PRA, is not fully appreciated. An example of this is that there is only one reference in the Haze RCA study report to data collected through DST despite the fact that a full DST exercise was done in parallel to the RCA. Commissioners similarly see these as two separate exercises with little integration:

> *DST was done after researchers finished with data collection. We found it hard to link the stories to the findings in the report. For example, we looked at striking findings in the report…but these were not revealed in the DST…And things that came out as striking in the DST, we cannot find these as much in the report…I'm not sure if the DST methodology use the same prompts as the RCA?* (KII, commissioner).

6.2.3 Respondent validation

Key Finding

1. There is little indication from reports or other documentation, or from interviews with RCA researchers, that respondent validation is a central part of an claimed non-extractive research practice, though some more experienced RCA researchers clearly do this to some extent.

RCA's approach to respondent validation

Respondent validation is purposive, and requires a researcher to act like a researcher; in other words, to not just pretend that they are there to have casual conversations, but rather to ensure understanding in a way that goes beyond casual "fact checking" and makes it obvious to respondents that they are part of a research process. It is not to be confused with probing, a technique used to elicit deeper or related information on a topic. Probing is a technique that appears to be adequately covered in RCA training materials. Respondent validation, on the other hand, is not covered adequately.

What we found

As we touched on above in 6.2.1 (unobtrusiveness) and discuss in more detail below in 6.2.5 (ethics of fieldwork), research participants are generally not told that they are participating in a piece of research, so are unable to truly participate in validation of those research findings. It was clear from speaking with more experienced RCA researchers that RCA's focus on discursive and

free flowing conversations can inhibit their ability to stop research participants to check meaning, and to present back to them what has been understood to check meaning.

Further, attempts to understand phenomena through accessing the meanings participants assign to them, as Guijt (2016) emphasises, requires explicit methodological focus and attention to the diversity and range of voices, uncovering respondents' structural positionality across gender, race, caste, poverty status, which we suggest the RCA does not do with any rigor or consistency.

A skilled researcher can explore in such a way that research participants themselves come to discover new meanings and new insights, even moving the process beyond one of validation to one of analytical co-creation, and thus ensuring the accuracy of voice. RCA has been criticised by others on this score. "*RCA seems to be seeing people as a "objects" to be studied, people with no agency…they are "the poor"…This just entrenches their unequal power relationships and elevates etic interpretations. You have to be skilled to lead people through their own analysis*" (KII, Qualitative research expert).

While this review suggests that there are clearly some RCA researchers who are able to do this, it does not appear to be an embedded practice, and there is clearly insufficient understanding of 'the poor' before starting a study.

6.2.4 Faithful and accurate recording

Key Findings

1. RCA note taking guidance is inconsistent, and researcher practice is variable.
2. Evidence suggests that memory and recall is complicated, and much less reliable than suggested by RCA practitioners.

RCA "rules" on note taking

RCA literature and documentation highlight "*the absence of note taking*" (Koleros et al. 2015: 7), and claim that taking notes and recording conversations would "*immediately distort the special spaces of interaction created*" (Masset et al., 2016: 174).

Further, note taking is identified as a source of researcher bias, whereas bias in recall is not seen as a significant challenge: "*researcher bias in writing field notes and discussions during the debriefing process [could lead to] conversations, experiences and observations made during fieldwork being interpreted by the researcher in a particular way, which may differ to some extent from the viewpoints of those in the community*" (RCA, 2014: 9).[43] It is the firm belief of RCA core team members that the debriefing process, if skilfully managed by the Team Leader, can trigger the most salient and valid memories of conversations, observations and experiences. A senior RCA core team member suggested that notes were less reliable than memory

[43] As previously noted, this document goes on to state that this risk would be mitigated by providing researchers minimal background or contextual information, as these details can encourage the researchers to make assumptions about the communities.

Note taking in RCA practice

The RCA guidance on note taking is not consistently communicated to researchers. A review of training materials and discussions with the RCA Senior Team members in Jakarta suggest that writing field notes discreetly during the evenings or other quiet times is an allowed (though perhaps not encouraged) practice. Yet this conflicts with what RCA researchers told us. Almost all of the RCA researchers with whom we spoke told us that note taking is discouraged full stop, There is, therefore, some considerable confusion amongst RCA researchers, and significant variation in researcher practices in this regard.

In practice, data recording appears to rely largely on researchers' memories of what they have heard, seen and experienced (subject to an internal process of what they select and choose to draw on, based on what they perceive to be relevant). Sometimes notes are jotted down prior to debriefing. However, we heard and observed during debriefing that some researchers do write field notes, either using mind maps or diagrams, or making notes by hand or on a phone. This is almost always done at night or other times when they cannot be observed doing so, generally at least once during the course of the 4 day field visit.

While some researchers shared with us how the debriefing process can trigger vivid memories, other researchers admitted that they cannot not remember details that they thought could be of importance to the research. One experienced RCA researcher explained, "*I figure I lose at least 50% of my data by not writing it down, and what stays with me is not necessarily the most important or the most representative. It might be the most surprising, or even the most confirmatory*" (KII, RCA researcher). Another researcher told us, "*Since we could not take notes while we are in the field, we can only take pictures secretly to avoid being noticed, and we download all information in the one day debrief, I think there are many things that cannot be captured more in-depth with RCA method*" (KII, RCA researcher). RCA researchers concurred that it is certainly not possible to recall accurately even the shortest and simplest of direct quotes if these are not written down immediately following the interaction or experience. This questions the veracity of some of the longer quotes in RCA reports.

To address this, some researchers – "*against RCA rules*" (KII, RCA researcher) - write notes as events or interactions are happening, or record conversations on their phones, both openly and surreptitiously, in order to augment their patchy memories. One experienced RCA researcher told us that she always used her phone to record conversations (with the permission of research participants), and that, as a result, the vast majority of boxed case studies in one RCA study were provided by her. Another RCA researcher explained that people sometimes wonder why they, as researchers, are not writing things down, suggesting that perhaps because what they are saying is not important enough to record faithfully and share with others.

How is this practice perceived by others?

Some commissioners were concerned with a lack of note taking, particularly combined with the relative inexperience of researchers:

> *While we appreciate the findings, we think the method fully relies on the memory and capacity of the researcher. There are no guidelines and checks. Recall is probably unreliable. We do value that field level information - discussions and observations - are very natural, without formality. But in 4 days there, if you do not take a lot of notes, or record information at all, this is a big risk. The capacity of the researchers has to be very good and we question that the training is of sufficient quality* (KII, Commissioner).

Another commissioner commented that he was unwilling to rely on the findings because he was not able to see the field notes nor to understand how the team moved from conversations, observations and experiences, to insights and implications (this is discussed below in 6.3). In their review of the RCA in Bangladesh, Pain, Nycander and Islam (2014) suggest that readers have no way of assessing the fidelity of the data RCA researchers collect, calling into question their conclusions.

What does the evidence on memory and recall tell us?

RCA senior team members argue that memory is greatly enhanced by experiencing daily life, in comparison to a focus group discussion, for example, which is out of context. While this may be true for a short, 1 – 2 hour encounter (that experiencing enables greater detail to stay in the memory longer than simply "interviewing"), there are a number of challenges to the claim that memory is sufficient to recall who said what, why, and in what context for more than a few key interactions and experiences. While this review has not gone into any depth in looking at memory, there is some interesting evidence about recall that we have gleaned from a very light literature review, and discussions with a behavioral insights specialist that call into question RCA's heavy reliance on memory in recalling conversations and events. These are presented in Annex 7.

According to a behavioural insights specialist, how important taking notes is, relative to allowing a less formal interaction to take place, is influenced by a number of factors. He suggests that the absence of note taking encourages interpretation, and that this requires an experienced researcher with a strong "mental framework":

> *The importance of taking notes, or not taking notes is going to come down to a number of factors. One is the nature and desired output of the research. Is the aim for more "factual" information? Or will the sponsors of the research be happy with the interpretation of the researcher. A massive factor is going to be the experience and quality of the interviewer. Are they experienced/competent enough and familiar enough with the subject to have a mental framework for it that will allow them to "store" the right things into their mem*ory? (KII, Behavioral insights specialist).

If the absence of note taking indeed does encourage interpretation, as is suggested, this works directly against another tenant of RCA, which is that researchers should not interpret information.

Common practice in ethnographic research

Short-term ethnography is by no means easy, and there are legitimate challenges to note taking. Those who have participated in RCA studies spoke of how intense and exhausting they are:

> *The researcher is exhausted after 4 days and 4 nights, and the hosts are probably pretty tired as well. You are not taking direct notes - you are trying to remember and note down in your free time. You are on call 24/4. You are also staying with very poor families, and this is not easy* (KII, Qualitative research expert).[44]

[44] TD

> *It can be really hard to just chat for 16 hours a day. So when we get saturated, we suggest to researchers to have some quiet time to walk around the village and reflect and have some time to themselves. This is also when they can reflect on what they might not have learned enough about* (KII, RCA core team member).

It is likely that this kind of fatigue and overload leads to significant memory loss, something that could be ameliorated by more frequent and structured note taking.

Experienced ethnographic researchers generally combine periods of intensive note taking with periods of little to no note taking. They generally do this based on the relationships that they have with research participants, the specific context, and with the way they think that note-taking is likely to be interpreted, and therefore whether it is likely to influence the data that they are able to collect (see Shah, 2018).

Qualitative research experts strongly opined that a more rigorous and formal note-taking approach was needed in the RCA. Anthropological good practice suggest that researchers write both technical and personal notes on at least a daily basis in order to minimise the otherwise inevitable data loss and bias that memory and personal impressions introduce. Shah (2018) suggests that this is of particular importance in short-term ethnography when a large amount of data has to be collected in a short amount of time and researchers have little time to build relationships and adapt to the context.

6.2.5 Ethical considerations in fieldwork

Key Findings

1. The RCA approach has paid greater attention to researcher safety in recent years, with a range of new guidelines and procedures for entering communities, explaining research, and obtaining consent.
2. While there have been changes in theory (and some practice) to how researchers introduce themselves in communities and obtain consent, it is clear that participants are often not fully aware that a "research process" is underway.
3. Despite the fact that the RCA recognizes that researchers' limited knowledge of the local context and support networks constitutes a risk, actions to mitigate this particular risk have not been initiated.
4. While there is an explicitly aim to include poor and excluded households in research, household selection procedures can exclude vulnerable populations. Compensation to households is not always sufficient to off-set losses from hosting researchers.
5. The RCA has weak child protection policies and procedures.

There are two primary concerns in regard to ethical considerations in fieldwork: 1. minimizing risk to researchers and 2. minimizing risk to research participants, particularly vulnerable populations. The latter is highly complex, and includes issues such as how households are targeted and compensated, how informed consent is obtained, and how vulnerable groups are included and protected in research.

Minimizing risk to researchers

The RCA's approach

There has been a recognition in recent years - presumably based on a number of negative researcher experiences - that risks to RCA researchers are potentially greater than risks associated with other types of field work.

There has also been a recognition that this increased vulnerability is due, at least in part, to how RCA research is designed and executed. RCA documentation suggests that researchers limited knowledge of the local context and support networks; potentially greater exposure to health risks; potential risks of being out of communication; risks associated with overnighting and hours of darkness; and undertaking work somewhat isolated from colleagues all contribute to risk (RCA, n.d.a)

RCA training materials claims that entering the community in low key ways - e.g. by foot with no prior arrangements - has "worked" in all the countries where RCA has been used (RCA, 2016d). The also claim that their approach builds trust, and "…*ensures* [our italics] that the power distances between researcher and study participants are minimized and provides enabling conditions for rich insights into people's context and reality to emerge" (RCA, 2016c: xiii). We are not aware of any discussions within the RCA team concerning the relationship between RCA's approach to entering communities and explaining their presence, and potential risks to researchers.

What we found

Despite the recognition that vulnerability can be exacerbated by research design and execution, we are not aware of the RCA team considering addressing potential sources of risk in this regard, for example addressing limited knowledge of the local context. None of these causes of increased risk are sufficiently addressed by new safety guidance (discussed in 6.1.5, above). Conspicuous by its absence in any RCA discussions of researchers safety is the recognition that female researchers face unique security risks, which can be significantly heightened in certain contexts.

Another cause of increased risk is lack of transparency on the part of researchers concerning their identities and the purpose of the research. There is ample evidence that lack of consideration of context and poor preparation, including failure to negotiate entry with local officials and pre-identify host households, has led to considerable challenges for both researchers and host households. For example, in the Haze Report (where it claims that power distances between researcher and study participants have been minimized), it also states that study researchers experienced suspicion in *a few of the areas* due to recent rumours and reports of criminal acts (such as kidnapping, illegal organ sales, and narcotics selling) which resulted in one team needing to move to a different area of the sub-district (RCA, 2016c). When we asked The Team Leader for this study about this, we were told that in 4 out of 8 locations "*researchers experienced suspicion… due to recent rumours and reports of criminal acts (such as kidnapping, illegal organ sales, and narcotics selling) which resulted in one team needing to move to a different area of the sub-district.*" (KII, RCA researcher) The Team Leader was unable to comment on what the effects of high levels of suspicious were on research carried out in those areas that teams did not move from (3 out of 8), but one can assume that this had a significant effect on data quality. Also, it is not clear if suspicion could have extended to the host families, during or after the researchers had left, as has happened in Ethiopia.

Risk to researchers and research participants

"It is not good if village officials…are aware of the researcher's arrival in the village"
My heart told me not to stay too long at the kiosk as the day was getting darker. Many people already noticed me there and, according to RCA, we were not allowed to speak too much with other people, otherwise our identity will be known quickly and reported to village officials. It is not good if village officials interfere and are aware of the researcher's arrival in the village.

– RCA researcher, Indonesia

"There is a very strict protocol"
In Ethiopia, you have to carry the correct letters with you, and negotiate entry at every level. You just can't walk into villages like you can in other countries. There is a very strict protocol, and local people and officials assume that if you have not followed it, it is because you are hiding something.

- RCA researcher, Ethiopia

"I was probably a spy"
When we were doing research on land title, which is a sensitive subject, many of the researchers were thought to be spies. This is common in Ethiopia anyway…I felt like I had built trust over the 4 days I stayed with my family and interacted with them and the neighbouring households. But I did hear that later, they were talking to others and said that I was probably a spy…Yes, I wonder now how much trust there really was.

- RCA researcher, Ethiopia

"I was taken by the police, and interrogated by armed villagers"
We initially visited the wrong village, because there is another village with similar name but different district. After I asked around, I got the correct village. But the head of village and his officials reported us to the local Police station there. They were worried that our activity related to religious tensions, as they saw foreigners walking around the villages. I was brought into the local police station for interrogation….eventually they took us back to the village. But at the village they yelled, interrogated and did a physical check and took my identity documents. They again questioned me, and group…came to interrogate me…I could not sleep well that night, afraid that those people might change their minds and grab me and take me out into the forest. In that same study, one of my friends also had a very unfortunately incident in his village…What I regretted from [RCA study leaders] is that they did not tell us before we left that this village and district is the sample. As researcher, I felt that I had been involved in the sampling process, and since I am local, I could have provided inputs that would have avoided us selecting such a sensitive area. They said it had been decided in Jakarta, using a formula to pick the sample areas…

- RCA researcher, Indonesia

"Every night my family was insulating me from the wrath of the villagers"
In one area, that was both Christian and Muslim and where there were religious tensions to do with land ownership (the subject of the RCA study), the researcher (not knowing) stayed in the one Muslim household in the village and this lead to huge tensions. The family another researcher stayed with was accused of being government spies due to him staying there. There were big arguments every night, and his household head had to defend him. Finally he was chased out of the village at 5 am on the last morning. He is still deeply concerned about what might have happened to his family as a result of him staying with them. Worse than that, he explained, one of the other researchers about 20 miles away was arrested an put in prison for terrorism after a gas cannister that some children were playing with exploded and killed two children. Local people thought she had planted a bomb.

– RCA researcher, Ethiopia

Researcher unable to explain her presence
There was a very inexperienced researcher, and I was really worried about her…[We] found out later that she never left the house where she was staying. Apparently the host family worried about her and were very protective of her. She could not explain why she was there, so she just shut down.

- RCA researcher, Indonesia

Minimizing risk to research participants

The RCA's approach

The RCA approach has come under considerable criticism for its covert entry into communities and presentation of the research to community members and, as a result, they have re-considered this aspect of fieldwork. As we see from the below, it is primarily out of concern for ethics regarding research populations that RCA has adjusted its approach, rather than risk to researchers, which should also be a concern.

From review of the implementation of early RCA studies in Indonesia, it was recognized that there may be some ambiguity in the way researchers explained their purpose to study participants and heterogeneity in researchers' behaviour regarding entering communities. This seemed to stem from a lack of confidence of inexperienced researchers in explaining their presence, a misplaced sense that they needed to be covert and a genuine desire to put people at ease and to ensure independence. A protocol for entering the community has since been developed and forms a key element of level 1 training. This was added to later with a more nuanced recommendation on covert/overt research informed by the continuing debate on the pros and cons of covert anthropological research which offers researchers advice on being more overt while maintaining informality. The approach now adopted ensures that study participants are aware that the researchers are indeed researchers and are there to learn from them and avoids any subterfuge which may have sometimes been present in earlier studies. Disclosure has to be balanced with providing an enabling environment for informal and open interaction and the independence and neutrality of researchers is fundamental to this.

- RCA, 2017a

RCA's household selection processes are highlighted in a number of reports. According to RCA documentation, household selection is intentionally meant to include poor people, and people easily excluded from research, such as female headed households, people with disabilities, children and youth, to ensure that the voices of those who are generally excluded from research are heard. Researchers are given a set of "poverty proxies" – sometimes simple, other times relatively complex - to try to find a host household for their 4 day visit. These might include - using the Household Financial Management RCA Study as an example - families with a range of children of school and work age who still live at home; families with one or more adults unable to work (incapacitated or in a caring role); families with 'missing middles' (e.g. grandparents looking after grandchildren full time); and families headed by women.

Ethnographers and other qualitative researchers have to take particular care in conducting research with vulnerable populations. The RCA has spent considerable time deliberating on issues of compensation, starting with the 2012 reflection process following 5 years of RCA studies in Bangladesh:

> *Subsequent RCA and immersion experience from other countries and other cultures including Indonesia has led to broad agreement which is felt to be respectful of the role the household plays. The principle adopted is to ensure the family is not out of pocket but reimbursement is made in a way which ensures no loss of face, is low key and does not compromise the informal non-guest relationship so carefully forged between researcher and host family. Even though costs are minimal, there are some costs associated with accommodating the researchers and cultural norms of reciprocity to be considered. RCA experience in other countries (including immersions experience in a variety of countries in Africa (Uganda, Sierra Leone, Malawi, Rwanda, Tanzania) suggests that the most appropriate form of compensation is to purchase basic food items locally and give these to the family on leaving as a gift and this has now been universally adopted as*

> *good practice across RCA studies. These may be supplemented with other small items e.g. the coloured pencils the team takes with them, possibly torches or mosquito nets etc. on a case by case basis. Copies of the photos taken of the family are given to them on subsequent visits and are regarded as important gifts. Before each study the RCA team discusses what form and value of these household compensation packages should take* (RCA, 2017a: 49)

RCA also claims to have rigorous policies and procedures in regard to child protection and safeguarding:

- *"As per American Anthropological Association Code of Ethics, RCA adopts an ethical obligation to people 'which (when necessary) supersedes the goal of seeking new knowledge'.*
- *Researchers 'do everything in their power to ensure that research does not harm the safety, dignity or privacy of the people with whom they conduct the research'*
- *All researchers are briefed on ethical considerations and Child Protection Policies before every field visit (irrespective of whether they have previously gone through this)*
- *All researchers sign two annexes to their contracts; i. Child Protection Policy declarations and ii. RCA study confidentiality and research data protection as part of their contracts*

-RCA, 2017a: 48

What we found

Despite a new, more candid approach to explaining researchers' purpose, current practice amongst researchers still appears to be variable. When we spoke with researchers, they were clear about not disclosing their real intent to the community, as this would interfere with people's free and frank discussions with them. Many researchers told us that when they introduce themselves to community members they are advised by the Team Leader to just tell people that they are from the capital, and want to learn about what life is like in this community.

RCA practitioners further claim that because RCA teams are "independent" that they need not reveal the projects for whom the research is being done, nor the identity of the commissioners of the research. This, we suggest, is quite a dodge, and a violation of the ethical principle of informed consent. One experienced qualitative researcher with whom we spoke has left the team due to concerns that "*uninformed consent*" was being obtained. Another researcher suggested that she felt that she was "*lying by omission*".

Further, as we noted in section 6.2.1 above, the "scramble" to find host households to stay with that often occurs with researchers arrive late in the day can only serve to exacerbate clear communication regarding the intent of the research. Shah suggests that "*this could make it difficult (and in some cultural contexts, inconceivable, for villagers to turn researchers away, thereby undermining any consent they may have given*" (2018: 21).

Taking ACFID's (2017) basic guidelines as an assessment tool, Table 3 presents a quick check of RCA practice in obtaining informed consent based on data collected for this review. There is clearly scope for improvement.

Table 3 –Informed consent quick check: an assessment of RCA practice

	Elements of informed consent	Fully met	Partially met	Not met
1	Research aims and objectives		✓	
2	Details of information that is being sought		✓	
3	How responses will be recorded and used			✓
4	The degree to which participants will be consulted prior to publication			✓
5	How findings will be communicated to participants (communicating the findings of the research to participants needs to be ensured, and the ways in which this will happen also needs to be communicated to participants during the consent process)			✓
6	Potential benefits and consequences of participation, including potential risks			✓
7	Reimbursements or incentives (if any) that will be provided for participating in the study - careful, culturally appropriate decisions need to be made on the nature of any recompense to participants or a community for participation, ensuring that any recompense is not perceived as an undue inducement		✓	
8	The name of the organization that is funding the research			✓
9	Contact details for someone independent of the research process for inquiries and complaints			✓
10	An explanation of the voluntary nature of the participant's involvement		✓	
11	The name of the investigator	✓		

While it is clear from reading RCA reports and speaking to researchers that researchers stay with a wide variety of households, staying with poorer households is not always possible, for a number of reasons. Some researchers are "forced" to stay with community leaders; some researchers arrive late in the community and have to stay wherever they are offered accommodation; sometimes poor household are unwilling or unable to host outsiders. Household selection can also miss particularly vulnerable populations - for example people living with disabilities, and households from marginalized religious, ethnic and sexual minorities - due to the ad hoc way in which selection criteria are often applied. The extent to which people with disabilities are included in research is not clear from reports, and when we probed on this issue in KIIs with core team members, they could only name two studies where they were confident that host households had included people with disabilities. This out of thousands of poor households included in RCA studies. Not only are these issues related to sampling problematic from an ethics perspective, we also explore examples of these issues with sampling and the impacts on research quality and effectiveness in Chapter 7.

Despite the clear attention paid by RCA study teams to the issue of compensation and evidence that this is largely appropriate in Bangladesh and Indonesia, there were reports that compensation is not always adequate in contexts that are, perhaps, less well-known to RCA teams. Researchers who conducted an RCA in Ethiopia suggested that families in famine hit areas experienced significant hardship feeding an additional person, and that the gifts of coffee and sugar left that

the end of the stay were not sufficient to compensate these households for the precious food they had lost.

A failure to reconcile non-intervention with the ethical imperative to protect children and safeguard other vulnerable populations in research has led to a number of serious incidents - we identified three during our Jakarta field visit - that researchers failed to report or to respond to appropriately and that, when reported, the RCA core team failed to follow up on. The suggests an extremely weak understanding of child protection and appropriate safeguarding processes at the time these interviews were conducted, even amongst the senior RCA core team and the RCA's appointed Child Protection Focal Point. The RCA's deep misunderstanding of ethnographic research ethics of moral relativism and non-interference is mainly, but not solely, to blame.

Part of the problem is that many commissioners (even those who have a core mission around child protection) appear to have skill deficits in relation to research governance, with particular weaknesses in assessing research ethics and ensuring accountability by researchers.

Due to DFAT's child protection policy on privacy requirements, details of the incidents cannot be discussed in this report. However, we note their gravity, and DFAT, on learning of the incidents, have been engaged with the RCA Team and Palladium in ongoing discussions. We understand a management response was prepared by the RCA Team to address the gaps and weaknesses in its child protection protocols and shared with DFAT. The reviewers have not seen this document and therefore are unable to comment on whether it sufficiently meets the concerns raised in this report.

6.3 The analysis and reporting process

6.3.1 Building, testing and evidencing hypotheses in data analysis

Key Findings

1. While RCA practitioners claim that iteration is one of the cornerstones of RCA rigor, it appears that iteration is not done with any rigor, and that RCA research does little to support collective analysis by the research team.
2. The debriefing process enables an impressive amount of information to be "downloaded" from the researchers to the Team Leader. However there appear to be considerable weaknesses, primary amongst them the somewhat extractive nature of the process.
3. Analysis in RCA follows a process of coding and charting that is unsystematic, or varying quality, and supports organization of data but not rigorous analysis.
4. RCA gives complete primacy to emic views over etic interpretation – rather than a judicious balancing of the two.

Before we "leave the field", as it were, in order to discuss the post-fieldwork analysis process, an important element of rigor, iterative analysis, needs to be considered.

Iterative analysis

What is iterative analysis and why do it?

Iterative analysis starts in the field and is ongoing during the process of data collection, allowing researchers to identify the most significant and relevant areas for further investigation, adjust their

focus, test and adjust hypotheses and introduce new methods as needed (Gittelsohn et al., 1998 in Shah, 2018). In short-term ethnography, as Shah (2018) notes, the limited time available to adjust and iterate has led to an emphasis on strategic, intentional, approaches to building cycles of feedback and reflection into the research process.

This ability to iterate in the field is seen as a hallmark of good qualitative research that utilizes a team approach. This is aptly described by McCallum et al. (2016). Their qualitative research team used a social mapping approach to guide program redesign in the Tingim Laip HIV prevention and care project in Papua New Guinea (funded by DFAT). They write:

> *At the end of each day, the team sat down together and share what they had learned and observed in their interviews, taking particular notice of the similarities and differences on a day-to-day basis. Through this daily reflection and as they met more people and visited new locations, findings that were either common or different across sites were notes* (McCallum et al., 2016: 132).

They report that the methodology developed has proven to be effective in increasing detailed program knowledge, as well as in shifting organizational culture relating to the manner in which programs are designed and put 'back on track' during their implementation (McCallum et al., 2016).

Do RCA researchers practice iterative analysis?

Qualitative researchers need a special class of analytical skills that can meet the demands of "messy analysis" in qualitative inquiry where context, social interaction, and numerous other inter-connected variables contribute to the "realities" researchers construct in the field, and eventually take away from the field (see Roller and Lavrakas, 2015). As RCA researchers generally play a minor role in analysis after debriefing – this being done by a small group of Senior RCA core team members, as we will see below – iterative analytical skills are critical.

While RCA practitioners claim that iteration is one of the cornerstones of RCA rigor, it is questionable how rigorously iteration is done. RCA researchers do not systematically use field notes to process early observations or to purposively plan further data collection. And nor do they systematically call on participants to help with interpretation (and nor is it certain that they have the skills to do so). And finally, RCA research does not go far enough to support collective analysis by the research team, as described above by McCallum et al. (2016). Any reference to iterative and adaptive thinking in RCA concerns researchers acting alone.[45] Many studies involve teams going to several sites[46], which provides, "*…a chance to discuss what is emerging and then explore these when you go back to the field. The Team Leader can flag issues to explore in second round. There is some really early preliminary analysis here*" (KII, RCA core team member), this does not provide the possibility of more deeply exploring issues that have arisen previously by returning to the same site.

Where research teams are unable to meet at the end of each day, it is common practice for lead researcher to "check in" with researchers by phone - to both conduct some iterative analysis and planning, and more generally support researchers - or for the research (in one site) to be punctuated by a mid-way check, which is what Shah (2018) recommends in the case of the RCA. This would be welcomed by at least some RCA researchers, who find the process of individual,

[45] http://www.reality-check-approach.com/related-resources.html
[46].

isolated research somewhat baffling in what is presumably a "team approach": "*I find it strange... that we do not consult with each other in the field to cross check, share information, and just generally support each other*" (KII, RCA researcher).

Debriefing

RCA's approach to debriefing

As soon as possible upon return from the field, and generally within a day or two, sub-teams conduct an "archiving day". Sub team leaders work with their teams to complete household information sheets and upload photos and any other visuals, and discuss what was challenging, what they learnt, what they would do differently. According to researchers, this process can help to jog memory, and some researchers use the time to jot down a few notes or do basic mind maps to prepare for the debriefing. The archiving day is followed closely – again, generally within a few days to a week – by a debriefing session.

Debriefing is the main moment in the RCA research process when data is "downloaded" from researcher to team leader. This happens in an intensive full day session facilitated by the study Team Leader. The Team Leader starts broadly, by getting an understanding of the study location and the host households. The bulk of the day involves researchers sharing conversations, observations, experiences related to the areas for conversation as well as expanding the areas for conversation based on people's inputs. This process aims to enable triangulation as the same topics can be explored through different researchers' experiences, conversations and observations.

RCA practitioners claim that the debriefing process requires continual reflection on bias, and is one where researchers have been empowered to challenge their own and others perceived bias, challenge findings, critically examine outliers and deviance, and ensure strong triangulation of findings. Further, the Team Leader is empowered to exclude findings which are felt to contain research bias (RCA, 2017a).

Researchers expressed how much they value the debriefing day, where they get to speak with fellow researchers who have been in the same community, but not in contact, about their experiences and what they have learned. This discussion processes was felt by a number of researchers to jog their memories.

What we found

Our observation, though admittedly of only one debriefing, was that an impressive amount of information is "downloaded" from the researchers to the team leader in the debriefing session. We did observe some considerable weaknesses in the process however, that we feel are critical to address in light of the foundational role that debriefing plays.

Firstly, we observed the process to be more extractive than discursive. There were a number of missed opportunities for probing, clarifying, and nuancing information. In particular, being clear who had said what and in what context. This was perhaps due to the fact that typical recording of data in the form of detailed field notes (and preliminary mental processing of these when writing and reflecting) had not been done by the researchers beforehand, so the one day debriefing needed to be "fast and light" in order to collect 12 days' of information from 3 different researchers on a range of different areas of investigation.

Secondly, while the RCA claims to explore multiple realities, it cannot do this if it lumps information together, over-categorizing, tending to dismiss differences between respondents, and

failing to ensure that 'deviant' experiences and observations are also captured. There are opportunities for identifying negative or deviant cases, or representative ones, if the debriefer probes carefully on whether there were "other cases" in any of the sites that were similar, for example, but, our observation is that the debrief process is too full, too rushed and too unsystematic to do this reliably. It is therefore not possible to know just how negative or deviant or representative cases are.

Thirdly, on bias, it is not clear how team leaders are able to determine what is factual and what is biased – beyond just asking "Did someone tell you this? Did you observe this? Did you experience this?" vs. "Is this your opinion". It should be noted that not all Team Leaders have been part of the field research process, so specific contextual and thematic expertise could well be limited, thus limiting their ability to "sense-check" information. Further, we saw no evidence that researchers were empowered to challenge their own or one another's bias during debriefing (nor indeed later, in the sensemaking workshop). While we only observed one debriefing session, this lack of reflexivity and consideration of positionality has also been noted by other reviewers:

> *The RCA reports show no evidence of reflection on the RCA team's p art on how their actions, identities and choices may have affected their hosts, or their findings. They do not consider how their own status, ethnicity, age, gender and other identifiers affect the relat ionships they have with their hosts…Other than in the reflection report from RCA Bangladesh (Lewis et al., 2012) there is no engagement evident in the reports I reviewed of the RCA team considering how their potentially ethnocentric perceptions, and their o wn experiences, backgrounds, relationships and preferences in t he field, may have affected their hosts or their findings. This kind of reflexivity is important for immersion- based research, as it works to mitigate the dual risks of conducting unethical research that harms participants and of obtaining inaccurate results due to unaccountable subjectivity* (Shah, 2018: 13).

We observed that there is a possibility of "host household bias" in that discussions centred around data collected from host households and a few other focal households. Several qualitative researcher experts who have participated in RCA studies raised this as potentially problematic due to the "*high risk that you become more empathetic to your close family members and see their views as more valid…the ultimate truth*" (KII, Qualitative research expert). Some RCA researchers also questioned the amount of bias that they felt could creep into the synthesis and analysis process: "*I did wonder, is what is emerging really what is important to people or is it what is important to the researcher?*" (KII, RCA researcher).

There was a feeling amongst a number of researchers and others close to the RCA process the Team Leader leading the debrief can unduly bias the process: "*It's all…[the Team Leader], or nothing. Her biases are not being interrogated by her. The debrief session is the crux and this is almost always led by [the Team Leader]*" (KII, Commissioner and Qualitative research expert). Interestingly, there has been an attempt to mitigate Team Leader bias in at least one RCA study. In the Household Financial Management RCA Study, they intentionally used 5 different de-briefers in order to minimize tendencies towards confirmation bias, which might have emerged if all researchers were de-briefed by a single de-briefer. This could, however, reduce the ability to do comparative analysis of information generated across sub-teams.

Synthesis and analysis

RCA's approach

The process for moving from a field immersion to a finished piece of analysis as presented in a report is described by RCA practitioners as using established framework analysis procedures involving the following stages of qualitative analysis:

i. **Familiarization** - This phase involves immersion in the data, re-reading case studies, field diaries and reported speech. Experience from RCAs conducted indicates the critical importance of the sub teams sitting together and reviewing and reflecting on the data extensively together and then sharing with the entire team to gain a comprehensive understanding of the data as a whole.
ii. **Identification of thematic framework** - This phase involves identifying key issues, themes and categories raised by the study participants, which emerge out of the familiarization phase.
iii. **Charting** - This phase involves re-visiting the entire set of data and placing summaries of the views and experiences shared by the study participants inside the chart of themes (categorization).
iv. **Interpretation** - The conventional fourth step in qualitative data analysis is 'interpretation'. RCA sources contradict each other on this score – while some claim to "*eschew interpretation*" (KII, RCA core team member), others state that the interpretation phase – while exercising caution not to overlay the researcher's interpretation of the insights -attempts to draw inferences from the charted summaries (RCA, 2016e).

The RCA core team has produced guidance for Level 3 practitioners on coding debriefing notes and clustering codes into themes. The next step is for the Team Leader and one or two other senior RCA practitioners to start to cluster codes identified in the debriefing notes into working categories that, at least in theory, form a working analytical framework.

Following this preliminary coding and charting, it is good practice for researchers who have participated in the study to come together for between a few hours and a day in a "sensemaking" workshop. The purpose of this gathering is to "*…start building a bigger picture using the stories of all the different communities. We take a day together for us to get a sense of the big picture, another layer of insights comes out at this level - memories are triggered, similarities and differences are probed*" (KII, RCA core team member). At this workshop, they are asked to take the position of study participants and identify emerging narratives from their studies. This process is thought to enable sense making and to ensure that researchers do not overlay their own interpretations on the findings.

Following the sensemaking workshop, there is some further coding and charting to incorporate anything that emerged from the sensemaking workshop, and further discussions between the two or three senior researchers to "*share ideas, and discuss how to structure the report*" (KII, RCA core team member). The study Team Leader also generally goes back to some or all of the researchers to fill in gaps in information, and to request "boxed stories" on different themes.

What we found

Much of this appears to follow minimum good practice, at least in theory, such as ensuring that "*At least two researchers should independently code at least some of the de-brief notes and compare their coding (in an effort to increase rigour in analysis)*" (RCA, n.d.b: 2). An examination of some of these materials suggests that coding and charting practices are variable and not as systematic as claimed. There is no doubt that there are a number of cooks in the kitchen, which can reduce bias and increase credibility; however practices around coding vary significantly, in terms of what codes are used by what practitioners, how many people code, and if these codes are ever compared and reconciled. Further, charting appears to be done in a way

that lends itself more to an organisation of data within a reporting structure than to the rigorous building, testing, and evidencing of hypotheses.

Based on what RCA researchers told us, sensemaking workshops are also extremely variable in terms of process and quality. We observed one sensemaking workshop; a description and our impressions are below.

Review team observations on a sensemaking workshop

The workshop that some of the members of the review team observed in Jakarta in December 2017 involved a number of participatory research style exercises that were aimed at enabling researchers to (*inter alia*):

- map where migrants from rural locations were migrating to and where migrants to urban locations were migrating from (Figure 1).
- categorise migrants based on the nature of the work they were engaged in, e.g. "cash casual commuter" (Figure 2).
- document what people in different groups (as per the above categorisation) told researchers were the drivers of migration, barriers to work, specific jobs they engaged in, what their lives were like (accommodation, food, costs and social life) (Figure 3), and what the pros and cons of migration are for different groups.

Figure 1

Figure 2

Figure 3

The synthesis workshop consisted largely of a series of participatory exercises (that, in our experience, a researcher would typically conduct in the field with research participants), with the team leader using researchers as "quasi research subjects" and doing rough counts of cases using different coloured stickers, for example for female seasonal migrants, and male permanent migrants. Without faithful and accurate recording of data in the field it is hard to assess how accurate these quasi statistic were.

One was left wondering why these exercises were conducted in the workshop, relying on researchers' limited understanding, lack of personal experience, and imperfect memories (particularly around numbers, when community members could have done this exercise based on their own personal experiences and thereby maximize participant and community control of the research and knowledge production processes.

Further, data was synthesized in a way that made it extremely difficult to segment based on socio-economic or demographic criteria and thus was not able to illuminate important differences in the views, experiences or priorities of different groups of participants. Our observation was that this process lacked depth and rigor, and did not get the best out of the researchers or the data that they collected.

Analysis following the sensemaking workshop – with 3 senior researchers working together – is presented as a way to off-set bias. Quite a number of commissioners and consumers felt, however, that there were significant biases in analysis, and some stridency around findings that reduced their confidence in RCA's objectivity. One suggested, "*The Team Leader already has*

hypotheses that she is discussing before the research, and then you see the report, and you see the hypotheses proved in the report (KII, RCA Commissioner and consumer). Another shared, "*I felt that that there was a great amount of stridency around interpretation and presentation of findings… I felt that they came to the findings with a particular view, leading them to read results in a certain way. They were not reflexive about this at all"* (KII, Commissioner).

Several RCA researchers and former researchers supported this view, with one remarking that, *"the Team Leader went in with strong ideas about what we would find, and, not unsurprisingly, this is what came out of analysis"* (KII, RCA researcher).

This issue of "who's knowledge counts" when it comes to "the final analysis" was also raised by several RCA researchers. One researcher explained, "*I feel like there is a distortion of local voices at the end because the RCA team leaders are the ones writing the reports. They all buy into the same ideology. They are all powerful and they are all non-Indonesian. There are many researchers that disagree with them, and who have really good analysis…This is a big issue* (KII, RCA researcher).

Implications of RCA practice for good quality analysis

The RCA claims to consciously include a range of study participants whose views *when combined together* will produce rich, granular detailed data (an emphasis on plural and multiple perspectives) (RCA, 2017a). But it is this combining together that is problematic for analysis. To be precise, it makes it easier for analysis because analysing according to segment is more difficult; but it makes it harder though to gain any interpretive depth. By presenting the views of "the people" or "villagers" or "farmers" rather than a more detailed segmentation based on sex ("female farmers"), age ("older people", "10 – 12 year old adolescent girls"), or other socio-economic or demographic characteristics, no one's views are faithfully represented or, if they are, it is unclear who's. Further, even in disaggregated categories, we would expect multiple views and heterogeneity.

There appears to be some confusion at the centre of the RCA about reflexivity and what this means for analysis and interpretation. Senior RCA practitioners appear to interpret Ruby's call to "stop being 'shamans' of objectivity" (Ruby 1980: 154 in RCA, 2016f 2016: 10) as requiring them to enable emic (insider) perspectives to emerge (a worthy endeavour) but to completely limit etic (outsider) interpretation or validation (Masset et al., 2016). Reflexive practice, however, requires "*…working from within institutional processes and social relationships in order to gain insights as an insider that could then later be analysed from the perspective of an outsider*" (Mosse and Kruckenberg, 2017:211). Roller and Lavrakas (2015) write that in order to derive meaning from data, qualitative researchers, as outsiders, need to evaluate a wide range of variables, including: the context, the language, the impact of the participant-researcher relationship, the potential for participant bias, and the potential for researcher bias.

RCA's presentation of descriptive data with little interpretation based on understandings of power, positionality and context, yields few meaningful ethnographic insights. The RCA approach generates a collection of narratives, interesting, at times informative, but not analytical, and prone to bias due to a lack of "*accountable and conscious theoretical interpretation*" (Shah, 2018: 19). A descriptive summary of what people say and do, rather than interpretive analysis – which demands an etic lens - is even less adequate in more complex circumstances where the issues under study are highly context dependent, as noted by Seeberg and Ray:

> *The relevance of ethnographic analysis rests on its alignment with what is important in the lives of people inhabiting the field under study. This*

> *does not necessarily mean that the ethnographer has to agree with the views of the interlocutors. A good ethnographer should be able to represent these topic of importance in such a way that it becomes possible to understand how they are perceived and practiced locally, and [our emphasis] how such perceptions and practices are embedded in the local political economy as well as in broader social and cultural contexts"* (2016: 99 and 100).

This shortcoming carries forward into RCA reports, to which we now turn.

6.3.2 Reporting

Key Findings

1. RCA reports are rich and include an impressive amount of detail. While RCA reports are interesting, and at times informative, without understanding power, positionality and context, little meaning can be generated, and this is where RCA reports fall considerably short.
2. The presentation of data – particularly in the form of quotes, boxed case studies, and tables – is, at times, problematic.
3. RCA reports present a new perspective, one which is often not seen by policy makers and programmers; however there is some blurring of lines concerning the presentation of "findings", " conclusions" and "recommendations" that can be confusing.

RCA reports are rich and interesting. Commissioners and consumers, as well as expert reviewers, are - for the most part - impressed with the level of detail that is presented. Shah writes that, "*Overall, whilst the team did not answer many of the questions listed as research aims, I was impressed with how much the research team was able to find out in a short space of time*" (2018: 25). We were similarly impressed with the level of detail in some of the case study reports that we reviewed, in particular the Village Law and Household Financial Management studies in Indonesia and the IP-SSJ report in Nepal.

Unfortunately, these rich and compelling accounts need to be seen in light of analytical weaknesses. We will not belabour this point, which we feel has been made adequately above, but instead turn to other aspects of reporting: presentation of data; and presentation of findings, insights, implications and recommendations.

Presentation of data

RCA reports include photos, diagrams, tables and boxed case studies. How data is presented is limited, however, due to the nature of the RCA methodology, which is based largely on conversation, observation and participation in daily life. While photographs are included - and these are effective ways to present information to those who are not familiar with the (often) rural settings in which research takes place – other visuals are somewhat limited, or are poorly presented. For example tables are often presented with no explanation of how the data was generated and by whom. A cogent example from the Haze RCA Study Report is presented in Annex 8.

While there are some examples where a more detailed segmentation of participants is presented in reports, for example the Adolescent Nutrition and Physical Activity RCA Study presents data in

tables according to socio-economic status, age and sex and remoteness, but analysis according to these differences is largely absent. However, even this presentation of data is not found in most studies that we reviewed, and RCA has been critiqued on this count by other expert reviews, such as Shah (2018) and Pain, Nycander and Islam, who write, "*Voices cannot just be 'heard' and 'quoted' as the RC reporting had a tendency to do, without careful consideration of how and why things are being said…This critical consideration and a more theoretically informed discussion of how interpretations are being made is missing from the RC reporting*" (2014: 24). The draft RCA study of the Senior Citizen's Grant (SCG) in Uganda has been similarly critiqued:

> *…The methodology is meant to interact with a wide range of people, including all those in households as well as others in the community. And, we know that the SCG [Ugandan Senior Citizens Grant] helps others, not just the older persons. But, there is very little in the report on the perspectives of others, the lives of other people in the households with older persons, the impacts of the SCG on others (i.e. the indirect beneficiaries). Where are the voices of working age people in the household? Where are the voices of children in SCG households? This definitely needs to be included* (pers. comm., Qualitative research expert).

Not knowing the prevalence of a finding is problematic, particularly for monitoring, evaluation and learning (MEL). RCA studies typically speak to hundreds, and sometimes thousands of people, and there is a missed opportunity to "tally" results and present them as quasi-statistics. This can only be done reliably, however, where there has been faithful and accurate recording of data in the field.

Absence of quasi-statistics

Failure to present quasi-statistics appears to have been particularly problematic for the Uganda Senior Citizen's Grant (SCG) RCA study, which was to serve as a baseline for some areas. This RCA study has been roundly critiqued for a range of short comings, including the failure to present quasi-statistics:

> *I'd also expect to see a lot more numbers in the report. Qualitative research should also include numbers: it's a great opportunity to check, in detail, what is going on in particular communities. Numbers could include, for example:*
> - *Numbers not registered or receiving the SCG who were eligible;*
> - *Lateness of payments;*
> - *Regularity of payments;*
> - *Distances to paypoints;*
> - *Number of grievances submitted to programme and number resolved*
>
> *These numbers can also be compared to the administrative data. Indeed…[for a baseline report] I'd have expected to see much more administrative data in the report: e.g. how many recipients in each community; when were payments actually made; how much was paid in each payment; how many complaints; etc.*
>
> - pers. comm., Qualitative research expert

While boxed case studies are very much appreciated by commissioners and consumers of reports in that they bring people's lives and stories to life, it is not clear how boxed case studies – which form an important part of data presentation in RCA reports - are selected, whether these present anecdotes supporting the writer's hypothesis or argument, or whether they give undue attention to surprising findings, at the expense of more common ones. One RCA researcher's claim that she provided the vast majority of case studies suggests that these would be representative of only a

very small sample of the overall research population. She implied that her case studies were over-sampled in the report due to her data collection methods (she used note taking and voice recordings during interviews) and to her writing ability, rather than to the application of any selection criteria.

As noted in previous reviews,[47] a central issue is the extent to which recording after the event is selective, if even only on grounds of interest to the study, and accurate, i.e. what is heard or seen (and not seen and heard) and how that is reported. Given the centrality of the use of quotes in reports selectivity and accuracy becomes an important issue. The RCA's use of quotes is problematic in that, unless researchers are recording conversations verbatim, these cannot be claimed to be accurate in any way. Some quotes are so lengthy that it is not possible that these are, in fact, quotes.

In Shah's review of the Papua Education RCA study, she notes that, although the report does improve considerably on the earlier Bangladesh reports in attributing quotes and findings, there are still a number of places where the report puts phrases in quotation makes and then attributes them to "students". This leaves the reader asking, "How many students? Did anyone actually say this 'quote'? Is it an amalgamation? Did gender affect this perspective?" (2018: 25).

Some quotes are not attributed, leaving out details that are critical to forming an evidence base that makes interpretation of the report's findings credible, meaningful and applicable. Without them, there is no way of knowing whether a particular quotation or perspective is an outlier, or how to interpret its implications (Shah, 2018). As with boxed case studies, there is some confusion as to why quotes are selected, and what their purpose is. As expressed by one key informant, "*I'm not sure whether quotes are illustrative of one person's view or are selected because they encapsulate the views of many different people, or because they are representative of an issue in the region*" (KII, Consumer).

There is some variation here between studies and they are not all uniformly lacking in segmentation. For example, in places the Haze report drew out differences between adult and child responses:

> *While adults told us that they felt children and the elderly could experience some minor discomfort with the haze, most parents told us their children were 'strong and did not get sick' during the 2015 haze. However, when talking directly with their children, they shared they had experienced coughs, fever, teary eyes, sore throats, had trouble breathing, and chest pain* (RCA, 2016c: xiv.).

The presentation of insights and implications

As with much of the literature on RCA which, at least more recently, appears to be focused largely on distinguishing RCA from other qualitative approaches, RCA tries to distinguish itself by actively not providing recommendations, but instead "insights and implications" that are grounded in what people themselves share and show. This is appropriate, as "*insight contributes a new perspective, whereas evidence contributes to new knowledge*" (Morse, 2005: 6). RCA is clearly trying to present a new perspective, one which is often not seen by policy makers and programmers.

[47] See, for example, Pain, Nycander and Islam (2014)

In an attempt to increase clarity concerning information report to the researchers, vs. what the researchers themselves observed, some reports present implications under two different headings: "what people told us, and "what we observed". We suggest that this is a good way to present findings, but that it currently has some shortcomings, as illustrated by the Village Law Study, which presents findings in this way. The report is confusing in that "what people told us" reports only what village officials told them, and what anyone else told researchers was recorded as what researchers observed. For example, "*In particular the youth shared their frustrations that their needs were not being addressed*" was presented under "what we observed" (RCA, 2016g: xviii). Being clear on whether insights and implications come from "what people told us" vs "what we observed" is important, as listening requires an ethical stance that seeks to resist judgment, while observation inevitably tends to bring in more of the observer's own subjectivity (Lewis et al., 2012). What is told, provided it is faithfully recorded, is therefore often more objective than what is observed (and needs to be interpreted to be told).

Though RCA core team members were clear that they do not present conclusions or recommendations, and reports often state, "*The following implications emerge from the in-depth conversations with families living in poverty and are presented from their perspective rather than from an external normative perspective*", there are numerous examples of "implications" presented in RCA study reports that appear to be the RCA research team presenting their recommendations. We give some examples below.

Implications or recommendations?

Haze Study, Indonesia
"As children seem to like wearing printed cloth masks more than the hospital-type masks, one way to influence them to wear masks potentially could be through promoting these masks as a 'cool' option and/or developing 'cooler' designs for more effective mask types. Simple messages explaining the need to wear masks and linking this to longer term health issues may also help adults understand the need for change. People need to know why preventive measures are important before they can change their behaviour."

Household Financial Management Study, Indonesia
"Families struggle when the chance to earn cash is circumscribed by child care suggesting a need for baby and toddler care within the community."

"While women are typically the decision makers in the family, the whole family would benefit from support to review together their expenses, especially on non-essential items."

"...these sorts of reviews might assist families in their financial planning..."

"…the example given here of the young man…suggests a model for peer mentoring by role models."

While some reports attempt to distinguish between "people's recommendations" and "researchers' recommendations", others provide recommendations that are neither attributed to "people" nor "researchers". For example this recommendation from the IP-SSJ report is not attributed: "*…strategies to make people feel the police are less 'outsiders' could be considered , including recruiting more police with local language proficiency and cultural context knowledge, provision of motorcycle transportation to allow greater mobility and more regular presence in the village and trust building efforts"* (RCA, 2015a: 45). It is highly relevant to understand where this recommendation came from.

Uptake of findings

RCA findings can be packaged in a way that commissioners find difficult to feedback on, disseminate, and action, as expressed by this key informant:

> *The ESP [Uganda] study was a big dump of 90 pages of all the information they collected, so not super useful in terms of format. We really needed a policy brief, but might not have asked for this in the contract. We didn't realise how difficult to consume the report would be though, so we've gone back and asked for a powerpoint and online discussion. We need to be able to share the findings with a wider group of constituents, and just can't do this with the current format* (KII, Commissioner).

Over the course of the last several years, the RCA core team have experimented with presenting findings in different ways. Having been criticised for their extremely long reports, they have produced short and punchy policy briefs to accompany longer reports, and are seeking to present findings in a number of different fora through presentations and discussions. Policy briefs are generally extremely well presented documents, and well-received by policy makers. Interestingly, their key findings (insights and implications) are sometimes not entirely consistent with the full RCA studies, so it can be difficult for consumers to trace summary findings in briefs back to supporting evidence in full reports. Some briefs also fail to capture important insights from full reports, which is a concern if the briefs are what most policy makers consume. For example, the Household Financial Management brief fails to capture key insights presented in the main report related to savings, debt and credit.

The absence of interpretations, the "so what" results of analysis, and the recommendations that emerge from this, makes it difficult for commissioners and consumers to take action, or indeed may encourage them to draw their own – potentially erroneous – conclusions based on bias and very partial data. This is a concern shared by other qualitative research experts:

> *The aim of qualitative research for development programming and policy is to surface the incredible heterogeneity of how programs and policies play out - different religions, cultural differences, geographies, ages, sexes, etc. It needs to reveal this complexity and then use etic experience to say 'so what does this mean for you, as program implementers and policy folks'. RCA strikes me as presenting somewhat of an ambiguous image – is it a rabbit or a duck? Is it an old woman or a young woman? You can't hand complexity and ambiguity over and not interpret, not tell policy makers 'it's a duck, and this is why you should care that it's a duck and not a rabbit'* (KII, Qualitative research expert).

However, we are even more concerned with conclusions and recommendations that emerge from weak analysis and interpretation by researchers who have little relevant experience or expertise. Indeed, as noted by Shah, making recommendations when you are not qualified to do so has ethical implications:

> *…the report frames the authors as experts who can represent indigenous "realities," and advise policy makers accordingly. If recommendations are acted on, this could have significant implications for people's lives…Recommendations on the basis of RCA findings – unless they are immediately apparent from the evidence collected – should be made by*

people who are experienced in development, knowledgeable about the context, and perhaps most importantly skilled at translating between different forms of knowledge, committed to doing to thoughtfully and ethically, and able to draw on a range of sources (Shah, 2018: 23).

We discuss RCA's effectiveness as a policy influencing tool in more depth in Chapter 7.

6.3.3 Transparency

Key Findings

1. Greater transparency has been achieved by introducing a more formal archiving process, but a lack of systematic note taking and rigorous debriefing leaves room for improvements in this area.
2. Reports do a reasonable job of highlighting research limitations; however, both a failure to properly consider how researcher positionality might have influenced the research, and a failure to consider the potential magnitude of limitations and impact of these weaken transparency.

RCA's approach

It is clear that improvements in the documentation process have been made since the 2014 evaluation of the RCA approach in Bangladesh found that "'*transparency in the method, the process and the results…' has not been met*" (Pain, Nycander and Islam, 2014: 22). With the addition of more formal archiving processes greater transparency has been achieved. However this review found that there is still considerable room for improvement.

What we found

Much of what we have discussed above speaks to the issue of transparency in reporting. There are significant challenges due to the fact that the RCA approach does not encourage systematic note taking in the field, and that the debriefing process is not rigorous in ensuring that information can be traced back to sources, nor that information can be tallied to present quasi-statistics. What we see in reporting – imprecise accounts of who said what, where data came from, etc. is largely a result of a lack of precision in relation to the documentation process.

Another way in which the RCA approach strives for transparency is through highlighting research limitations in their reports. How well this is done is quite variable across different reports, with the Household Financial Management and Village Law RCAs doing a reasonable job of describing limitations and mitigating measures. Little acknowledged as a limitation in reports, which is surprising, is how the positionality of the researcher might have affected the research process. The caste, religions, ethnicity, sex and age of the researcher, amongst other things, cannot fail to influence the research process. Shah (2018) raises a similar issue in the context of the Papua Education RCA. She notes that some of the RCA research team were Indonesian, and "*given the politics of Papua and West Papua, especially the tensions and perceived injustices that exist between many indigenous Papuans and the Indonesian government, the ethnic identity of the researchers will have had a significant impact on the findings and may even explain some of the inaccuracies…However, there was no reflection on this dynamic in the report*" (Shah, 2018: 25). The same could be said of the IP-SSJ research in the Terai region of Nepal, with tensions between Madhesis and Pahadis.

There is an acknowledgement in the Haze RCA study report that "*One of the locations (peri-urban West Kalimantan) had a more pronounced male dominant culture which meant that groups of women were more reluctant to speak directly with male researchers, and that wives would often defer to their husbands during, or withdraw themselves from, conversations with a researcher where both husband and wife were present*" (RCA, 2016c: 8) and that "*In a few of the locations, people used local language limiting the opportunity for researchers to take part in group conversations*" (Ibid), but whether these were serious limitations, how specifically they affected research, and how limitations were mitigated is not discussed. A further limitation – that it was difficult to speak with men about haze –was highlighted by RCA core team members in key informant interviews; however this was not highlighted as a limitation.

6.3.4 Ethical considerations in analysis and reporting

Key Findings

1. The RCA archiving processes and secure data storage are adequate to ensure privacy and confidentiality.
2. The RCA has no feedback mechanisms to communicate feedback to research participants and to enable them to contest how they are being represented. While it is clear that the intention is to present research participants respectfully in reporting, this is not always the case.
3. The RCA has an internal peer review process that serves the purposes of client quality assurance but does not review content from a contextual or technical perspective. External peer review processes that could serve this purpose are absent.

Ethical considerations in analysis and reporting include privacy and confidentiality, cultural sensitivity and consideration of context, and ethical review processes. Our discussion here focuses on how people are represented in reports, as we have already suggested that the lack of respondent validation and participation in iterative analysis is problematic both in terms of rigor and ethics.

Archiving processes and secure data storage are adequate to ensure privacy and confidentiality. RCA core team members are aware that preserving anonymity requires that there be no link between the data and the source, and work hard to prevent the identification of communities, households or individuals in reporting.

RCA, again like many other qualitative research approaches, has no process or methods for communicating research findings back to participants nor for actively engaging participants with the findings and their implications. From an ethical perspective, this further reinforces conventional power imbalances found in research, with RCA falling into the (common in qualitative research) trap of presuming to represent a whole host of diverse people with diverse interests, experiences and perspectives on the basis of only brief encounters. It gives researched people no chance to represent themselves or to contest the findings. This appears to be due more to commissioners seeing little value in this, and therefore being loathe to pay for researchers to return to communities to communicate and discuss findings, than a lack of appetite to do this on the part of RCA practitioners.

Respectful representation in reporting is an important ethical consideration. While level 3 training materials suggest that those who write reports "*…need to be respectful of every word we write*" (RCA, n.d.b: 3), there are a number of places where reports refer to research participants in

reductionist and somewhat derogatory manner, for example "The Punk Boy" or "The Milk-Fish Girl" or “The Tomboy” in the Adolescent Nutrition and Physical Activity RCA study (2016b: 20 and 21). In the Uganda senior citizens grant RCA study, one expert reviewer commented: "*I found the images of older persons in the diagrams a bit off-putting (and potentially disparaging). I also don’t understand the wheelchair symbol: is it referring only to those in a wheelchair, or to disability, or to what? And, if an older person was blind, but active, where would they fit?*" (pers. comm., Qualitative research expert).

In terms of peer review, quality assurance of RCA study report is carried out through internal peer review, with special concern to ensuring the research retains the positionality of people themselves. We were told by RCA core team members that those who are peer reviewing have often not been involved in the research, so it is hard to understand how they could know whether the research retained the positionality of people themselves. One can only assume that this refers to ensuring that the reporting guidance in terms of representation is followed, to at least present what was heard, observed and experienced as reflecting the positionality of people themselves.

Good practice is for research that addresses sensitive issues or topics, involves vulnerable groups, uses considerable participant time, or is innovative or largely exploratory to employ a more formal ethical review and approval process (ACFID, 2017). Clearly, RCA qualifies for a more formal ethical review and approval process; however, this is generally absent[48]. Using study Reference Groups as a mechanism for external peer review appears to be completely ineffectual.

[48] One of the case studies for this review was subject to ethical review. Overall, we were told that 8 RCA studies has been through ethical review. We are not able to judge the rigor of these review processes.

7 Outcomes and impacts: Does RCA deliver and in which contexts?

The quality of the approach and methodology both in theory and practice is important in itself, but ultimately matters most for how it translates into the outcomes and impacts set out in the ToC in Chapter 4. RCA's ability to deliver on the theory of change can be looked at through two lenses. The first lens relates to RCA as a research methodology, it's simplicity, flexibility, depth, and respect for voice; its ability to build trust, and to surface information not usually generated through normal monitoring or other research approaches (e.g. unintended and indirect effects and negative changes, sensitive topics and changing social norms, surprising and counterintuitive findings). The second lens relates to RCA's ability to effectively communicate research participant experiences and views so that empathy is built amongst policy makers, thus ensuring development interventions' relevance to the poor.

This chapter will focus on the research outputs, communications, and policy engagement processes to establish RCA's effectiveness first as a research exercise – specifically, does the research reflect the voices of desired groups, and is it valid and relevant - and secondly as a tool for influencing policy and – ideally – translating into changes in programs or policies. In Section 7.1 we look at the effectiveness of RCA as a research exercise. In Section 7.2 we ask, "in what contexts and for what purposes does RCA works best?" We conclude with looking at RCA's effectiveness as a policy influencing exercise in Section 7.3.

7.1 Effectiveness of RCA as a research exercise

Key Findings

1. RCA claims to reach poor and marginalized voices, but the sampling criteria are often poorly conceived – making it unlikely that the poorest or most vulnerable will be reached – and implementation of sampling on the ground compounds these issues. This raises questions with both the validity and relevance of studies.
2. The RCA is a simple and flexible approach that generates large amounts of detailed data; its interpretive depth is weak, however, and it's respect for voice is largely limited to "good listening".
3. Some findings are presented as being surprising and counterintuitive, but primarily to those who have limited contextual understanding and knowledge. RCA is more valued for its ability to confirm, than for surfacing new knowledge.
4. RCA is strong at communicating some research participants' own perspectives through presenting a variety of case studies and voices; however this is often very aggregative, with little sense of how case studies or indeed overall conclusions fit within the whole set of data collected, let alone how findings should be seen to relate to the wider context.

The best way to assess effectiveness of RCA as a research exercise is simply to examine the reports that are produced. However, because research does not occur in a vacuum – commissioners will have particular motivations for undertaking a particular piece of research in a

particular way, and will value it accordingly – we also contextualise our document review with findings from KIIs.

Assessing the 'quality' of a research output is inherently somewhat subjective, but we can make the review as transparent as possible by setting out some objective criteria that should serve as reasonable benchmarks for determining whether the research meets a high standard of quality. These are:

- **Inclusiveness -** Did the research effectively include the desired population groups, and were their voices reflected?
- **Validity**[49] **-** Are the findings valid, meaning (i) is there useful and meaningful description of the data; (ii) are findings presented objectively (transparently); (iii) is interpretation careful, evidence-based, and nuanced; (iv) does the analysis suggest where, when, and how the findings might be transferable to other contexts; and (v) are voices/perspectives reflected authentically?
- **Relevance -** Are the findings relevant: do they (i) answer the research questions and (ii) add to the knowledge base?

Each of these questions will be addressed in turn, using the case studies as the basis for the assessment. Within these discussions, we also highlight where the RCA claims to reflect the voices of desired groups, and present findings that are valid and relevant.

7.1.1 Inclusiveness

What RCA claims

One of the main purported advantages of RCA is being able to reach marginalised populations who normally would not have a voice at the policy table, allowing their perspectives to be heard. Masset et al write, "*By purposefully staying with the most disadvantaged, older people, children and poorer, marginalised community members, often 'unheard' in conventional evaluations and studies, are brought to the fore*" (2016: 173). RCA study reports frequently claim that the RCA approach is particularly suited for research with the most marginalised and that researchers are given clear instructions to prioritise interaction with those who best fit the target population criteria: poor, marginalised populations (see, for example, RCA, 2015a).

What we found

Sample selection

The studies tend to set out selection criteria for how the host participants will be stratified. These purport to allow the research to reach poor, vulnerable, and marginalised groups who are of interest to the commissioners. This is similar to the approach used in most qualitative work. With the RCA this becomes challenging for two reasons, however: (1) from the outset, the criteria for stratification are highly problematic - in practice there is likely to be some significant self-selection bias that makes implementation of even ideal criteria difficult; and (2) in practice the criteria often need to be abandoned by researchers in the field out of practicality or indeed even sometimes for their own safety.

We can look at some case study examples to better understand how these challenges play out in reality. The Adolescent Nutrition and Physical Activity study states that it was motivated by worrying national-level indicators on two extremes of nutritional problems amongst adolescents:

[49] See Tierney and Clemens 2011

under-nutrition on one hand, with high rates of stunting[50] and wasting, and over-nutrition on the other, with rates of obesity a concern. This should have led to a study design that included the very poor, as poverty is likely to be highly correlated with under-nutrition. As is a hallmark of the approach, described above, the study did not undertake any background contextual research in order to determine the criteria for stratification. Instead it took some criteria based on previous RCA work in the country, as shown below. It then claims to have reached 8 households of poor socio-economic status, 8 middle, and 4 high, all further stratified by location (rural/peri-urban/urban).

Examples of sample selection criteria

Adolescent Nutrition and Physical Activity Study Selection criteria

1. Dependency ratio, a result of comparing number of working members in the family with the number of family members it has to support;
2. (Ir)regularity of job and/or income (dependence on salaried versus waged or self-employed work);
3. Level of indebtedness;
4. Ownership of particularly expensive assets, e.g. cars, at screen TVs, laptops/computers;
5. Opportunities for diverse diet (often a proxy for having cash);
6. Social status in the community, e.g. village leadership, of noble descendent) which in turn infers connectedness and networks (but not necessarily economic prosperity).

Household Finances Study Selection Criteria

1. All families, at least theoretically, meet the eligibility criteria for social assistance by being families living in poverty.
2. All families have at least two generations living together in the house with at least one dependent child.

In addition, purposeful selection to include the following groups, indicate a household is particularly poor:

- Families with 'missing middle' generation (grandparents and grandchildren without parents)
- Families where an adult is unable to work
- Families where a mother is at home unable to work because of childcare
- Women-headed families

The Household Finances Study does not clearly articulate any background or context motivating the research aside from a desire to understand how households access and use cash especially given the fact that there are social assistance programs delivering transfers in cash, but there is at least some idea that it is the finances of the poor that are of interest, given the criteria that are set out (right-hand side in the call-out above).

These criteria are woefully inadequate in both studies, and reflect a lack of understanding about poverty and vulnerability in general, let alone how this might be relevant for adolescents. The dependency ratio, regularity of work, and indebtedness are unlikely to allow identification of the poor, who may have regular but very low-paid employment, and indeed the poor often carry less debt than those who are better off because they have much less access to credit. In the Household Finances study, it is not clear how they determine who would be 'theoretically eligible' for which social assistance program[51]. The remaining demographic characteristics, while they could be

[50] While alarming, stunting amongst adolescents would normally reflect under-nutrition in early childhood and actually is not particularly relevant for a study of adolescent nutrition practices.

[51] No further explanation is given in the report, but the RCA team stated in personal communication that they used criteria related to the Unified Database, which is a national-level proxy means test for several social assistance programs. This was intended to cover roughly the poorest 40% of the population, but targeting errors are extensive, as

useful when combined with other relevant criteria, are poor proxies in themselves for identifying very poor households. Instead, in both case, much better criteria should have been selected, based on a thorough understanding of livelihood strategies and poverty/vulnerability in the study areas, differentiated for rural/peri-urban/urban locations.

It also begs the question how these could have been effectively operationalized in the field, as part of the household selection process: are researchers really asking about how indebted a household is, or their social status, in a preliminary conversation while they are 'hanging out' in a public place, before asking if they can then stay with that household for several days? How do they establish whether households are 'theoretically eligible' for social assistance programs exactly? Social status is sensitive and highly nuanced and requires some understanding of a community in order to gauge; this would be very difficult just through casual observation to ascertain. And it could be awkward or not welcome for a stranger – especially an outsider – to be asking questions about debt, or even whether a household owns expensive consumer goods.

This is an issue that goes far beyond this study to the approach as a whole. It is questionable to what extent researchers can really consistently identify the poorest, most vulnerable, and marginalised households, and whether these households are in a position to host an out-of-town guest for several nights. For example, adolescents who are in school, with leisure time, were far more likely to be able to host the researchers in this study, compared to those who work outside the home (and who therefore might not feel comfortable asking their employer if they could be 'shadowed' by an outsider at their workplace) or even inside the home (including out-of-school girls with heavy household burdens, some of whom would already be married).

The report acknowledges some issues with this, "*In some locations, particularly rural, it was difficult to find adolescents hanging out outside their house to start engaging with. Most of them were inside their homes or if they were outside, they would be riding on a motorbike.*" (RCA, 2016b: 7 & 8). However, there is no discussion of how this might have biased the sample or the implications for including the poorest. Nor is there adequate reporting back on how the actual sample reflected the criteria, so no way to objectively assess what exactly 'poor' 'middle' or 'high' actually means. All that was included was 'asset' holding (namely 3 kinds of consumer goods: TV, motorbike, and mobile phone), but not social status, indebtedness, livelihood sources of parents, etc. Somewhat suspiciously, all the poor households had TVs, motorbikes (except 1) and mobile phones as did the better-off; indeed, the patterns of these goods shows no real differentiation between the 'low' 'middle' and 'high' groups at all.

We can, however, assess the sample against other external data sources to get a sense of how well it represents poor adolescents. From a recent OECD report on education (OECD & ADB, 2015: 148[52]), we know that while enrolment in lower secondary school is quite high in general, it drops off markedly in upper secondary (which is not part of the mandatory basic education), with gross enrolment rates of 76% (and net enrolment would be considerably lower). Enrolment is highly correlated with poverty status and age, as school fees, distance to schools, and the higher opportunity cost in terms of lost employment is greater the older the child. So to understand nutrition amongst poorer children, it would have been essential to reach those who are out of school (having dropped out early) not just those in school. However, it appears that all the adolescents were either in school or had graduated from at least some level of schooling. Furthermore, the sample is heavily skewed (23 of 32) towards older adolescents (15 and above)

in any PMT (see Kidd and Wylde, 2011), and in any case questions related to malnutrition should probably have been aimed at a much narrower segment of the poor than the bottom 40% to be effective.

[52] OECD. & ADB., 2015

and in the West Lombok sample, all were 16 or over, with the majority aged 18-20 years. This is problematic partly because children of this age who are still in school are much more likely to be from better-off households but also because issues related to adolescent under-nutrition are important to study in younger adolescents.

In the study, the analysis and findings are overwhelmingly skewed towards youth in school, nearly all of whom have a motorbike and significant amounts pocket money for food. (Although surprisingly, the selected children highlighted in the 'Meet some of the adolescents' are mostly 'low income' and out of school and working, although it is not clear whether they were all host children or others who were encountered).

Another indication that the sample was highly skewed away from the poorest is the amount of pocket money that the participants in the Adolescent Nutrition Study all cited as being 'necessary'; only a few got less than 5,000 per day, with the majority getting 5-15,000 per day. Similarly, a case study participant in the Household Finances Study mentions needing 50,000 for his three daughters per day. However, the official poverty line around the time of the research was 374,500 per capita per month (Hasbullah, 2017).[53] Pocket money expenditures of even 5,000 per day, assuming only 20 days per month, would therefore be over a quarter of the entire per-capita expenditure of the poor (which includes housing, utilities, as well as the basic caloric requirements), and 44% of total food expenditure. By contrast, even on average (i.e. skewed by the non-poor), in the two provinces of the RCA study total household purchases of prepared food was under 30% of food expenditure[54]. It is therefore extremely unlikely that *poor* adolescents have the kind of pocket money found amongst study participants here[55].

The sample – or at least the sample used for the bulk of the findings and all of the conclusions– is therefore almost certainly not representative of poor adolescents for whom under-nutrition would be a potential concern – one of the key motivating factors for the study.

In order to actually reach poor adolescents, the design would have needed to be much more cognisant of how poverty influences decisions about schooling, work, and marriage, how to identify out-of-school youth (whether working outside the home or not), and what kind of hosting arrangements would have been appropriate, given the context of their daily lives. This context is presumably very different for poor and vulnerable girls compared to boys, and yet the study design makes no consideration of gender issues at all.

There may be some cases where the RCA approach does allow representation of marginalised groups – for example the Nepal RAP3 study, where staying with host families allowed the inclusion of very remote communities that might otherwise have been excluded from research – but the overall approach appears to be problematic with respect to actually ensuring that the voices it purports to seek out are included. This is a clear example of how a lack of basic contextual research, and indeed a strong lack of reflexivity, can seriously compromise the quality of the study. Shah and Arvidson both offer up harsh critiques of RCA on this count:

> *…it is difficult to get away from the notion that power remains with the researcher…while the 'building of rapport' and 'intimacy' are pri marily interpreted as ethically informed approaches, they have become*

[53] Hasbullah, 2017
[54] Data from BPS-Statistics Indonesia (2016).
[55] One other qualitative study by Nurbani (2015) did find that pocket money was a significant and growing part of household expenditure, but, again, it is not at all clear that this finding necessarily represented poor households.

> *commodified and professionalized…which is initially presented as an ethically motivated approach can suddenly appear as unethical practice* (Arvidson, 2013: 284).
>
> *RCA would do well to drop the patronising rhetoric about how it "gives" people "voice" and instead engage with the unjust power dynamics that it may well inadvertently perpetuate* (Shah, 2018: 22).

Synthesis, analysis and reporting

There are also major concerns with the way in which the voices of different groups are synthesised and reported, as already discussed in Section 6.3.1 above. Rather than focus on the process as we did above, here we can examine the way in which issues related to methodological rigor manifest themselves in terms of the quality of the output.

Continuing with the example of the Adolescent Nutrition and Physical Activity Study, we have considerable concerns with the quality of the synthesis, analysis, and reporting as it pertains to reflecting the voices of poor, vulnerable, or marginalised adolescents. In this report, data, analysis/interpretation, and conclusions are presented in a few sections: a chapter on 'changing contexts' and one chapter each on physical activity and nutrition (the two main research themes); and then a short conclusions/recommendations chapter. In the chapters presenting the initial data and analysis, assertions/findings are sometimes disaggregated by location or sex, and there is at least one paragraph on children who are out of school, but the data and interpretation are jumbled and tend to be highly generalised, with little meaningful analysis across different groups of adolescents. The conclusions are drawn entirely from the findings on the adolescents who are in-school (and it would appear to be heavily biased towards peri-urban and urban contexts).

Implications

RCA therefore fails fairly comprehensively to deliver on one of its main claims about amplifying the voices that might otherwise be missed; not only are they systematically under-represented in the sample as a result of the (lack of) rigor in the design process, their voices are then systematically under-represented in the synthesis and analysis that appears in the report. This practice is not isolated to the Adolescent Nutrition and Physical Activity and Household Finances studies, but was found to be a widespread characteristic of the reports in general. We found there was a tendency to view 'the poor' as an undifferentiated group, with no nuanced understanding of class, caste, ethnicity, gender or livelihoods.

7.1.2 Validity

Assessing validity

Assessing validity in qualitative research is somewhat contested, but there are ways to usefully consider the extent to which research represents the situation and context it has set out to investigate.[56] The first way is to assess the methodological process, as research that follows a rigorous methodology is more likely to result in research that is valid. We did this in Chapter 6 in our discussion of how findings are presented in reports (6.3.1) and our discussion of transparency (6.3.2). Our discussion here focuses on the end product, in particular the ability of RCA reports to provide useful 'thick' description, authentic representation of voices/perspectives, and conclusions that are in line with the data presented.[57]

[56] See, *inter alia*, Morse, 2005; Tierney & Clemens, 2011
[57] See, for example, Tierney & Clemens, 2011

What RCA claims

The RCA claims to be a simple and flexible approach to research[58] that respects voice and achieves depth of description (Lewis et al., 2012; Masset et al., 2016). It also claims to be good at surfacing insights on sensitive topics and on changes in social norms. (Masset et al., 2016; Lewis et al., 2012). For example, "*As data is primarily gathered through informal conversation, researchers are able to discuss the many aspects of life relevant to security and justice, including community dynamics, political events, crime and safety, service provision and social norms and expectations"* (RCA, 2015a: 15).

What we found

Detailed description

RCA reports provide significant descriptive detail; however true depth, both descriptive and analytical, is largely absent. There were strong views amongst some qualitative research experts that, "*The information included in reports is not very deep; it's rather superficial. I think 2 or 3 days of fieldwork in any community can find these things out. Researchers don't really have a chance to write things out, mull things over, hypothesize, They are just supposed to be sponges, squeezed out when they get back to Jakarta*" (KII, Qualitative research expert). Two commissioners (whom we interviewed together) suggested that, "*If you want to go really deep, why not do longer ethnographic studies?... and also use different methods?…If we want depth, we use more conventional qualitative approaches that will enable us to capture difference, segmentation, variation in a bigger sample, to at least try to surface more heterogeneity. This is why we might pick another approach rather than RCA*" (KII, Commissioners). While not particularly deep, we do find that the amount and breadth of descriptive detail generated in such a short amount of time by researchers is impressive.

Authentic representation of voices/perspectives

Claims to be able to collect and analyse data on sensitive topics are not supported by the evidence. While some researchers told us that they were able to achieve a level of trust and intimacy with some host household members over the course of longitudinal RCAs, they further suggested that sharing of sensitive information was largely unrelated to the topic of the RCA: "*I think that the intimacy of living together can encourage people to share sensitive information, but this is not always related to the topic, so cannot translate to insights relevant to the study*" (KII, RCA core team member). Indeed, a majority of the RCA researchers consulted for this review strongly disagreed that this was a strength of the RCA. Those familiar with RCA and approaches to qualitative research in Indonesia suggested that "*RCA is not good at looking at sensitive issues. In some areas [of the country] where there are significant sensitivities, such as in tribal areas, we have to stay 2 or 3 months in communities to get any credible information at all. Building trust takes a long time. So how can RCA know the 'reality' in 4 days?*" (KII, Qualitative research expert). Shah (2018) also reports that, for the Papua Education RCA, the report acknowledged that discussions around sensitive issues such as bullying and teasing were difficult.

With respect to findings related to social change, the IP-SSJ RCA study in Nepal failed to present significant insights concerning social norms in relation to security and justice, and a specialised

[58] This flexibility of not unique. Roller and Lavrakas (2015) suggest that a defining characteristic of qualitative research is the flexibility built into the research design. They give the example of a focus group moderator who only understands which topical areas to pursue more than others or the specific follow-up (probing) questions to interject once they are actually in the discussion. Participant observation takes this flexibility to the extreme, in that they have little control over the activities of the observed and, indeed, the goal of the observer is to be as unobtrusive and flexible as possible in order to capture the reality of the observed events (Ibid). Reliant as it is on participant observation guided by "Areas of Conversation" RCA is indeed a simple and flexible approach.

study on norms needed to be commissioned following the RCA study. While longitudinal RCAs can pick up some changes in attitudes, beliefs, behaviours, identifying and understanding normative changes is a very specialised area of research. Lewis et al. suggest that,

> *When researchers undertake longitudinal work that aims to 'follow change', it can become difficult to separate actual change from change in the way that researchers themselves see things. For example, how far did social norms actually change during the five year period, and how far did the teams' longitudinal presence simply allow them to better see things that had been there all along* (2012: 46).

Evidencing of conclusions

The general lack of rigor with respect to the synthesis, analysis, and reporting processes has already been presented in detail in Section 6.3, and this, unsurprisingly, leads directly to weakness in the validity of the research presented in the review case studies.

How these issues with validity manifest themselves in the review case studies

Returning to the example of the Adolescent Nutrition and Physical Activity Study, the report can certainly be said to provide some degree of interesting detail, and the reader is provided with some sense of the activities and food choices made by young people. Unfortunately, however, the quality of the description is very poor from the perspective of validity; there is very little 'thick' description, instead quite a lot of aggregated or generalized data is provided, in a fairly unstructured way; there is no systematic presentation of findings against the key research questions.

Perhaps more problematic is the fact that evidence is not produced to convincingly back up the findings that are made. Instead, there is an impression of very selective use of examples, with no real analysis by group or context (e.g. location, age, livelihood or employment situation of the parents, etc). For example, although there is a very minimal amount of description of the situations of adolescents who are not in school with respect to activity and nutrition, the conclusions and recommendations derive *entirely* from the experiences of adolescents who are in school. This raises very big questions about the validity of the findings: not only is the reader left without any concrete description of, for example, the eating habits (and the contexts in which these occur) of different kinds of children beyond a few generalisations – in other words we are missing the 'thick' description that would actually be very useful – the lack of introspection and careful, nuanced discussion about for whom the results apply, when, and why creates major questions for evaluators about the very large generalisations – let alone policy conclusions - that are ultimately claimed.

This is characteristic of the other reports included in the case studies. A similar lack of attentive and thoughtful analytical approach was found by Shah (2018) in her assessment of the Papua RCA study on education.

Writing of the RCA study on education in Papua, Shah, an education specialist who lived in Papua for a number of years commented that, "*More seriously, the report does not use ethnicity as an analytical lens in its presentation of findings, despite how important cultural differences (particularly between western Indonesians and indigenous Papuans) are to understanding education in Papua. The report fails to even mention the ethnicity of teachers and parents that findings are attributed to*." (2018: 24). She continues that, "*As an experienced ethnographer of education in Papua, some of the findings in the report seem to me to be outright inaccurate. Other*

findings, whilst broadly accurate, are superficial and when presented without meaningful interpretation, encourage ethnocentric assumptions. Some of the report's findings even contradict each other" (Shah 2018: 25 & 26). She cited a number of examples of where the report was inaccurate, superficial and contradictory (Shah 2018: 26 & 27).

We also find a similar lack of analytical rigor in regard to gender and social inclusion issues in the Nepal RAP3 studies and the IP-SSJ RCA study that formed one of the case studies for this review.[59] Our detailed assessment of analytical rigor in regard gender and social inclusion issues in these two studies can be found at Annex 9

Of note is that a lack of understanding of gender and social issues in case studies has led, on a number of occasions, to ill-informed etic overlays. This can be found, for example in findings that are touted as 'unique' but are actually likely to be the result of a combination of poor evidencing of data combined with lead researcher bias towards a particular finding. For example, the Household Finances, Adolescent Nutrition, and Millennium Village evaluation all find that there is a new 'need for cash'. Were these all really examples where, until recently, households were mainly self-sufficient, with few cash transactions? If this were the case, it would be important to explore why that was happening – what recent changes had led to that? – and the time period under which the change was happening. For those familiar with rural household economies, this seems extremely unlikely; even very poor households in remote rural areas are very likely – except in rare circumstances – to be integrated with local markets – in cash – to some degree.[60]

In any case, there is almost no evidence cited to back up these findings aside from some vague quotes about "in the past …" – no timeframes given, no sense of magnitude of the change in own-produced versus marketed produce, etc. And yet this 'finding' is repeated in many places (indeed, cited across these reports) and then becomes taken as 'fact'.

7.1.3 Relevance

Assessing relevance

Issues related to relevance partially flow from the prior issues with sampling and validity; in many cases, the sole purpose of the RCA studies is to represent the voices of particular groups on particular research topics. Where it fails to do this because it does not include the relevant groups, and where those voices are not reflected with careful objectivity and authenticity, relevance is *de facto* compromised. It is further compromised, however, by the lack of contextual research – which would help to situate findings within what is already known and therefore what value the research adds to the knowledge base – as well as the lack of a clear research framework.

What we found

Taking the example of the Adolescent Nutrition study, one of the stated reasons motivating the research was the issue of under-nutrition amongst young people, and yet the research provides no insight whatsoever into this issue. Similarly, caste and gender issues were only weakly investigated (and sometimes inaccurately presented) in the IP-SSJ Study, despite this being a

[59] The Lead Reviewer completed 2.5 years of ethnographic research for her Masters and PhD degrees in Anthropology in Nepal, and has subsequently worked in Nepal for approximately 7 years between 2002 and 2018, researching and advising on gender and social inclusion issues.

[60] See for example a growing list of studies across a range of country context related to cash vs in-kind social protection, where the circumstances where local markets do not function are quite rare (Gentilini 2016 provides a good overview). For Indonesia specifically, a Food Security Assessment by WFP (2008) identified that in fact the most food insecure households were those who relied on daily cash wage labour and net sellers of crops, suggesting a high degree of market integration even amongst those who should have been the focus of the Adolescent Nutrition study.

stated aim. Even in studies where RCA research was better able to meet study aims, there are issues around surfacing relevant new knowledge.

The RCA claims to present findings that are surprising and counter-intuitive, challenging commissioner and consumer assumptions. While it is sometimes the case that research findings are counterintuitive and surprising, and could only have been gained through immersive participant observation, some findings are simply resurfacing old knowledge. A number of examples from case studies can be found at Annex 9. These examples represent significant missed opportunities; if stronger background research were conducted in order to inform research design, these 'findings' should have been the *starting point* for the research, and then much more time and effort could have been put into the *what*, *why*, *when*, and *how* questions where the studies could have actually added value.

While this speaks to RCA as a methodology in that it fails to undertake sufficient secondary research and recruit seasoned researchers and analysts who have sound contextual knowledge, it speaks even more to the erosion of institutional knowledge within commissioning agencies and implementing partners, where insights are discovered, lost, and (hopefully) rediscovered, with changes in key personnel. It also speaks to the increasing disconnect between policy makers and the people who most feel the effects of these policies. Qualitative and ethnographic research approaches can help to bridge this gap.

In interviews with commissioners, raised many more times than ability to uncover counterintuitive and surprising findings, was RCA's value in confirming things that commissioners suspected or had some, but insufficient, data on. This finding echoes Pain, Nycander and Islam's findings in their review of the Bangladesh RCA Study (2014).

> *Everyone knew that the Socio-economic Development component in RAP3 was not working. RCA confirmed it. The people's voices were like the nail in the coffin* (KII, Evaluation expert).

> *Well, maybe but I already knew a lot about the area. It was not much of a surprise for me when I read the report. The way RCA presented the findings is more important. Though we had been talking to government about these issues, hearing the RCA story really confirmed it, and helped to convince my government counterpart* (KII, Commissioner).

7.2 Where does RCA work best?

Key Findings

1. RCA studies have made valuable contributions to mixed method monitoring, evaluation and learning efforts, particularly at the beginning (scoping and baseline), but also as a retrospective lens, to better understand why particular changes may or may not have taken place.
2. Stand-alone one off exercises are less rigorous than longitudinal mixed method exercises and are not recommended, with the exception of high level "journalistic" pulse-taking exercises that can fill information gaps in regard to how policies are affecting local people, and highly exploratory landscaping exercises followed up by more rigorous research methods.

Within the overarching objective of surfacing accurate, unique and authentic insights that will best respond to opportunities for influence and to communicate these to policy makers in a way that generates understanding, and empathy, RCA has been used in a few different ways: as part of longitudinal mixed method monitoring, evaluation and learning, and as stand-alone one-off research to feed into design or provide insight into the effects of policy on people.

7.2.1 RCA as part of longitudinal mixed method approaches

RCA's intentions

There are a number of ways in which RCA has been combined with other approaches in an attempt to strengthen its effectiveness and relevance. RCA has been used alongside both quantitative and other qualitative approaches in longitudinal third party mixed method monitoring, evaluation and learning (MEL) components of programmes. As with any qualitative component of mixed method MEL, the aims of RCA studies were primarily two-fold:

1. inform the design of survey instruments and ground truth and qualify the findings of quantitative surveys; and
2. offer an interpretive lens to provide understandings of people's attitudes and behaviors, and why change did or did not happen as expected.

While RCA has been used extensively as a stand-alone qualitative research approach, the original intention of RCA was to work alongside other approaches, filling gaps. Indeed most respondents continue to argue that RCA should be seen as a complement to other sources of data, particularly in evaluation.

"RCA studies while insightful and relevant cannot provide firm conclusions regarding the effects of development interventions... it does need to be combined more systematically with other sources of data to generate more rigorous and generalizable findings" (RCA, n.d.b: p. 12).

"RCA works best when it is used alongside other studies, like a quantitative survey, or other forms of research" (KII, RCA core team member).

"I would strongly urge RCA to be combined with other qualitative approaches. I don't think that RCA replaces good qualitative research" (KII, RCA core team member).

"The approach was never set up to be on its own - a 'tell all' of what's going on in a programme. Its purpose is to provide insights about what is going on in a particular area on a particular issue…It was 1 of 5 or 6 different lenses applied on Village Law. Of course it can't tell the whole story. It's a way to document what is going on in a village" (KII, Commissioner).

Using case studies to assess whether RCA delivers in relation to longitudinal mixed method MEL

Review case studies that help to assess whether RCA aims in relation to longitudinal mixed method MEL have been met include RAP3, and IP-SSJ. We also conducted a lighter touch document review and two key informant interviews related to the Millennium Villages Evaluation in Ghana.

An RCA scoping study was undertaken prior to the baseline for RAP3. This study was reported as being useful in informing the design and implementation of the quantitative survey (see Annex 10 for a summary of RCA's contribution to baseline questionnaire development and analysis).

However a good qualitative researcher with knowledge of Nepal and the road sector could have made many of these contributions without undertaking a large scale research exercise.

Most evaluation experts suggested that RCA's usefulness as part of mixed method MEL was greatest at baseline:

> *At baseline, RCA really gave great insights about communities and how they work in a deep way. There were issues about the way the program distributed fertilizers, for example. Other qualitative work revealed tensions, but with RCA there was actual evidence, because researchers observed project staff being very aggressive, for example. The RCA also informed where the survey was weak – there was a whole commentary on this – which we couldn't change, but at least it showed where data might not be robust* (KII, Evaluation expert).

While many of the "significant insights" from baseline RCAs are already known to those with geographic and subject matter expertise or could have been revealed by other qualitative methods, some are insights that non-ethnographic qualitative research methods would not reveal: insights that require observation and are not easily evidenced through interviews or conversations alone.

This has value for money implications. In some contexts, it might suggest that a desk based review by a geographic and subject matter specialist of surveys might be sufficient. In other context, where little is known and ethnographic insights are important for filling an evidence gap, there is a stronger justification for more immersive research. For example, framing security and justice issues in ways consistent with how they are understood by communities was seen to be an important contribution of the IP-SSJ RCA study. The IP-SSJ RCA study helped to develop the baseline questionnaire by suggesting a number of additions, including examples of positive/negative behaviors exhibited by the police, indicators of safety, questions related to dowry and bride price, as well as suggesting a number of replacements of Nepali terms by local language terms that people actually use for the concepts included in the questionnaire.

RCA core team members frequently cite the importance of RCA in developing a "people's theory of change", challenging outsider project bias in terms of what results and changes – both intended and unintended - are desirable to local people. Key informants shared mixed, though largely negative, views on the RCA's value in this regard:

> *I do not really know what to say about a People's ToC – it's not part of the evaluation at all really. We did produce sector-based ToCs, and I know these are partial and can miss bits, but it was the best way to tackle this monster of a program. [The RCA team leader]…has been developing a complexity model that they also call a People's TOC, but this has not landed. It's something that came at the end and did not contribute to analysis. [The team leader]…said it came out of the RCA, and keeps saying it takes the project forward, but it does not. It's not helpful* (KII, Evaluation expert).

The story of the RCA's contribution to the IP-SSJ and RAP3 Theory of Change is similarly contested. While RCA core team members claim that the RCA critically illuminated intergenerational migratory aspirations hitherto unknown and not taken into account, others involved in the longitudinal MEL component suggest that, *"the 'people's theory of change' language is really disingenuous. I couldn't really figure out what RCA's contribution was…[The*

team leader] created this…in a vacuum so it never really landed anywhere" (KII, Evaluation expert.) This suggests that while there is the potential for RCA to help build a Theory of Change that reflects ground reality and people's aspirations much more closely, but that a key challenge is in working together with other team members to ensure that this is useful and more broadly used and owned. Regarding the Theory of Change produced for IP-SSJ, the review team finds little evidence that the RCA added anything in terms of depth, for example better explaining the quantitative results and making them more 'real' by grounding them in actual experiences of different groups of people, making sense of counter-intuitive or contradictory quantitative findings, identifying causal assumptions underpinning social norms change, showing the relationship between behaviour change and other main organising elements of the ToC.

The RCA's relative value over time

From the Millennium Villages Evaluation, we are able to gain an insight into how RCA can be effectively combined with other qualitative approaches. As well as RCA and a quantitative portion to the evaluation, evaluators also conducted Focus Group Discussions and PRA exercises in 20 villages. One evaluator reflected at length on the unique contributions and complementarity of these different qualitative approaches.

> *PRA is a much stronger comparative approach. PRA gets at community level and relational issues and comparisons. But RCA brings the detail, and the insights you can only get through observing and interacting with people over time. People revealed different things* (KII, Evaluation expert).

Non-RCA team members involved in both the Millennium Villages Evaluation and the RAP3 MEL component expressed that RCA's utility decreased over time, particularly in relation to other qualitative approaches:

> *By midline, RCA was still somewhat useful in interpretation but not so much so. But by endline, it became more difficult to reconcile different types of evidence, and things we thought were surprising and useful at baseline now seemed to be presented as more extreme than perhaps they were in reality. The whole team knew more at this point, so to us RCA findings appeared to be over-egged .... With other qualitative methods, like the larger sample PRA, the evidence seems more* balanced (KII, Evaluation expert).

Some key informants, however, suggested that the RCA increases in value over time, as well as having value as a retrospective tool. Of the RCA studies done for RAP3, one key informant suggested that,

> *RCA helps to build a story that enables better contextual understanding of how and why change did or did not happen or why unexpected changes occur. But did it make a big difference to the project? Probably not. It was interesting, but not highly relevant. RCA did help shift the focus more towards resilience and reducing vulnerability, not just lifting people out of poverty.* (KII, Evaluation expert).

While key informants are unanimous that by midline the RAP3 MEL was moving towards being mixed method from an earlier "multiple method" approach, the RCA's ability to provide explanatory depth within a mixed method approach still appears variable. In the midline report,

there are a number of places where qualitative insights provide some explanatory flavour, and where real life examples aptly illustrate quantitative findings. However, RCA insights are conspicuous by their absence in other places in the report where one would expect qualitative insights to be absolutely critical, for example in the Demography and Social Change section, where there is a large amount of sex disaggregated quantitative data, but no RCA insights, and in the Section on revisiting poverty and vulnerability in Road Building Groups, where there are no RCA insights.

A number of commentators (c.f Lewis, 2018, Chambers, pers. comm.) suggest that RCA has value as a longitudinal approach both to breakdown power differentials and establish trust, and to better understand change over time. There are a number of challenges to this claim. This first is evidence that trust takes longer to establish than is claimed by RCA practitioners (c.f. Arvidson, 2013). The second is the suggestion by key informants from the RAP3 and the Ghana Millennium Project Evaluations that RCA's value diminishes over the course of time in relation to other methods (i.e. as others involved learn more and are able to challenge what are perceived as extreme cases and unrepresentative findings). Other key informants from RAP3 suggested, however, that the RCA has "longitudinal power", for example describing important changes to livelihoods over time (e.g dietary changes), that would not have been picked up through regular monitoring. Finally, while the RCA can be useful for collecting data on observable phenomena – such as what people are eating and what shops are selling - and information that communities are conscious of, and are willing to share (Shah, 2018), there are questions and issues that just cannot be systematically and rigorously explored through a four day field visit, even if these are repeated periodically. These have to do with understanding the meaning of, and making sense of institutions, practices and relationships, particularly where these are sensitive and embedded in social norms.

The RCA's contribution to summative evaluation

Good evaluation requires judgement –the judgement of both those who are supposed to benefit from the initiative as well as those who are evaluating it. It also requires

> *…a very good ability to hold conversations and not say what you think. So RCA probably performs better as a diagnostic tool rather than evaluative one. Also, as an evaluator you need to have a good knowledge of the technical issues related to what you are trying to investigate, otherwise there is a tendency to interrupt a lot to get clarification and understanding, and possibly follow the wrong lines of conversational inquiry…* (KII, Research and evaluation expert).

RCA, on the other hand, "*purposely seeks to avoid judgement" and is "concerned about description of 'what is' not what 'should have been'*" (RCA, 2016a: 11). Further, "*RCA researchers are independent and not evaluative*" (ibid). While this somewhat draws into question RCA's use as an evaluative tool, some did see the benefits of RCA being applied in this way, For example, in the Ghana Millennium Villages Evaluation, RCA was able to highlight at endline that what mattered to the project in terms of "results" and what mattered to the people was really different: "*RCA has some utility as a retrospective explanatory tool, undertaken after a quantitative survey, revealing unintended outcomes that can help to break through programme bias. It can point out where people do and do not value things, in a way very different than that anticipated by the project*" (KII, Evaluation expert).

.

Somewhat problematic to RCA being integrated with other approaches to MEL RCA's "evangelism". Masset et al. write,

> *Ethnographically informed interpretive studies that seek to understand people's local realities and aspirations, and the context and relevance of interventions with the minimum of etic distortion, offer greater local insight than other qualitative methods in a mixed method evaluation study"* (2016: 183).

A number of commissioners and consumers suggested that the RCA core team has somewhat of a "hammer and nail" tendency, however. When we asked RCA core team members when another method would be preferable to RCA, the only answer we received was that in some contexts it would be preferable to combined RCA with other approaches. This suggests that, while the RCA core team recognises the need for other tools, there is always a need for a hammer.

Further, key informants were of the view that RCA has a "*tendency to over claim*", especially amongst a set of what were described by several people as "*RCA true believers*". This has been suggested by some to have occurred with the increasing "*commodification of the RCA*" as a distinct approach or methodology, rather than a "*bundle of ideas, a set of principles, rather than a set of tools or an approach*" (KII, Qualitative research expert), as it was first conceived of in Bangladesh. Commodification, some have suggested, has led to a somewhat evangelical group of RCA true believers, as explained by a qualitative research expert who has both commissioned and consumed RCAs: *"[Some in the RCA core team] see the RCA as all singing and all dancing…say there is no time not to use RCA. There is a real blind spot about when not to use RCA. The real skill and craft is about when to use RCA in what situation…*" (Qualitative research expert).

7.2.2 RCA as a stand-alone one-off exercise

RCA's intention

The above section has made it clear that RCA should not be seen as *"…a stand-alone coherent methodology in its own right"* (KII, Qualitative research expert) and that triangulation with other approaches is a key element of rigor. Similarly, rigor is sacrificed where studies are one-off: *"…the RCA takes place annually over five years, tracking change and people's perception of change, and is repeated each year, in the same locations, at approximately the same time and, as far as possible, with the same households. Multiple opportunities to talk and observe add greater depth."* (Lewis et al., 2012: 18). Yet stand-alone one-off exercises are a common type of RCA study.

Stand-alone one-off studies are carried out for two different reasons, to feed into design – which mixed method MEL RCAs also do - and to "pulse take" to inform broader policy. In terms of the case studies for this review, the Haze and Adolescent Nutrition and Physical Activity studies were both stand-alone one-off RCA's to feed into program design. The Household Financial Management and Village Law RCAs were pulse-taking studies, to help policy-makers better understand the impact of policies and suggest interventions to support for effective roll out. DFAT's investment in RCA has primarily been focused on stand-alone one-off studies.

The value of the RCA as an exploratory, gap-filling tool

A number of key informants suggested that RCA studies are most instructive when the purpose is exploratory, and the questions are broad. Where there is a big gap in knowledge around a certain issue – "*when you don't know what you don't know*" as one commissioner put it – RCA can be a useful source of information. A range of commissioners and consumers concurred on this.

"RCA studies while insightful and relevant cannot provide firm conclusions regarding the effects of development interventions... it does need to be combined more systematically with other sources of data to generate more rigorous and generalizable findings" (RCA, n.d.b: p. 12).

"RCA works best when it is used alongside other studies, like a quantitative survey, or other forms of research" (KII, RCA core team member).

"I would strongly urge RCA to be combined with other qualitative approaches. I don't think that RCA replaces good qualitative research" (KII, RCA core team member).

"The approach was never set up to be on its own - a 'tell all' of what's going on in a programme. Its purpose is to provide insights about what is going on in a particular area on a particular issue…It was 1 of 5 or 6 different lenses applied on Village Law. Of course it can't tell the whole story. It's a way to document what is going on in a village" (KII, Commissioner).

RCA's exploratory nature is not fully understood or received well by all commissioners, however. And, interestingly, RCA's lack of ability to get at granular and nuanced issues of importance to commissioners was specifically called out as a weakness:

> *What I realised, in coming back to the study, at the end of it, was the limited number of households they visited, and the limited number of days. I realised that, the RCA by itself, is not enough to get at the information we need. It's good for exploratory research and understanding the main issues, but not very specific and or good at getting at nuanced questions. They explained to us, the RCA team, that they cannot direct the conversation with participants. At the time I didn't really understand what they can and cannot do. At that time I was trying to push them to get more granular in terms of what we needed to know* (KII, Commissioner).

> *With many of the studies we commission, we have specific questions we need answering. Of course we get more information than we are looking for, but we start with some specific questions in mind. This does not fit so well with RCA. They leave the household to lead the conversation. As a result they come back with very wide information, but not always addressing the questions that we want. So this goes to when to use the RCA. We chose carefully between the methods* (KII, Commissioner).

While it is certainly true that the findings the RCA generates are appreciated when they fill a void, it is essential that the void be filled by high quality data. When RCA is used as part of a mixed-method approach, we heard that there are frequently challenges to RCA findings from other sources of data and a need for reconciliation. This suggests that RCA's own findings need to be triangulated by other methods in order to increase their reliability and validity, and calls into question whether RCA should ever be used as a stand-alone approach. A qualitative researcher with a long association with RCA concurred, "*I'm less convinced about stand-alone uses. Longitudinal RCA's challenge very formalistic evaluations and research and the very short-termist way in which people think about development projects. It's about a longer commitment. If RCA is going to be one-off, you might as well do FGDs, SSIs, participatory discussions"* (KII, Qualitative research expert).

7.3 Policy influencing

Key Findings

1. RCA is valued for providing insights into service provision from a service provider's perspective, something seen to be unique and that is not surfaced by normal monitoring.
2. RCA is seen by some commissioners to be a "blunt instrument" providing insufficiently detailed and disaggregated analysis of ground realities and actionable insights.
3. RCA reports have produced some insights that have been important for program development, implementation, monitoring and evaluation, as well as for policy. Case studies where this is more apparent are where there is a strong institutional home for RCA insights

7.3.1 Does RCA "punch above its weight' in terms of policy influence?

RCA's claims

Perhaps RCA's biggest and most ambitious claim is that RCA, "*'punches above its weight' in terms of policy influence*" (RCA, 2016d: 27). RCA claims to be unique tool to better understand and communicate the complexity, interrelationships, and dynamic context to policy makers, and thus to *ensure* that development interventions are relevant (RCA, 2017a).

Ascertaining the veracity of this claim is challenging. Firstly, policy influencing is a long and convoluted process; the policy landscape is complex and determining attribution, or even contribution is difficult. We have not attempted to assess the impact that RCA studies have had on policy, though we do cite some examples of where study insights proved helpful to policy makers.

Secondly, commissioners are not always the most reliable informants: they can suffer from positive bias towards research that they themselves have commissioned and invested time in ("of course it was valuable and good value for money!"), and do not always have the skills or experience necessary to judge whether research findings are trustworthy . Though the novelty, colour and detail of RCA reports are appreciated by commissioners and help to stimulate consumption, for desired policy impact to be achieved findings need to be relevant, valid and reliable.

Policy-makers' views

We heard from a number of commissioners that RCA reports can provide them with information that is not surfaced by normal monitoring and can lead to greater understanding of ground realities. While RCA is not always able to achieve depth, commissioners certainly appreciate the breadth of detail provided in RCA reports.

An area that was raised by RCA commissioners and consumers as being one where RCA adds value concerns understanding the challenges of the supply side of service delivery from a provider's perspective. While not often raised by proponents of RCA as a unique and important contribution of the approach, indeed it was this aspect of RCA that most excited commissioners in Indonesia in 2009:

> *Various kinds of PRA and RRA methodologies have been around a long time, What was very attractive about RCA was that they didn't just…look at villagers. They looked at how service delivery facilities like schools and clinics actually worked, including how people working inside of them viewed their users. That was much rarer, especially at that time* (KII, Commissioner).

While more conventional qualitative approaches such as key informant interviews have been used to explore service providers' experiences and views, RCA's immersive approach to understanding service delivery from the perspective of providers is seen to be relatively unique. A good case study example of this is the Village Law RCA study. It appears that this would be a difficult area to research well in other ways, and the RCA worked well due to its naturally exploratory nature and lack of a formal structured approach.

Village Law RCA Study, Indonesia

I would use this approach again to look at front-line service delivery staff. I think that RCA is good for challenging assumptions around service provider motivation and performance barriers and incentives. My main concern is that it must have a very specific purpose and be part of a broader policy dialogue where there are other sources of information being used (KII Commissioner).

It has been effective at getting at insider "service provider" perspectives, those of village government officials. There were a lot of things about how the new Village Law was landing out there that we did not know before the RCA (KII, Consumer).

Other commissioners – primarily those with a strong sectoral or research background - are more sceptical of the RCA's ability to provide high quality analysis of ground realities that is useful for policy makers:

> *There is no analysis, There is a lot of detail, possibly some of it accurate, but there are no rich explanations. These are totally missing. How can you influence policy if you don't explain things to people. It's like showing them people Chinese characters. They are beautiful and elegant, but they mean nothing if you can't understand them!* (KII, Commissioner).

> *Sure, the RCA reports present case studies and people's voices, but this is a blunt instrument. I want to see some comparative analysis, so that I understand different viewpoints and perspectives* (KII, Commissioner).

> *There is no analysis in RCA reports – it's completely missing…There is also the issue of a complete lack of political economy analysis in the approach which is troubling* (KII, Commissioner).

There is an inherent tension between commissioners and consumers who want the RCA team to do the interpretive "so what" heavy lifting around findings, and the desire of the RCA core team to change the policy making discourse:

> *We want findings to be a teaser for wider discussion, not giving them [policy makers] an easy answer. Reality is complex, so we don't want to make it simple and give simple recommendations. Indonesia is complex and we need to make sure that people don't over generalise and identify one solution to a complex and area specific problem. We're really trying to change the discourse* (KII, RCA core team member).

This strikes us as a strange perspective. Early reviewers of the RCA were clear to point out that policy making is driven by simple narratives that at the best of times struggle to handle complexity (Pain, Nycander and Islam, 2014). "Making it simple", without losing granularity, is surely the job of qualitative researchers: "*Human being are story tellers - and personal stories are important - but decision makers also need analysis that tells them priorities, scale, choices, opportunity costs*" (KII, Qualitative research expert). Government of Indonesia policy makers have asked the RCA team to provide more concrete recommendations on a number of occasions, and the RCA core team has provided these (discretely) at least once. We suggest that responsible research needs to engage in the difficult project of presenting complex reality in a way that does not lead to one-size-fits-all solutions. Successfully policy influencing requires not only the generation of evidence, but the provision of advice.

7.3.2 Conditions for actioning RCA generated insights

Where there are strong commissioners and an institutional "home"

Despite the difficulty that some commissioners and consumers have in actioning insights, RCA reports have produced insights that have been important for program development, implementation, monitoring and evaluation, as well as for policy. Case studies where this is more apparent are where there is a strong commissioning team, for example the KOMPAK commissioned Village Law Study.

Village Law Study, Indonesia: 3 key actionable insights

- **Technical Content of the Training:** The study revealed that village officials had been confused with some of the technical concepts and had limited detailed understanding of Village Law and Village Government. This was shared with the MoHA and KOMPAK in Jan 16 and by March 16 the Master Training modules for village officials were revised. The training modules were revised to not be so focused on technical terms, compliance and prescriptive regulations.

- **On training and support:** Several Village Officials felt on-site mentoring is more helpful than large scale training provided in hotels. Village Officials were disappointed not to be able to contact trainers for clarifications after the formal training. These insights were used by KOMAPK to highlight the importance on the use of various ICT and tools that will further support village officials in terms of self-development and capacity assessment outside the classroom, so that they can learn in their own time. In order to address the need for capacity support to District and Village officials, MoHA and KOMPAK crafted a provincial model of Teaching Assistant support and referred to the RCA study report during its development. This model will be used to support provinces in delivering training for village and district officials.

- **On Knowledge and Understanding of Village Law**: Village Officials told the RCA research team that they need specific clarifications on a wider range of issues, including (inter alia): pooling of Village Funds; fund allocation formulas; exactly what Village Funds can and cannot be used for; what district and national funds are intended to cover; the official process for accessing Village Fund tranches; what means for complaints and grievance mechanisms exist for district and subdistrict services provided to them; the role of the Village Facilitator and what they should expect from them in

terms of support and frequency of visits; the future status of Village Secretaries as village employees or civil servants. Government of Indonesia and KOMPAK have acted on many of these issues to clarify the details, procedures and legislations.

While these are not doubt important insights, it appears that they come almost exclusively from conversations with Village Officials, and could therefore have been gathered using other qualitative methodologies.

There is also evidence that, where RCA is embedded within longitudinal MEL processes, it suffers less from a central challenge that has plagued it from its inception: how to successfully inform and influence policy makers - who are used to more formal types of data - using this type of information. "*A more structured research is seen as a more robust approach, and it is easier to convince people because you have a framework, can compare, can quantify the qualitative work. If trying to influence policy, you maybe need tool, or maybe better, ensure that it is better married to other approaches"* (KII, Commissioner). It is well recognised within the RCA that packaging people's views and experiences as generated by the RCA together with data generated through other methods makes it much more palatable and usable for policy makers.

Where the RCA is used as part of context specific engagement, it can at least partially overcome inherent challenges related to policy makers' distrust of "small scale" qualitative approaches. For example, RCA studies commissioned by programs that are drawing on a number of sources of information to inform their policy dialogues with government, such as the Village Law Study commissioned by KOMPAK and the Household Financial Management Study commissioned by KSI, perceive RCA has having value as "*a different piece in the jigsaw puzzle, and one that is received well by government because they feel it gives them a glimpse into how people are experiencing policy on the ground*" (KII, Commissioner). In some contexts, it was felt that the push-back on RCA-generated data from government would be so strong that RCA data was not explicitly presented: "*The RAP baseline was presented to government, but we did not present RCA to the government. We just knew what the response would be. So we used it to back things up and give case studies*" (KII, Commissioner)[61].

Where policy makers are engaged throughout

In addition, programmes with longitudinal mixed method MEL, like RAP3, have often institutionalised mechanisms, such as sector groups and steering committees, that can link insights contained in the reports with wider policy processes on a regular basis. The RCA has tried to do this from the first set of studies in Bangladesh, with the aim of creating a 'feedback loop' between the RCA and the management of the wider sector reform process (Pain, Nycander and Islam, 2014). The RCA+ project further tried to embed this through establishing Reference Groups for each study, with varying success.

Establishing mechanisms to feed RCA findings into policy dialogue appears to be more challenging with studies not embedded within larger programs. RCA has made significant efforts in this area through the establishment of Reference Groups. Reference groups are formed in order to create formal spaces for the research design, findings and implications to be discussed collectively.

> *Typically the reference groups consist of government counterpart agency (who host the meetings), wider government ministries, donors and donor*

[61] It should be noted that the RCA core team refute this, and report that a 30 minutes presentation on RCA findings specifically formed part of the baseline presentation to Government.

> *programmes and other relevant research institutions. The reference groups convene at two key stages in the process: Firstly, during the design, to have inputs into shaping the design, to share recent evaluation data and other research studies, to help identify synergies with other research being conducted and to raise insights and concerns about ambiguities and conundrums obtained from other research and data sources to be explored further. Secondly, in the analysis and dissemination of the findings, providing feedback on areas to explore further in the findings, reviewing the draft report to ensure it addresses the needs of the policy makers and users of the research, and suggesting appropriate platforms to disseminate the findings* (RCA, 2017a: 46).

Haze Reference Group

The Reference group for the Haze study included experts from UNICEF, Ministry of Health, WHO, BRG, UNICEF's local development partners, and other international non-governmental organizations (INGOs) working in Indonesia to promote and support children's rights. The RCA team consulted with them at key moments in study design to provide background and advice on the study design, highlight their areas of interest in the study, provide input in the areas of conversation and suggest parameters on which to base selection of study locations. Preliminary findings were also shared with the Reference Group to ensure that they had access to field information soon after collection, and to elicit feedback on findings prior to the report.

Success has been mixed. The RCA core team suggests that formal reference groups have proved to be effective ways to create ownership for the findings and ensure the knowledge produced meets the needs and is used by the policy makers (RCA, 2017a). Those who have longitudinal knowledge of the RCA clearly see the importance of stronger steerage for policy uptake:

> *The RCA now brings senior policy makers into the process from the very beginning to find out what their assumptions, interests and information gaps are. This has given RCA more structure and more focus. This is vastly different from the original approach in my view and a great improvement. And the more recent reports are a lot more focused, targeted and relevant based on feedback from…[a range of different stakeholders]* (KII, Commissioner).

However this appears to be amongst a very small group of progressive policy makers who are frequently brought into steering committees and reference groups set up as part of RCA studies. "*By and large, GoI is frustrated with RCA as a source of policy relevant research*" (KII, Commissioner).

Our fieldwork in Indonesia suggests that it is difficult to get the right people on these reference groups and engaging sufficiently with the RCA study. Reference Group members are not always subject matter specialists, and nor are they generally knowledgeable concerning context, so this can limit their ability to guide design or challenge study design and findings. Further, as noted above, Reference Group members from government are empathetic, but are not always part of decision-making structures. One RCA core team member commented that, "*Our engagement with Reference Groups is very opportunistic and really dependent on the commissioner of study and the effort that they are willing to put in*" (KII, RCA core team member). Indeed, the importance of an engaged and informed commissioning "team" was seen by early commissioners in Indonesia to be critical to RCA's success.

> *… between the RCA researchers – who were bright and really savvy - and policy makers – who were hungry to better understand why things weren't working - we had a very large team of mostly Indonesian operations people who were interested in this sort of problem solving and also had good access. These conditions really helped us to crack through the wall of intellectual resistance to the idea that textured interpretations of why services were being delivered or not mattered for improving planning as much as the quantitative questions did, even if the methodologies would be different* (Pers. comm., Commissioner and consumer).

Where there is a significant gap between policy makers and poor people's lived realities

In addition to RCA being part of a context specific engagement, policy influencing appears to have been aided by the fact that these studies have as their purpose not to prove or disprove specific assumptions about cause-effect relations, incentives that drive change and what constitutes improved performance, but rather to illuminate how policy decisions are playing out in the lives of people on the ground. The disconnect between policy makers in Jakarta and rural Indonesians is a significant problem raised by numerous key informants, and one that commissioners feel RCA helps to address:

> *RCA plays an important role was providing a reality check to senior officials who rarely to get to the field. Eschelon 2 and above…are only allowed 3 days to go out of office, so with travel, that leaves only about 1.5 days in the field every year. The furthest many are able to go is the district office. RCA studies really help me to understand the context on the ground, and how national policy is being implemented* (KII, Consumer).

But is it information, or empathy, or both that are important here? Commissioners in Indonesia had a range of views on the role of RCA in raising empathy, and the importance of empathy in policy making. A number of commissioners suggested that empathy is important in the early stages of policy dialogue, to get the issue on the table, so to speak, but less important further into the process. Some commissioners also suggested that building empathy was easier, but less impactful, with some policy makers than others: "*In Bappenas they are used to speaking the language of donors, and already have some empathy. They like the RCA for this reason. But they are not part of the decision making structures. For decision-makers, the empathy is not strong, and RCA is difficult to get on the table. It takes a great deal of skill to engage with them*" (KII, Commissioner). Almost all commissioners were of the view either that RCA is unique in the "*humanness and realness*" of reports, "*like a field trip being brought to you*", or that the RCA is stronger than other approaches in building empathy amongst consumers.

It's clear that the ability to translate RCA findings into policy-relevant insights is not a simple technical linear process, but depends all three actors in the system – policy makers, commissioners, and researchers (Lewis et al., 2012). Where RCA has been able to "punch above its weight" in relation to policy influence (and these instances have been relatively rare) it has been due less to RCA as a method that is able to generate knowledge and translate this into actionable evidence, only very partially to connections that RCA has been able to make with policy actors, and more to the spaces into which RCA has been invited, and the opportunity that these have afforded.

8 Is the RCA more effective, efficient and economical than comparators?

In the previous chapters we assessed the RCA against its own stated aims and Theory of Change. The question for this chapter is, given this assessment, whether and to what extent RCA delivers value for money (VfM). However, such an assessment of VfM cannot be undertaken in a vacuum; it explicitly requires an understanding of what the alternative – or counterfactual – would be. While this counterfactual could of course simply be a 'do nothing' scenario – a choice between undertaking an RCA exercise or none at all – in reality RCA is one of many different qualitative approaches that could be used. A more useful understanding of alternatives therefore requires an assessment of RCA in the context of comparators.

Section 8.1 presents these comparators to provide readers with an overview of their approach, as well as an assessment of their methodological rigor and their effectiveness. This understanding of the comparators is then used in the assessment of VfM in Section 8.2.

8.1 RCA in the context of comparative approaches

Although proponents of RCA tend to see it as a wholly unique solution to issues with 'standard' research methods, there are in fact a range of approaches that are comparable with respect to purpose/objectives and methods, even if there is not complete overlap. We begin with a brief overview of each (PEER, Human-Centred Design, Sensemaker, and a 'general' qualitative approach), followed by a detailed comparison of objectives (the 'why'), and methods ('the how') before moving on to an assessment of the relative effectiveness of each.

8.1.1 Overview of comparators

PEER

The Participatory Ethnographic Evaluation and Research (PEER) approach is based on the ethnographic method, but uses members of a community to generate data, as 'peer' researchers. It is more rapid and 'light touch' than traditional ethnographic approaches, which require long field research phases to allow the (outside) researcher to gain trust and understanding. The stated strength of the method is in its ability to tap into the existing established relationships with the individuals whom peers interview. Importantly, the method uses 'third person' interviewing techniques, whereby respondents are asked about people like themselves in their communities, rather than about their own personal experiences and opinions. In doing so, it can yield rich narrative data to help understand health and risk perceptions and behaviours from an insider's point of view, generate detailed understanding of the context in which these behaviours occur, and provide a more intimate engagement with the realities of young women's lives (Price and Hawkins, 2002)

Human-Centred Design (HCD) 'Inspiration' phase

HCD has been used for many years to design products and services in the commercial sector, but more recently has been used in the social sector by organizations such as IDEO.org. to design solutions to "sticky problems", such as adolescent sexual and reproductive health, exclusion of people with disabilities, and urban poverty. The stated goal of HCD is to produce desirable, feasible, viable solutions that achieve impact at scale. HCD takes a suite of diverse methods and tools and with the aim of gaining a deep understanding of the needs and desires of those who are being designed for, the context in which they live and the relationships that matter to them (ITAD 2017: 13).

The entire HCD process involves the following stages:

- Inspiration: desk-based and qualitative "design research" to understand the desires of target users and the context in which they operate
- Ideation: insights from the inspiration phase are translated into a set of design options, which are refined until a few rough prototypes are developed. These are then rapidly developed and field tested.
- Iteration: using learning from field testing to continue to refine the design.

As such, it is only the initial inspiration phase that is applicable as a comparator.

While the methods and tools used in design research do not differ significantly from those found in a participatory qualitative research toolkit, a human centered designer executes these in a particular manner, using a guiding philosophy best understood as 'design mindsets'. The core mindsets are' empathy, optimism, iteration, creative confidence, making, embracing ambiguity and learning from failure. There is a belief that designing with these mindsets will create solutions that are durable because they are desirable.

Sensemaker

While 'sensemaking' more broadly is often referred to as an approach to synthesizing and analyzing qualitative data (indeed many approaches include a 'sensemaking' stage), here we are referring to the application of an approach developed with accompanying software by Cognitive Design, which draws on complex adaptive systems thinking as well as cognitive science and anthropology.

The crux of the approach is the idea that participants provide short stories related to the topic of interest – 'micronarratives' – and in so doing they reveal a range of diverse perspectives. These are then interpreted by the participants themselves using a set of pre-defined questions (or the 'signification framework'). The software then filters and analyses these micro-narratives and signifiers to identify patterns and trends that may be of interest, allowing an understanding to emerge from a large amount of different experiences that might not be possible using other methods. Visual patterns and a set of individual stories are then used in an iterative sense, to understand what is significant.

Qualitative research in general

Beyond these particular approaches, there is a long tradition of qualitative research more generally, which can use a range of techniques and methods, from structured and semi-structured interviews to focus group discussions to any number of participatory exercises or, indeed, a combination of all these. This can be flexibly designed to respond to whatever scale and scope is required, depending on the research requirements. While this could potentially encompass a wide range of studies, here we refer specifically to a sub-set that meets a high quality standard and employs a methodologically rigorous approach.

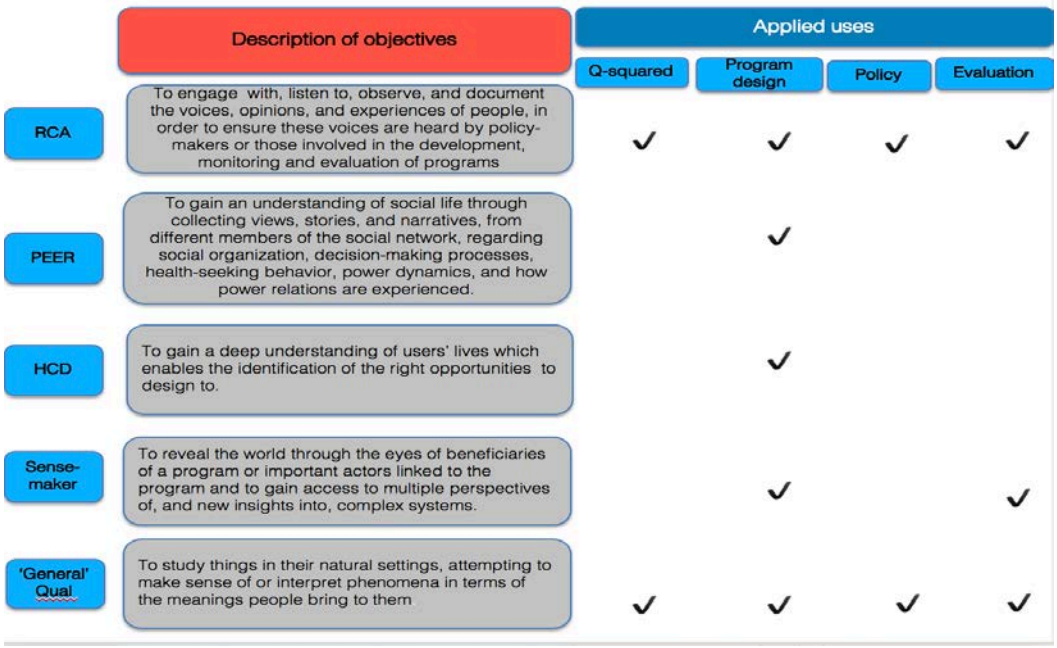
8.1.2 Comparing objectives: the 'why'

The language used to describe the comparators is often chosen to make them sound entirely unique (as many of them vie for attention amongst commissioners), but in fact there is quite a lot of similarity in the ultimate objectives, which can be summarized as providing insights from the perspective of users/beneficiaries amidst complex and nuanced contexts in order to improve programming and service delivery. These improvements can be targeted at various phases of programming, from design to implementation to monitoring and evaluation, but the overall focus

is on ensuring greater relevance and effectiveness of development activities. Across these approaches, there is an explicit aim of overcoming gaps in information that would be left by using either quantitative data alone or qualitative data that insufficiently captures the perspective of beneficiaries themselves.

Figure 6 illustrates the similarity in objectives, and the areas of overlap in terms of applied uses. PEER and HCD tend to be used more narrowly in practice, with an emphasis on identifying policy and programming needs, while Sensemaker does this as well as being used for evaluation. 'General' qual approaches and RCA both are applied more broadly, including in 'Q-squared' endeavors intended to inform quantitative survey instruments (and ideally be used more iteratively with quantitative data in interpretation as well, for example in a 'qual-quant-qual sandwich').

Figure 6 – Comparison of stated objectives and applied uses



While they may differ in the 'how', the key point here is that they are all similar enough in terms of overall objectives and the underlying theory of change to provide reasonable 'counterfactuals' for assessing value for money below.

8.1.3 Comparing methods: the 'how'

It is in methodology that there is more divergence across the comparators. Annex 11 (Table 1) provides an overview of how each comparator approaches key stages in the methodology. While the process reflected in these stages of the methodology itself give some immediate sense of the rigor of the approach, we can examine this in more detail here by assessing each comparator against the same criteria used for RCA in Section 6 above. (Further description against each criteria is also provided in Annex 11, Tables 2-5).

Rigor in the research preparation and design stage

Good general qualitative work involves the most rigorous process during research preparation and design of the comparators, with substantial background research informing context and understanding, and feeding into both a framework to guide the inquiry as well as the design of specific tools/instruments used. It is implemented using experienced researchers, or where experienced researchers are not available training is put in place. Acknowledgement of power dynamics and the position of the researcher vis-à-vis the participant are explicitly made, and indeed incorporated into the design of research techniques and instruments used. Teams are not always extensively multi-disciplinary in terms of composition, however.

PEER is probably the next most rigorous approach, differing from general qualitative work in the sense that researchers are by definition not experienced. However, lead researchers are normally highly experienced in anthropology. Teams are not normally multi-disciplinary; however, all the other aspects of methodological rigor are met.

Sensemaker is somewhat interesting to assess in terms of rigor, with the use of background research to develop the framework for inquiry, so it scores highly in these aspects. Field researchers are not necessarily highly experienced in either qualitative work generally or the Sensemaker approach specifically, but there needs to be extensive capacity in the approach and software within the team. In practice, the capacity needs are often under-estimated, and therefore there can be a significant reliance on external consultants. Team composition-is not particularly multi-disciplinary, nor is the approach especially reflective about issues related to power and relation, although in practice these things should not matter, since participants themselves are the ones who tell their own story and signify what it means. In practice, however, the position of researchers is always important to understand the context in which stories are told, and this is not generally adequately addressed.

In theory, HCD rigor should be high, emulating the practices leading to rigor in general qualitative approaches, since it essentially can apply any qualitative technique that would be appropriate. In practice however, it tends not to be implemented in a highly rigorous manner, often with only cursory background research undertaken ahead of time. Analytical frameworks and instruments are highly tailored to the specific design question at hand – as is understandable given the objective – but are not necessarily as rigorously designed as they could be. Fieldwork teams are large and span a range of areas of expertise (designers, program staff, etc) but the methodology does not explicitly address multi-disciplinarity (although the examples in the case studies do include different disciplines). There is generally little expertise or training in issues related to understanding position, as the HCD research is often conducted by very young non-locals who have a limited background in developing country contexts and formal research approaches and can bring significant (unrecognised) bias to research. The international team members also tend not to understand the issues related to position, and therefore miss critical contextual understanding and nuance.

Rigor in fieldwork

Again, when it comes to fieldwork, general qualitative approaches are rated highly in terms of rigor against most of the criteria. Careful and experienced researchers will be careful to be as unobtrusive as possible, and – crucially – to be reflexive about the extent to which this is possible and the potential impacts on the research; there is often extensive triangulation across methods, instruments, researchers, and in longitudinal studies over time; and detailed records of data are kept. One aspect where it tends to be less robust in practice is perhaps in terms of respondent validation, as ‘human sense-making’ exercises, where findings are presented back to participants is often not done due to time and resource constraints.

PEER is also rated highly across most criteria: by nature it focuses on being unobtrusive through the use of peer researchers and asking questions in the third person; responses are triangulated across peer researchers and contexts; peer researchers are actively involved in the interpretation of findings. Where PEER is somewhat less rigorous is respect to recording of findings: although there are no notes during interviews by design (to maintain trust), they should be 'downloaded' to the lead researcher in detail within 48 hours of an interview but in practice the lag might be longer, which could compromise accuracy of recording. This is similar conceptually to the issues with rigor in recording accuracy of RCA articulated in Chapter 6 above, but is less severe as a result of the rigor in the development of the research framework and instruments, so while a lag might lead to recall bias, the recording will still be much more detailed, against the pre-defined set of research questions.

Sensemaker is, again, somewhere in between when it comes to rigor; the unique and very specific approach should in some sense insulate it from issues of rigor in the fieldwork: triangulation is implicit across individuals since all are asked the same question; data is captured in the software; and validation by participants is inherent to the approach (as they interpret their own stories) and indeed in practice a further round of 'human sensemaking' adds significant value. However, there is almost no triangulation across methods, and there does not seem to be much acknowledgement of the way in which obtrusiveness of the process might influence the stories people tell or how the interpret it.

HCD makes attempts if not to be unobtrusive *per se* but to actively understand the perspective of participants in their daily lives. In theory the use of local researchers, especially youth, should help with obtrusiveness but in practice they have not been well integrated into the team and issues of trust and power are not addressed in a reflexive manner.

Rigor in synthesis, analysis, and reporting

Rigor in synthesis, analysis, and reporting follows the same general pattern as the previous two stages of the research process: general qualitative and PEER follow rigorous iterative processes to analyse data, develop hypotheses, and assess the representativeness of cases, the latter of which is aided by the strong contextual foundation and research frameworks developed during the preparatory stages. They are both also transparent in their description of the processes used.

Sensemaker is, again, rigorous in some ways, as an integral feature of the approach embedded in the software: analysis is inherently iterative, through the use of overall views and 'deep dives' into individual stories; the representativeness of cases reflects the sampling approach; and transparency is an inherent aspect of the process, with data stored directly in the software. However, the particularities of the approach mean that the extent to which meaningful iteration can yield deep insights depends entirely on the strength/quality of the initial prompting question and signifier questions used to interpret the stories. If these are not well articulated, then iteration itself is unlikely to yield meaningful findings. Similarly, if sampling is not done initially in a rigorous manner, the analysis will be unlikely to assess the representativeness of cases in the dataset.

HCD scores fairly low in terms of rigor across the board here. In theory it should score highly in terms of the iterative nature of analysis, as there is a premium on ensuring that the research team understands the perspectives of users, but in practice this can be done in a geographically disaggregated manner, limiting synthesis across sites, and can be too reductive to allow thorough interrogation of the findings. While this can be a strength, given the commitment to continual iteration - "test, fail, learn, iterate, test, limited success, learn, iterate, test, more success, learn,

iterate" - which perhaps makes up for initial depth through on-going iteration, the lack of depth in synthesis could mean key insights are missed, or that they are more costly to identify because it takes more iterations for the same things to surface. Beyond syntheses, there is little in the way of hypothesis building and testing, and researchers have often lacked the skills to do this. Transparency is generally low, with little emphasis placed on this in the approach overall, and often very little transparency about why a particular design is preferred by the IDEO.org team over another.

8.1.3 Comparing outcomes: the 'what'

Chapters 6 and 7 outlined our findings on RCA's outcomes (research quality) and impacts (policy influence). We now ask, "How do the comparators stack up in this regard?" Unsurprisingly given our findings on methodological rigor above, PEER and general qualitative approaches score highly in terms of representation, validity, and relevance, but HCD and Sensemaker also were found to be valuable by commissioners, with insightful and impactful research reports, although with more variability in quality.

PEER

PEER, as would be expected, has been found to score highly with respect to representation, being good at getting at voices that are generally not heard, and also at surfacing insights on unintended consequences Similarly, one of the KIIs interviewed said that: "*It might be odd that you as a researcher are asking questions that the interviewee knows you already know about. It might give less detailed response because they "know you know*"". However, what it may lack in drawing out things that are obvious to the researcher, PEER does do well at reaching very vulnerable voices. The KII went on to say that, "*But this group is really exploited and have a whole diversity of experiences, and so this approach gives them space and permission to share their stories with someone who knows the world and is non-judgemental*."

An example of the depth of insights is shown in the call-out, below.

HCD

HCD has been found to be effective in some circumstances in surfacing the perspectives of users whose voices may not otherwise be heard by programs.

> "*Some HCD processes and tools are especially effective in helping to understand and represent the unique contexts and circumstances that shape user behaviours. For example … by mapping a woman's 'journey' from meeting her first partner through her various reproductive milestones, it is possible to learn a lot about a woman's relationships and her ability within them to communicate and/or negotiate choices about family size, contraceptive use and future aspirations for herself and her family*" (Tolley, 2017: 20).

However, it has generally been found to be insufficient to generate deep insights around norms, power, gender issues etc.

> "*Traditional socio-behavioural research (SBR) may also be more appropriate when seeking to identify determinants of behavior change and/or evaluate program implementation and effectiveness. First, HCD should not replace the in-depth qualitative SBR that generally aims to identify, describe and compare end-users' individual (i.e. attitudes, motivations and behaviours), partner relations and larger clinic, community and/or cultural contexts related to whatever outcome (e.g. pregnancy, HIV or other*

disease acquisition) a product is being designed for. HCD research is not designed to provide rigorous evidence on these kinds of question" (Tolley, 2017: 21).

With respect to outcomes and impacts, using the Marie Stopes and IDEO.org example (see call-out, below),HCD was successful in terms of the key outcomes of relevant, impactful research reflecting the voices of girls, as well as an impact on programme delivery and therefore the lives of those adolescent beneficiaries. However, although HCD sees the inspiration phase as an opportunity to 'surface deep insights and enabling a deep understanding of a girl's world' in practice it did not provide insights valuable for the wider body of knowledge on adolescent sexual and reproductive health, instead focusing more narrowly on the design task at hand. The evaluation found that "*evidence suggests that it is unrealistic to expect the inspiration process on its own – the field portion of which generally lasts less than two weeks – to be able to generate new, deep and generalizable insights a complex, social challenge, e.g. for an entire population segment.*" Bearing this experience in mind, in later HCD work planned for the Sahel a formative research component was added (ITAD 2017: 23).

Sensemaker

Applications include evaluation of the integration of marginalized farmers into value chains; girls' empowerment; water, sanitation and hygiene interventions; empowerment of people with disabilities; and Internally Displaced People (IDP) reintegration and peace-building. It has been used as a complement to quantitative research (for example used with DHS for Girl Hub (Walker et al., 2014)) as well as stand-alone (such as in VECO's[62] evaluation of its programmes on agricultural value chains (Deprez et al 2012) and IRC's evaluation of its WASH programming (Casella et al 2014)).

Outcomes have varied widely depending on the application, with some case studies showing highly successful results – new and valuable insights – and others producing findings of low value. Very much depends on implementation and the quality of the design of the instrument and facilitation and appears to be best combined with quantitative survey.

General Qual

General qual exercises of the sort that meet the definition here – that are by definition high-quality – are, unsurprisingly, generally quite good at both triangulating with existing findings from other studies and surfacing new findings.

Taking an example from a study on adolescent girls' empowerment (in keeping with several of the other studies above on adolescents and/or reproductive health), a recent research exercise into adolescent girls' economic empowerment in Rwanda (Calder and Huda, 2014) made a number of findings, including those on the challenging topic of social norms, and, from these, specific recommendations for programming.

Examples of outcomes and impacts from selected comparator approaches

Outcomes from the PEER Study by CARE in Zambia on Sexual and Reproductive Health Services for young people (Price and Hawkins, 2002)

The study identified groups who were not being reached by youth friendly services, in particular: young women engaged in commercial sex; out-of- school youth; men who spend time in bars; and bus-drivers and conductors. Also identified the need for youth educators to understand more about key issues raised by the

[62] VECO is a Belgian NGO.

research including nature of commercial sex activity in the compounds; unsafe abortion and treatment of complications of abortion; accessibility of condoms and condom usage; and access to drugs for STI treatment.

As a result, potential new programme approaches to reach more vulnerable and marginalized young people in the compounds were also identified. One such approach was the use of peer networks to reach young people in locations where high-risk behaviours take place, such as in bars, nightclubs and around mini-bus stops. As a result of PEER the peer evaluator/researchers developed their own proposals to continue the PALS project in their compounds.

Outcomes & impacts from the HCD exercise by Marie Stopes (Kenya and Zambia) and IDEO.org on Family Planning/Reproductive Health services for adolescents (ITAD, 2017)
MSI recognized that the innovative approaches used during the inspiration phase helped to understand the desires and aspirations of urban unmarried girls in a way that standard qualitative research had not, putting girls 'front and centre' in the design. In both places the HCD solutions appear to have been effective in reaching adolescents and contributing to an increase in uptake of FP/RH services by adolescents, although results were preliminary and the evaluation was too early to determine if the solutions would be scalable

Outcomes from the 'general' qualitative study on adolescent girls' economic empowerment in Rwanda by Girl Hub and the Nike Foundation (Calder and Huda 2014)
The study identified some important findings and corresponding recommendations for programming, including the need to:

- Work with power holders, in particular men and boys to support positive change for girls, to change the 'value' that households, communities, and others put on girls, so that future economic calculations are in her (and her younger sister's) favour. They need entrepreneurial and business skills as well as access to capital, and training and micro-finance should be tailored specifically to adolescent girls, as they currently were not.
- Address the risks to girls dropping out from school and engaging in risky behaviours by tackling their immediate economic resource needs within economic empowerment initiatives, and focusing on getting girls back in school and keeping them there.
- Build social capital for girls within families and friends, especially for 13-15 year olds who may be attending school less regularly or not at all.
- Provide appropriate savings vehicles as an essential part of initiatives, as girls expressed a desire to save but a lack of safe places to do so.

More generally, it also identified the period between 13 and 15 years old appears to be a critical one for girls, particularly in terms of building self-efficacy and livelihood capital: social, human and economic. It further provided concrete program design options to address the above recommendations.

8.2 Assessing Value for Money of RCA

Key findings

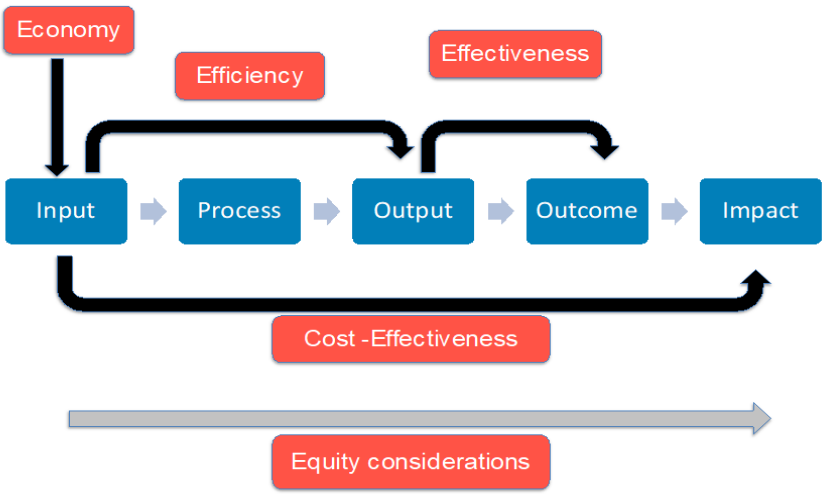
1. RCA sits towards the middle of the range of comparators in terms of costs.
2. RCA also is in the middle range in terms of research and policy outcomes.
3. Taken together with the relative costs, RCA therefore appears to be relatively weak in terms of both efficiency and effectiveness, as well as equity, as a result of shortcomings in the rigor of the approach.

Before turning to our assessment of VfM in Section 8.2.2, we first provide an introduction to VfM concepts.

8.2.1 Understanding Value for Money

Value for Money refers mainly to the '3Es' – economy, efficiency, and effectiveness – which are related to particular stages in the results chain. Economy is achieved through the minimisation of the cost of inputs, while efficiency involves maximising the amount of output achieved for a given input. Effectiveness relates to the extent to which the intended outputs lead to the desired outcomes. Finally, overall cost effectiveness relates to the extent to which the whole causal chain, from inputs to outputs to outcomes, results in the desired impacts. Finally, a 4th' 'E' is equity, which cuts across the entire results chain. This is illustrated visually in Figure 7.

Figure 7 - The causal chain and the 3Es



In other words, assessments of the value for money come down to an analysis of, firstly, whether the hypothesis underlying the programme design (articulated in the Theory of Change) is correct and, secondly, whether the desired impacts were achieved at the least cost. The objectives of VfM analysis are therefore to:

- Understand 'what works' through hypothesis testing: are the assumptions in the causal logic valid? Do the activities lead to the right outputs, outcomes, and impacts?

- Understand the key cost drivers: what kind of trade-offs can be made on cost, quality, and scope without sacrificing results? Are there ways to manage resources more effectively to achieve the same results at lower cost?

The answers to these questions are not always straightforward, however, because there are often trade-offs to be made across the objectives of the 3E's. Especially important are the trade-offs between efficiency and effectiveness, where driving down the cost of achieving outputs (for example by minimising the use of inputs) leads to deterioration in the quality of outputs, so that outcomes are not achieved as expected. In other words, output quantity might be traded off for

quality, which could jeopardise results. These trade-offs are especially relevant in the context of RCA and its comparators, where greater efficiency in the research – whether in terms of the scope of fieldwork, the experience level (and hence salaries) of researchers, or the level of inputs of senior researchers - may compromise quality. Clearly, the objective is to ensure that the trade-offs in the 4Es are all balanced so that overall value for money is achieved.

8.2.2 Assessing Value for Money of RCA: Thinking about VfM metrics and how to assess

Given the nature of this formative evaluation and the fact that there is little quantitative data available on outcomes and impacts of either the RCA or its comparators, we do not attempt to be definitive here (providing, for example, a full cost-benefit analysis), but will rather use this opportunity to:

- Flesh out the cost drivers and potential VfM metrics that could be used for future assessments;
- Articulate whatever rough sense of costs is available from the literature; and
- Set out an initial view of how RCA compares in terms of overall VfM.

Outputs: cost drivers and efficiency

At output level, the focus is mainly on how efficiently inputs translate into outputs. This therefore entails simply analysing the costs on a per-output basis. Across all four approaches, the basic output is a research product (report, briefing, etc), which has essentially three elements: the research findings themselves, the presentation of those findings, and communication/advocacy of the findings.

To get to these outputs, the basic set of inputs is the same, used to varying degrees in each:

- Research design: developing research framework, instruments
- Training researchers
- Fieldwork/data collection
- Synthesis/debriefing
- Report-writing
- Presentation/dissemination

The basic set of ingredients for cost drivers is also basically the same, including:

- Lead researchers (national and international)
- Local researchers
- Fieldwork transport and subsistence
- Materials
- IT

The table below shows the relative size of these cost drivers in each approach. In terms of senior research expertise, HCD is the most resource-intensive, as it involves large teams of experts from a range of disciplines, normally internationally-based. Sensemaker is also resource intensive, as it requires a fair amount of capacity from international[63] experts (whether internal to an organization or externally hired), while RCA, PEER, and general qualitative approaches all have a lower level of senior research inputs. These might be somewhat higher in general qualitative approaches, depending on the specific methodologies used, if large amounts of data are analysed.

[63] We distinguish here between international and national experts because they tend to have different cost levels associated with them; international here refers to any expert working outside their home geographical area and paid at the higher international pay scales.

The cost of local researchers is lowest in PEER, HCD, and Sensemaker, which all rely to varying degrees on input from peer researchers/facilitators and participants themselves.[64] The amount of experience is therefore lower than RCA and general qualitative approaches, which rely on more highly educated researchers with, in the case of general qualitative work, more experience in research.

Training costs for each research exercise are similar in RCA and general qualitative work in terms of duration (and therefore researcher costs), even if the actual scope/depth of the training is greater in general qualitative work. Training requirements for Sensemaker can be significant when staff within an organisation are new to the approach, as the staff who are doing the research itself must undergo a significant period of learning on the technical aspects of implementing the approach, and also wider consumers (for example M&E officers, programme staff, senior management) who need to be able to understand the approach in order to appreciate the findings. Training costs for PEER are relatively low,

Fieldwork costs themselves are fairly low in RCA, PEER, and Sensemaker, since the amount of time in the field is quite short and in the case of RCA researchers stay with participant households. Fieldwork costs can be somewhat higher for general qualitative approaches, depending on the scale of the research.

The only method with a large IT cost is Sensemaker, which requires investments in software licenses and proprietary software customization.

Table 8 - Relative scale of cost drivers by comparator

	RCA	PEER	HCD	Sensemaker	General Qual
Lead researchers	+	+	+++	++ (if implemented within an organization) +++ (if external experts required)	++
Local researchers	++ Trained college-educated national researchers	+ Local peer researchers	+ local peer researchers	+ Stories from participants Local facilitators	++ Field researchers are experienced
Training	++	+	++	+++ (if implementers are unfamiliar with the approach) + for field research	++
Fieldwork transport/ subsistence	+ stay with local people, but reasonably large team in field	+ limited external field time, with majority done	+ limited period of immersion	+ limited external field time, but often large sample	+ to ++ depending on scale

[64] The composition of research teams in HCD is very variable, and can often consist of only international team members without any local researchers. Our description here is based on what was reported in the case studies (See Annex 13 for further description of these).

		by peer researchers			
IT				Investment in Sensemaker software +++	
Overall costs	++	+	+++	++	+ to ++
Estimated cost/study	$120,000 (Indonesia average within DFAT programme); Other studies for UNICEF and World Bank ranging from $50,000-$120,000 depending on size of study and whether overheads included or not._	$27,000 (CARE Malaysia)	$400,000	No Data	$30,000 - $70,000 for a single country study

In theory, the approach that was able to produce research for the lowest cost would be the most efficient. This implies that efficiency can be gained from either: reducing the scale of the research (fewer locations/smaller sample), using 'cheaper' researchers, and minimizing international expert time.

Across the comparators here, PEER would therefore be the most efficient, as it is implemented at an overall very low cost, and general qual work can also be similarly efficient. RCA and Sensemaker are both more expensive and therefore less efficient, and HCD is very expensive indeed (over 10-20 times the cost of PEER, for example) and therefore fairly inefficient from the perspective of research alone, although this is part of a wider design phase that yields additional outputs beyond the research itself.

Outcome: Effectiveness

At outcome level, the VfM metrics are able to account not just for efficiency, but also effectiveness: what is the quality of the research insights, presentation of findings, and advocacy that results from each approach?

Given that we don't have an actual real-life counterfactual (two different research exercises are never undertaken simultaneously in the same place to compare the findings of each), it is difficult to assess this with precision. But the question of the counterfactual is still of central importance, namely: would a different (potentially less expensive) method be able to achieve the same results as RCA?

The key to assessing VfM in terms of effectiveness ultimately rests on the quality dimension, which is encapsulated in terms of the rigor of the exercise and, relatedly, the strength, validity and relevance of the findings. Further details of example studies from comparators are provided in Annex ZA. From a VfM perspective, the ideal scenario is an approach that is high quality and low cost. We find that on balance, it is the 'general' qualitative and PEER approaches that

achieve both higher quality and lower cost. By contrast, HCD is very high cost and tends to be lower quality[65]. Both Sensemaker and RCA are in the middle, with a range of quality in terms of research output, and also reasonably expensive.

Impacts: Overall cost effectiveness

To assess overall cost-effectiveness – in terms of the impact on policies and programmes on people's lives – would require research exercises to report on their ultimate impacts, which could be several years after a study has concluded. Unfortunately, this 'evaluation of research' is almost never done in practice. It would also be highly dependent on the commissioner, and the to what extent they actually use the findings and advocate for change as a result.

Based on our findings about RCA rigor and effectiveness, the working hypothesis would be that RCA impacts might be lower than some comparable methods.

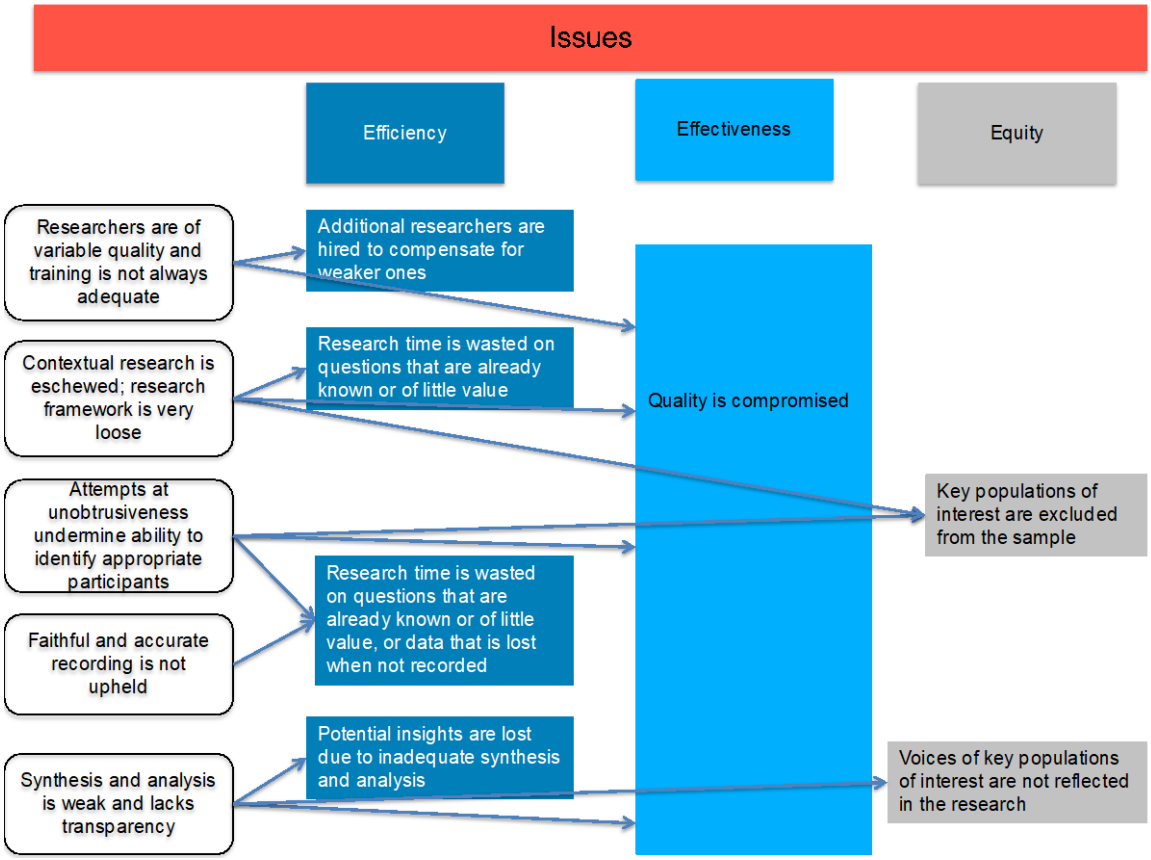
In the absence of actual concrete information on this kind of research impacts, it highlights the importance of commissioners being clear from the outset what the whole theory of change is for a research output, and beginning to commission some follow-up/monitoring of impacts.

8.2.3 Summary of VfM issues with RCA

Although a full VFM assessment is limited by our available data, the emerging picture based on what we do know is that RCA appears to be relatively weak in terms of both efficiency and effectiveness, as well as equity, as a result of gaps in the rigor of the approach. The figure below summarises the findings, using the rigor framework from Chapter 6 above.

[65] In theory, HCD should use a rigorous approach, since it simply draws on a range of qualitative tools, but in practice tends not to always be implemented with a high degree of rigor. For that reason, we depict the quality against a fairly large range.

Figure 9 – Summary of Findings



9 Conclusions and recommendations

Findings	Conclusions	Recommendations
Research design and preparation		
Experienced, reflexive and well-trained researchers		
1. There is significant variability in experience of RCA researchers – some have no qualitative research experience prior to joining RCA studies - and training is insufficient for those with little to no background conducting qualitative research. 2. While training materials for Level 1 are of reasonable quality (and would serve as a good refresher and RCA orientation for experienced qualitative researchers), training materials for other levels are insufficient to support progression from novice to more experienced researcher levels.	**The lack of experience and training amongst RCA researchers is problematic for a number of reasons.** Primary among them are that a lack of understanding of social and power dynamics limits the ability to know how to interpret what is being said, and inexperienced researchers are likely to be ill-equipped to know when and how to effectively probe issues (to 'know what they don't know' and to know what is important). Similarly, while inexperienced researchers may have different assumptions than more seasoned researchers, there is no reason to believe that they might have fewer assumptions. Indeed, more experienced researchers have spent years having any assumptions they might have had being thoroughly challenged in the course of research, years developing reflexivity, and years honing deep listening and probing skills. Level 1 training materials do provide a good foundation for growth, but it is questionable whether further levels of training support this adequately.	<u>Practitioners:</u> **We have three overarching recommendations for practitioners:** 1. Place a much stronger emphasis on recruiting RCA researchers who have some previous qualitative research experience as well as knowledge of the issues that RCA studies will explore 2. Ensure that training is of sufficient length 3. Revamp training materials to improve content, and ensure a stronger progression of learning and competencies be established For level 1, we suggests more extensive training on a range of issues. These include: understanding power and social (including gender) dynamics; methods for recording and reflecting on data and using this to plan further field activities; the range of participatory approaches that can be used to aid in local people sharing understanding; ways to encourage people's own emic analyses (beyond probing); and research ethics and child protection. Level 2 does a very quick review of Level 1 training on day one, and misses an opportunity to go much deeper in to technical research issues, as well as to expand researcher's repertoire of research tools. A robust training module on research ethics and child safeguarding is absent and should be included. Level 3 should include much stronger synthesis and analysis training, and this should be ramped up even further in Level 4.

		Much more time on this in pre-field briefing – where reflexive practice can be related to actual research locations and themes - and in debriefing - is recommended. While a short training (5 days according to RCA reports but this can be shorter, and these 5 days include a 2 day immersion) might be a sufficient "refresher" on key aspects of qualitative research and an orientation to RCA as a distinct approach for more experienced researchers, it is assessed to be inadequate for inexperienced researchers. We recommend a full 5 day workshop, followed by a 2 day immersion, and a further 2 day workshop to debrief thoroughly. <u>Commissioners:</u> We recommend that commissioners request and scrutinise the CVs of qualitative researchers to ensure that they have sufficient skills and experience to carry out the research required.
Contextual Understanding		
1. RCA actively excludes contextual understanding, claiming 'virtue in ignorance' as a way to eliminate researcher bias. 2. RCA researchers therefore, unsurprisingly, neither seek nor achieve good contextual understanding prior to conducting research.	**There is inadequate research and orientation on key sectoral and other relevant issues prior to studies being undertaken.** **A lack of good contextual understanding makes it difficult to be fully reflexive and aware of how one is interpreting things and how they should be interpreted within a particular context. This can lead to bias.** Listening to people in their own context, as RCA core team members admit, is "surprisingly hard to do well" and needs not only a keen attention to detail, curiosity, and a good memory, but also an understanding of the context in which people are sharing their thoughts and experiences, and a deep knowledge of the ways that issues such as social norms and power dynamics influence attitudes and behaviours. While RCA claims that "people who are really knowledgeable go into studies with these pre-determined biases", we suggest that properly trained and experienced researchers are well aware of this possibility, and address it through rigorous reflective practice rather than avoidance of important contextual information.	<u>Practitioners:</u> **RCA would benefit from significantly more background research to inform the entire study design and execution**. This would aid contextual understanding so some extent. We recommend that this include both desk-based research and discussions with expert practitioners. **The regular training components (Levels 1 – 3) should be complemented by an expanded pre-study briefing that adequately covers relevant contextual issues.** **We recommend that the very simple and short (1 day) briefing prior to undertaking a study be complemented with a further 1 day "deep dive" into the context and the issues under study.** This would enable researchers to develop a more robust understanding of: 1. specific contextual issues related to their field site (for example, are there ethnic or religious tensions there? Is this an village that has had previous poor experiences with extension workers? Has a recent disease outbreak lead to a

	Further, we suggest that there is a real risks in using inexperienced researchers who have little contextual or issue-based knowledge. They are more likely to: 1. Be unable to discern which information is most relevant to retain and why, 2. "re-discover" what is already known by more knowledgeable researchers and 3. fill gaps in their knowledge with their own assumptions. **While awareness of own positionality and biases appears to be a key component of RCA Level 1 training, it does not go deep enough in relation to gender, or local contextual issues**.	number of infant deaths? Are there particular norms regarding gender that researchers need to consider?) 2. specific issues to be researched, for example agricultural livelihoods, or education 3. specific research populations, for example, adolescents, and 4. particular approaches to research that might work best given the context, issues and population to be researched, for example participatory tools with younger or less literate populations <u>Commissioners:</u> We recommend that commissioners request and scrutinize the CVs of qualitative researchers to ensure that they have experience of the issues and contexts in which research will be conducted. Commissioners should also ensure that a secondary research phase be planned prior to primary research being carried out.
Framework to Guide Inquiry		
1. RCA uses a series of questions, captured in "areas for conversation" that enable an open, exploratory approach to research; however these are often very vague, leading to low-quality data in terms of relevance and depth. 2. At the same time, there is little evidence that this approach either reduces potential bias or leads to more meaningful conversations.	RCA's method of using an "Areas of Conversation" checklist to remind researchers of the kinds of areas that might be relevant to the overall research theme enables researchers to conduct open and exploratory conversations, led by research participants, and this is to be welcomed. However, the complete lack of focus in a vast majority of these checklists is extremely worrying. Further, there is no reason to believe that areas of conversation contain any less bias than research questions or lead to richer and more meaningful conversations. It also reduces the comparability of data collected across multiple interactions and multiple sites. **Without a good conceptual framework and a set of research questions that have emerged from this, a concern is that researchers having little to no background in the context or issues being explored often fail to identify what lines of inquiry are worth pursuing and what are not; what statements might need to be challenged or probed further; and when attitudes, beliefs and behaviors conform with or challenge norms**. One need not develop a rigid theoretical frame to guide research; a conceptual framework	<u>Practitioners:</u> **Future RCA studies could pilot the use of a conceptual (rather than a theoretical) framework for inquiry.** **Significantly more attention needs to be paid to formulating "areas for conversation".** A more comprehensive review of previous research on the issue under study be undertaken, including identifying assumptions that need to be tested through more immersive research. From this, a research framework that is sufficiently detailed to ensure that key areas are being investigated, but not so detailed that it would be impossible for any team of researchers to investigate during a short immersion. This research framework should be reviewed by at least one, and preferably several people, who have expertise in the issue to be explored, and the population segment with which and geography in which the research will take place. It cannot be assumed that commissioners have the relevant expertise to do this.

	also allows linkages and relationships between issues to be explored in a more open manner. For example, exploring the relationship between formal and informal institutions related to land inheritance, or between social capital and human capital. A conceptual framework can also help to reveal and be explicit about assumptions, through a deeper pre-field discussion of how, and how well linkages and relationships within the framework are evidenced.	Commissioners: Commissioner ToRs should be informed by a light touch review of previous research on the issue under study be undertaken, including identifying assumptions that need to be tested through more immersive research. Commissioners should include sector and research specialists where possible in review of research frameworks and approaches, either from their own staff, or brought in to provide quality assurance.
Multi-disciplinarity		
2. RCA seeks to achieve multi-disciplinarity through putting together large diverse teams. However, the RCA fails to achieve multi-disciplinarity, as researchers' disciplinary backgrounds are largely irrelevant to specific RCA studies.	**Simply including researchers with different academic backgrounds, does not make RCA research multi-disciplinary**. RCA researchers are not recruited for the disciplines they practice: indeed, this appears to be irrelevant. And, they are not expected to utilise frameworks or experience from their professional training.	Practitioners: **There would be advantages to deploying teams with relevant experience across several disciplines, even if this requires RCA teams to be smaller, and to perhaps cover two or three sites in one study.** The views of these researchers should be actively sought not only in developing a research framework, and helping to increase the contextual and issue-based understanding of the entire team prior to fieldwork, but should also be seen as a valuable resources during fieldwork and analysis. There are a number of comparator approaches that capitalise on this approach. Strong researcher training on bias and reflexivity should enable researchers to bring their multi-disciplinary perspectives and working experiences to bear on research design and analysis without unduly biasing the research, and without muffling the "voices of the people". Instead of avoiding this risk by not seeking multi-disciplinary perspectives, skilled qualitative researchers take this responsibility very seriously, and guide and support their teams to do the same. Commissioners: Commissioners should demand research teams that have relevant multi-disciplinarity.
Ethical considerations		
3. Training and orientation do not inadequately prepare field	**While RCA's own ethical guidance is taken from the American Anthropological Association Code of Ethics (RCA, 2017), we do**	Practitioners:

researchers, sub-team leaders and team leaders to act in the best interests of children and other vulnerable research participants in the field. 4. Ethical review prior to fieldwork is not standard, and it is clear that commissioner approval processes are insufficient in this regard.	**not consider these to be appropriate or sufficient ethical guidelines for conducting short-term research for development with large teams of researchers**, many of whom have not conducted field research previously, or whose only exposure to qualitative research is through the RCA. The RCA, and indeed any qualitative research approach, cannot claim to be inherently ethical and should never be seen as such. **Training and research preparation have gone some way to addressing ethical issues, but there are considerable gaps and shortcomings that render RCA design and preparation unable meet high ethical standards.** The training materials and new procedures we reviewed – including several risk assessments generated by these new procedures - are adequate to mitigate some of the risks experienced by researched populations and researchers that we discuss in this report. **The responsibility for building ethical foundations rests not only with the RCA core team, but also with the entities within which RCA is "housed", and commissioners.** If ethical foundations are not laid during researcher training and research design and preparation, there is a much greater likelihood of ethical problems emerging during research execution, putting both researchers and research participants at risk.	As RCA research involves "common" vulnerable populations such as children, there should be a **comprehensive standalone researcher training on child safeguarding, adequate orientation to child protection policies and procedures that guide the research, and discussion of ethical issues related to other vulnerable populations such as women experiencing domestic violence,** regardless of whether this is a requirement by the commissioner, who may not be informed enough to know this should be a clear expectation. This will also support key personnel, such as Child Protection Officers, and researchers to understand and appropriately address safeguarding issues during fieldwork. **A key means to mitigating risks associated with "limited knowledge of the local context and support networks" would be to increase researcher knowledge of the local context and support networks.** This would both help to address rigor shortcomings and improve effectiveness and relevance of research outputs, and, provided researchers are reflexive, should not unduly bias them. Commissioners: **A more formal ethical review and approval process is recommended** in light of the fact that RCA research addresses sensitive issues or topics, involves vulnerable groups, uses considerable participant time, and is largely exploratory. This review should consider contextual issues related to research design.
Fieldwork		
Being unobtrusive		
3. RCA prides itself in researcher unobtrusiveness, and on some counts it has good reason to do so. Considerable time is spent preparing researchers to "fit in" while staying in local communities.	**Considerable time is spent preparing researchers to "fit in" while staying in local communities.** This is a good practice and is to be applauded. **However, a lack of proper preparation and careful selection of households prior to the full immersion not only undermines the achievement of selection according to poverty proxies, but can make researchers more obtrusive, not less**. It is unrealistic to	Practitioners: **Continue the good practice of training and orientation on how to be less obtrusive in the field, but use much deeper contextual knowledge to support this, and completely re-think the strategy for entering the community and obtaining consent.** Much more careful and reflexive sampling design and implementation strategies are also required, to grapple with and ultimately balance the need to be

4. We find, however, that circumventing local protocols is a poor strategy. A lack of proper preparation and careful selection of households prior to the full immersion not only undermines the achievement of selection according to poverty proxies, but can make researchers more obtrusive, not less.	assume that researchers can enter under the radar in any community, where the radar is generally very good indeed! It is a fallacious assumption that researchers can ever truly be unobtrusive and that fieldwork can ever be anything but an intervention. **A covert, one-size fits all approach is disrespectful, and ineffective and, as we discuss in the report, can do harm to households and researchers**. Part of doing ethnographic fieldwork entails time and care taken to explain to local authorities the purpose of the research, and the importance of living with basic households. Flouting these conventions is not only disrespectful, it reinforces, rather than challenges, power dynamics by the researcher assuming that "normal" conventions don't apply to them. The RCA core team acknowledge that there are some trade-offs between complete openness of the presence of researchers on the field with the effort to gather genuine and candid responses from study participants (RCA+ transitional Report). **We suggest that the trade-off is unacceptable, and fails to meet minimum standards for informed consent**, including making people aware of any potential implications stemming from how their views on a policy or program are positioned in a public document, and the implications of a commissioner's response (or non-response) . **Suspicion, and lack of clarity concerning the purpose of the research and, potentially, host households who feel that they had to consent to researchers staying with them, all erode good feelings and trust.** RCA practitioners claim the RCA enables a level of trust to be built between researcher and research participants that is absent in other approaches. However difficulties in building trust raised in RCA documentation, and by researchers themselves suggest that there is some over-claiming of a close and trusting relationship between researchers and household members.	unobtrusive with the need to have a good understanding of the context, and the participants who are being selected. . This should consider pre-selection of households or, at a minimum, allowing more time for researchers to select a household in the community. This would entail arriving earlier on in the day. RCA should also adopt community entry strategies that are tailored to the local customs. <u>Commissioners:</u> As part of an ethics review, ensure that good practice is being followed in terms of community entry and informed consent.
Triangulation		
3. We find that RCA achieves "triangulation over time" (particularly in longitudinal RCAs) and, to a very limited extent, "triangulation of people"	**The RCA approach of using conversation, observation and participation in daily life is a much-needed approach to social research**. It enables researchers to have more natural and free flowing discussions, gain empathy for research participants, and use observation to compare what people say with what they do. It also	<u>Practitioners:</u> **As part of an overall more careful research design, RCA methodology would be greatly strengthened by the integration of more standard qualitative tools such as interviews, discussions, and PRA. This would aid triangulation, comparison and**

(as the distinct voices of different research segments is often lost due to not faithfully recording conversations and aggregation in analysis). 4. Assessing RCA's achievement of "triangulation of methods" is somewhat more complex, but we find that this is weak and that this is a missed opportunity.	enables researchers to access types of information that cannot be gained through surveys or through other qualitative and participatory approaches. **However, despite the fact that the RCA enables researchers to speak to a large number of varied individuals, it is weak in terms of triangulation of people** as the distinct voices of different research segments are often lost due to not faithfully recording conversations and aggregation in analysis. **A failure to fully integrate other qualitative and participatory approaches renders it's triangulative power in relation to methods weak.** RCA cannot claim triangulation of method through merely practicing something akin to anthropological participant observation, especially as practiced with little to no record-keeping.	**aggregation of data collected across multiple sites by multiple researchers.** The use of more standard tools need not be at the expense of researchers "following their noses" or spending at least part of their research time without an agenda, and interacting in a more natural way with participants. There are numerous examples of mixed qualitative methods research that does this well. **Sufficient training and orientation of researchers should be undertaken in PRA and any new methods that are integrated to ensure rigor.** More thorough documentation would then allow accurate triangulation across participants to be undertaken in the synthesis and analysis phase.
Respondent validation		
1. There is little indication from reports or other documentation, or from interviews with RCA researchers, that respondent validation is a central part of a claimed non- extractive research practice, though some more experienced RCA researchers clearly do this to some extent.	**While clearly some of the more experienced RCA researchers do have skills in eliciting respondent validation as a natural part of conversations, respondent validation is not actively and purposively sought by RCA research teams. This has implications for both rigor and ethics.** Respondent validation is an issue of rigor because when a researcher is relying heavily on emic (insider) analyses of complex issues, the researcher needs to ensure that they have fully understood the research participants. This becomes even more important when the researcher themselves know little about the issue under discussion, as they will need to check both "factual" and "interpretive" data with participants in order to ensure that they insights that they deduce and present as implication's or recommendations to policy makers are valid. Respondent validation is also an issue of ethics because it helps to ensure that participants, and their views and experiences, are not misunderstood and misrepresented, and that the research process is empowering and not extractive. An important ethical foundation for this is that research participants understand how the information will be used and with whom as this affects the context for individual decision-making on consent; people might not mind being represented	<u>Practitioners:</u> **The RCA should ensure that respondent validation be more actively sought by researchers, and that it is an embedded practice with the approach**. This is absolutely critical in terms of both rigor and ethics. Additional training and practice would be required to do this well. Post-fieldwork debriefing and sensemaking processes should more rigorously "check" data through the lens of respondent validation. More purposive respondent validation could also aid with memory recall, as articulation and repetition of key pieces of data (called "acoustic encoding") can help these to move from working memory into short term memory. <u>Commissioners:</u> Donor timeframes and budgets often mean that the last step, returning to the field to check draft analysis and interpretation, is rarely done. This reduces both rigor – reducing validity of findings - and adherence to ethical good practice, and should seriously be considered as part of any research exercise. It is important to ensure that time, space, and resources are allowed for this.

	in one way to one audience, but very much mind being presented in another way to another audience. Validation can be sought in the course of the conversation ("am I understanding you correctly that…?", "so it seems to me that what you are saying is….?"), at the end of conversations or research exercises ("You have taught me so much. I think that what I have learned is x, y, and z. Is that right? Did I miss anything important from what you said?") and at the end of the research process or afterwards, by holding a group or community debriefing, or returning with draft analysis and interpretation to check this with respondents. Good practice suggests that all of these different levels of validation should be sought. But this would require that RCA researchers present themselves much more openly as researchers.	
Faithful and accurate recording		
1. RCA note taking guidance is inconsistent, and researcher practice is variable. 2. Evidence suggests that memory and recall is complicated, and much less reliable than suggested by RCA practitioners.	Ad hoc and inconsistent note taking practices significantly undermine RCA's claims to accurately report data collected in conversations, in particular, but also observations and experiences. Memory recall is deemed to be unreliable and unnecessarily introduce researcher bias and unsupported interpretation.	Practitioners: **We recommend that much more rigorous note taking, and sufficient training on when, and how, to do this well, be incorporated into RCA practice.** Note taking can be done in a way that does not interfere with rapport building, but would enable a much more real-time (either immediately following or later that day) and accurate, recording of information, in particular, who said what, and in what context. Note taking would also enable researchers to better start processing their experiences, observations and conversations, as both an aid to post-field analysis, as well as an aid to utilizing time in the field more effectively and efficiently. Commissioners: As routine practice, commissioners should require that researchers submit their raw data whenever possible.
Ethical considerations in fieldwork		
6. The RCA has paid greater attention to researcher safety in recent years, with a range of	**The RCA pays insufficient attention to critical ethical considerations, including consent, power and positionality.** The RCA's attention to ethics is neither broad enough nor deep enough,	Practitioners: **Involve local researchers in site selection to ensure a robust and informed discussion of any potential issues with research sites**. In

new guidelines and procedures for entering communities, explaining research, and obtaining consent.

7. While there have been changes in theory (and some practice) to how researchers introduce themselves in communities and obtain consent, it is clear that participants are often not fully aware that a "research process" is underway.
8. Despite the fact that the RCA recognises that researchers limited knowledge of the local context and support networks constitutes a risk, actions to mitigate this particular risk have not been initiated.
9. While there is an explicitly aim to include poor and excluded households in research, household selection procedures can exclude vulnerable populations. Compensation to households is not always sufficient to off-set losses from hosting researchers.
10. The RCA has weak child protection policies and procedures.

particularly considering that researchers are living in the homes of study participants, not merely meeting them in tea shops, field, or schools, or community meeting places. Nor is it nuanced enough for the wide range of different contexts in which research is conducted.

New guidelines and procedures on entering communities and obtaining informed consent are welcome. However, a less-than well thought out and executed community and host household selection process means that researchers are not familiar with local context within which the data is to be collected. Without prior knowledge of local dynamics and proper introductions, researchers and the household in which they stay can be exposed to risk, during, and after the researchers have left.

A lack of completely informed consent can result in suspicion, which reduces openness required for research participants to engage fully in research processes, as well as placing research participants at risk.

The assumption that the poorest and most vulnerable can take in a stranger – who comes from the city – might indeed be borne out in many cases, but RCA, at a minimum, needs to have a good grounded discussion of potential issues with this approach. One might also query whether this specific research in Ethiopia would pass the 'benefice' test of ethical research, which sets a standard and justification for research that prioritises the benefit to participants.

As noted, training and orientation on child protection policies and procedures is extremely limited and does not equip RCA researchers to reconcile their responsibilities for child protection with RCA's quasi-anthropological non-intervention policies. RCA researchers are confused about what their responsibilities are in relation to responding to child safeguarding incidents, or indeed even what would constitute a child safeguarding incident.

addition, ensure that contextual information on research sites is sufficient to inform selection and to mitigate risks.

Immediately revamp the approach for entering communities ensuring that proper protocols are followed and permission is sought prior to research. Discuss with the local community the purpose of the research, including the type of households that researchers need to stay with. Ensure that households are not coerced into hosting researchers, and are adequately compensated (particularly in new and unfamiliar contexts). This will potentially require researchers to work with local mobilisers, and entail an additional time investment in household selection.

Immediately develop and implement a vulnerable population safeguarding plan based on accepted ethical standards of research with vulnerable populations and in particular with children. As part of this,

- ensure that a safeguarding plan is flexible enough to enable RCA research teams to consider how ethical principles might apply differently in different contexts.
- Ensure that roles and responsibilities for implementing this are made clear and that those responsible in the core team have adequate knowledge and support to do so.
- Incorporate discussions of child and vulnerable population protection issues in debriefing days to ensure that potential issues and incidents are being robustly discussed, and ensure any necessary reporting and follow-up.

Based on this, refresh the researcher training and train all researchers on the new policy and plan.

<u>Commissioners:</u>
Put in place a comprehensive ethics review process that reviews the above and ensures that good practice is being followed.

Analysis and Reporting

Debriefing, synthesis and analysis

5. While RCA practitioners claim that iteration is one of the cornerstones of RCA rigor, it appears that iteration is not done with any rigor, and that RCA research does little to support collective analysis by the research team.
6. The debriefing process enables an impressive amount of information to be "downloaded" from the researchers to the Team Leader. However there appear to be considerable weaknesses, primary amongst them the somewhat extractive nature of the process,
7. Analysis in RCA follows a process of coding and charting that is unsystematic, or varying quality, and supports organization of data but not rigorous analysis.
8. RCA gives complete primacy to emic views over etic interpretation – rather than a judicious balancing of the two.

Iterative analysis processes when researchers are in the field are absent and this is a missed opportunity. Researchers take few notes (so are not processing information in the field as part of early analysis) and do not meet up while in the field to discuss emerging findings and other key issues.

Collective and participatory analysis processes are weak, and Team Leaders, who may not have been in the field, conduct the bulk of analysis.

Debriefing, while enabling an impressive amount of information to be "downloaded", suffers from a number of flaws. It is relatively extractive, which leads to limited and arbitrary probing, clarifying, and nuancing of information; little attention to reflection on possible biases; and overcategorization and limited exploration of different respondents' experiences and viewpoints.

The RCA's sole reliance on emic views, rather than judiciously balancing this with etic interpretation, yields descriptive data but largely fails to surface meaningful ethnographic insights. The RCA approach generates a collection of narratives, interesting, at times informative, but not analytical, and prone to bias due to a very weak understanding of power, positionality and context.

The RCA does not analyse issues of gender and social difference, rendering it unable to achieve interpretive depth. Not only that, RCA uncritically collects people's views without deeper interrogation of their positions, biases, blind spots, which can magnify researcher bias and be erroneous and misleading to policy makers who may assume, for example, that the lack of explicit reference to gender issues means that these are absent and therefore do not need to be considered.

Though RCA practitioners claim that grounded theory can emerge from RCA analysis, it is not clear how this is possible without the analytical process supporting the building, testing and evidencing of hypotheses. RCA's use of "framework analysis"

Practitioners:

The RCA core team should reflecting further on how iterative analysis could be better supported through the fieldwork process. This might be through stronger note taking, reflection and fieldwork methods to encourage research participants to engage in more purposive interpretation; or through check-ins with other team members and sub-team leaders during the course of research.

In terms of iterative analysis, the RCA core team could even consider experimenting with slightly longer immersions (6 - 10 days, for example) that would enable sub-teams to meet up part-way through the immersion to discuss challenges and emerging findings, and to adjust their approach on the basis of this learning. We agree with Shah (2018), that the first phase of research could be very open-ended and exploratory (akin to the current approach) but that the second phased could use more structured methods alongside RCA's looser immersion approach in order to stress test hypotheses through greater triangulation and soliciting participant feedback to validate, refute, or qualify findings. Leaving a research site and returning can be very good for building trust, and so a "punctuated" approach would also have this advantage.

Drawing on more detailed and comprehensive field work notes would enable the debriefing and synthesis process to generate richer and more robust data. We also recommend seeking a judiciously balance between emic and etic analysis.

We suggest longer, more structured and more collaborative debriefing and sensemaking sessions. We recommend a 1.5 day debriefing per sub-team. The debriefing should also be more structured, and should illicit segmented data. Collaborative charting and coding would enable researchers to generate shared analysis. This would be a valuable input to the sensemaking session, meaning that sensemaking could focus more on making sense of data generated through debriefing rather than on generating more data, which it does currently.

	appears to be useful for the purposes of organising data (particularly in relation to report writing), but does not appear to be fully utilised as an analytical tool.	**All RCA studies should ensure at least a 1 day, and preferably a 2 day sensemaking session with all of the researchers who conducted fieldwork, as well as any researchers who will be doing further analysis and report writing**. This would enable report writes to develop a stronger connection with the data and for researchers to more effectively influence analysis. **We recommend that the RCA completely re-think it's stance on interpretation.** We suggest that more experienced and well-trained researchers and much stronger processes to support local analysis, combined with processes for debriefing synthesis and analysis that are able to carefully combine emic and etic perspectives and include much stronger gender, social and power analysis. RCA practitioners should utilise the most relevant recognised frameworks to enable this analysis. <u>Commissioners:</u> Commissioners should demand high quality gender, social and power analysis in all of the qualitative research they commission, including evaluations. If internal capacity to assess this is insufficient, then robust quality assurance should be commissioned.
Reporting		
1. RCA reports are rich and include an impressive amount of detail. While RCA reports are interesting, and at times informative, without understanding power, positionality and context, little meaning can be generated, and this is where RCA reports fall considerably short. 2. The presentation of data – particularly in the form of quotes, boxed case studies, and tables – is, at times, problematic.	**RCA reports do capture the interest of their readers, and provide a great deal of detailed information. Without interpretation and clear recommendations, however, RCA reports do not live up to their full potential.**	<u>Practitioners:</u> Unless the RCA evolves to become more analytically rigorous and to include expert interpretation – whether etic or emic but preferably a combination of both – **it should steer away from providing conclusions or recommendations**. **RCA reports would be strengthened by presenting more and better quality data obtained through PRA methods, as well as more rigorous presentation of quasi-statistics**. This would also strengthen the triangulative power of the RCA. <u>Commissioners:</u> Quality assurance of final report products should be carried out be qualified individuals, either as part of an internal review process or an external, commissioned, review.

3. RCA reports present a new perspective, one which is often not seen by policy makers and programmers; however there is some blurring of lines concerning the presentation of "findings", " conclusions" and "recommendations" that can be confusing.		
Transparency		
3. Greater transparency has been achieved by introducing a more formal archiving process, but a lack of systematic note taking and rigorous debriefing leaves room for improvements in this area. 4. Reports do a reasonable job of highlighting research limitations; however, both a failure to properly consider how researcher positionality might have influenced the research, and a failure to consider the potential magnitude of limitations and impact of these weaken transparency.	**It is clear that recent improvements in the documentation process have been made** in the last several years. With the addition of more formal archiving processes greater transparency has been achieved. **However this review found that there is still considerable room for improvement.** In particular, a lack of systematic note taking and rigorous debriefing reduces transparency. **Report do a fair job of acknowledging limitations, but there is scope for improvement in this area,** particularly concerning what they methods they use can and cannot achieve, and how their approach, and their sampling, affected findings, and how far they were able to mitigate limitations.	<u>Practitioners:</u> **We recommend that note taking in the field be more rigorous and that these notes be archived along with other materials generated**. A more rigorous use of PRA methods would generate valuable visual material that should likewise be archived. **There needs to be greater transparency around limitations to studies** and how the research itself was situated within – and indeed could have perpetuated or exacerbated - local power dynamics. <u>Commissioners:</u> Commissioners should request access to archived materials. In some cases where research is novel or particularly sensitive, commissioners may want to consider a review of these materials by qualified individuals.
Ethical considerations in analysis and reporting		
4. The RCA archiving processes and secure data storage are adequate to ensure privacy and confidentiality. 5. The RCA has no feedback mechanisms to communicate feedback to research	**Archiving processes and secure data storage are adequate to ensure privacy and confidentiality**. RCA core team members are aware that preserving anonymity requires that there be no link between the data and the source, and work hard to prevent the identification of communities, households or individuals in reporting.	<u>Practitioners:</u> **RCA study teams should lobby tenaciously for the budget and time required to conduct feedback sessions** whereby local people are able to debate, confirm and contest research findings.

participants and to enable them to contest how they are being represented. While it is clear that the intention is to present research participants respectfully in reporting, this is not always the case. 6. The RCA has an internal peer review process that serves the purposes of client quality assurance but does not review content from a contextual or technical perspective. External peer review processes that could serve this purpose are absent.	**There is value in listening to a range of voices and presenting these authentically, but the RCA is plagued with a range of shortcomings that undermine respect for voice.** An approach whereby researchers "just listen" to what people say about their situation and provide what is perceived to be a relatively unfiltered perspective is valued by many. However, there are weaknesses in terms of faithful documentation and reporting of these voices, and tensions concerning informed consent that continue to undermine RCA in terms of respect for voice. Further, we argue that it is not ethically acceptable for an approach that claims to give primacy to local people views and interpretations to have no process or methods for actively engaging participants with the findings and their implications. **Internal peer review processes are inadequate** as they do not involve researchers with the experience and expertise necessary to challenge findings and recommendations (insights and implications).	**We recommend that research involve subject-matter specialists with experience conducting research in the geographies and with the populations with which the study is concerned**. <u>Commissioners:</u> Commissioners should ensure adequate budget and contract length to enable feedback to research participants.
Effectiveness of RCA as a research exercise		
5. RCA claims to reach poor and marginalized voices, but the sampling criteria are often poorly conceived – making it unlikely that the poorest or most vulnerable will be reached – and implementation of sampling on the ground compounds these issues. This raises questions with both the validity and relevance of studies. 6. The RCA is a simple and flexible approach that generates large amounts of detailed data; its interpretive depth is weak, however, and its respect for voice is largely limited to "good listening".	***Inclusiveness*** **RCA fails fairly comprehensively to deliver on one of its main claims about amplifying the voices that might otherwise be missed;** not only are they systematically under-represented in the sample as a result of the (lack of) rigor in the design process, their voices are then systematically under-represented in the synthesis and analysis that appears in the report. ***Validity*** **RCA reports provide significant descriptive detail; however true depth, both descriptive and analytical, is largely absent.** There is very little 'thick' description, instead quite a lot of aggregated or generalized data is provided, in a fairly unstructured way; there is no systematic presentation of findings against the key research questions. Evidence is not produced to convincingly back up the findings that are made. Instead, there is an impression of very selective use of examples, with no real analysis by group or context.	<u>Practitioners:</u> Where they stem from the lack of methodological rigor, many of the issues with respect to effectiveness would be addressed through the recommendations already listed above, in particular those related to sampling (whose voices are included) and the more careful and well-documented approach to synthesis, analysis, and reporting (authentically reporting those voices). If stronger background research were conducted in order to inform research design, these 'findings' should have been the *starting point* for the research, and then much more time and effort could have been put into the *what*, *why*, *when*, and *how* questions where the studies could have actually added value. Successful policy influencing requires not only the generation of evidence, but the provision of advice. <u>Commissioners:</u>

7. Some findings are presented as being surprising and counterintuitive, but primarily to those who have limited contextual understanding and knowledge. RCA is more valued for its ability to confirm, than for surfacing new knowledge. 8. RCA is strong at communicating some research participants' own perspectives through presenting a variety of case studies and voices; however, this is often very aggregative, with little sense of how case studies or indeed overall conclusions fit within the whole set of data collected, let alone how findings should be seen to relate to the wider context.	**Claims to be able to collect and analyse data on sensitive topics are also not supported by the evidence.** ***Relevance*** **Relevance is severely compromised by the lack of contextual research – which would help to situate findings within what is already known and therefore what value the research adds to the knowledge base – as well as the lack of a clear research framework.** The RCA claims to present findings that are surprising and counter-intuitive, challenging commissioner and consumer assumptions. While it is sometimes the case that research findings are counterintuitive and surprising, and could only have been gained through immersive participant observation, some findings are simply resurfacing old knowledge. We also, importantly, conclude that the RCA should not, for reasons of rigor and ethics, be used to explore sensitive issues or issues that are highly complex and rooted in social norms.	More should be done to address the erosion of institutional knowledge - where insights are discovered, lost, and (hopefully) rediscovered, with changes in key personnel- so that more informed oversight and quality control of the research output can be provided.
Where does RCA work best?		
1. RCA studies have made valuable contribution to mixed method monitoring, evaluation and learning efforts, particularly at the beginning (scoping and baseline), but also as a retrospective lens, to better understand why particular changes may or may not have taken place. 2. Stand-alone one off exercises are less rigorous than longitudinal mixed method	**This review finds that if RCA findings are used to complement data collected in other ways, and are seen to provide a view onto a broader landscape than merely the landscape of "the project", then RCA – with the ethics and rigor caveats discussed in Chapter 6 – has the potential to provide an important contribution to conventional evaluation.** **RCA's particular value is in its ability to provide a glimpse into people's everyday lives**. In contexts where little is known and ethnographic insights are important for filling an evidence gap, there is a stronger justification for more immersive research. For example, framing security and justice issues in ways consistent with how they are understood by communities was seen to be an important contribution of the IP-SSJ RCA study.	<u>Practitioners and Commissioners:</u> **Stand-alone RCA studies are not recommended**, with the exception of high level "journalistic" pulse-taking exercises that can fill information gaps in regard to how policies are affecting local people, and highly exploratory landscaping exercises followed up by more rigorous research methods. **RCA studies should only ever be implemented when combined with other qualitative methods**, and preferably embedded within a mixed qualitative quantitative exercise. **RCA generated findings should not only be placed within a wider evidence base, but should be part of MEL efforts that include team members with deeper contextual knowledge** in order to

exercises and are not recommended, with the exception of high level "journalistic" pulse-taking exercises that can fill information gaps in regard to how policies are affecting local people, and highly exploratory landscaping exercises followed up by more rigorous research methods.	**The RCA's weakness is its inability to provide real interpretive depth**, which comes from a judicious combination of emic and etic. Shah suggests that, "*RCA findings should be presented as limited, though able to provide considerable insights and triangulation when interpreted within a wider evidence-base*" (2018: 23). **We similarly conclude that RCA, rather than offering greater local insights than other qualitative methods, can offer different and complementary insights.** **RCA has potential in contributing to theories of change that reflect more closely local people's views and aspirations for change.** However changes are need in terms of how these are used, and by whom.	counter the lack of experience and expertise within RCA research teams. **TOCs need to be owned and used by program teams – NOT RCA research teams – in order to be useful.** TOCs that include RCA generated insights could be used within longitudinal MEL as frameworks for further RCA enquiry. Presumably, if the ToC is informed by RCA generated insights, it is not seen as a preconceived research framework. This suggests possibilities for a slightly different type of RCA study following initial studies in longitudinal research and evaluation, with implications for increasing RCA's relevance and integration within mixed method research. <u>Commissioners:</u> Seek mixed method research whenever practical and feasible, even if one of the Qs relies on secondary data.
Policy influencing		
1. RCA is valued for providing insights into service provision from a service provider's perspective, something seen to be unique and that is not surfaced by normal monitoring. 2. RCA is seen by some commissioners to be a "blunt instrument" providing insufficiently detailed and disaggregated analysis of ground realities and actionable insights. 3. RCA reports have produced some insights that have been important for program development, implementation, monitoring and evaluation, as	It's clear that the ability to translate RCA findings into policy-relevant insights is not a simple technical linear process, but depends all three actors in the system – policy makers, commissioners, and researchers (Lewis et al., 2012). Where RCA has been able to "punch above its weight" in relation to policy influence (and these instances have been relatively rare) it has been due less to RCA as a method that is able to generate knowledge and translate this into actionable evidence, only very partially to connections that RCA has been able to make with policy actors, and more to the spaces into which RCA has been invited, and the opportunity that these have afforded. RCA has been most successful when i. it has strong commissioners and an institutional "home"; ii. Policy-makers are engaged throughout; iii. there is a significant gap between policy makers and poor people's lived realities	<u>Practitioners:</u> Continued attention should be given to understanding how evidence influences policy to be more pro-poor and inclusive. While political saliency and policy relevance of RCA study topics is good, this could be significantly increased by: 1. ensuring sufficient knowledge of previous research on the issues to be studied, increasing contextual understanding at the local level, and engaging researchers who are experienced in formulating questions to get to complex and highly nuanced information. 2. Continue to work closely with "evidence translators" that have political savvy and credibility. 3. Work with commissioners to plan for and dedicate significant time and effort to policymaker engagement, relationship building and co-creation. 4. Reflect on and address the relationship challenges that have plagued the RCA throughout its history.

well as for policy. Case studies where this is more apparent are where there is a strong institutional home for RCA insights. 4. The RCA+ team's approach suggests only a very nascent understanding of how researchers need to think and work politically in order to influence policy. 5. RCA has a checkered record in relation to working effectively with key evidence translators.		5. Continue to increase the accessibility of research and at the same time address credibility of research issues through some of the recommendations provided in other sections. <u>Commissioners:</u> Make a concerted effort to address demand side absorptive capacity issues, as recommended in DFAT's ODE Review of Research (2015). Increase the use of qualitative research evidence to improve the effectiveness of aid through ensuring a stronger link between research and M&E, and through the strong leadership at Post, perhaps linked to incentives around "doing development differently".
Value for Money		
1. RCA sits towards the middle of the range of comparators in terms of costs. 2. RCA also is in the middle range in terms of research and policy outcomes. 3. Taken together with the relative costs, RCA therefore appears to be relatively weak in terms of both efficiency and effectiveness, as well as equity, as a result of shortcomings in the rigor of the approach.	**In theory, the approach that was able to produce research for the lowest cost would be the most efficient.** Across the comparators here, PEER would therefore be the most efficient, as it is implemented at an overall very low cost, and general qual work can also be similarly efficient. RCA and Sensemaker are both more expensive and therefore less efficient, and HCD is very expensive indeed (over 10-20 times the cost of PEER, for example) and therefore fairly inefficient from the perspective of research alone, although this is part of a wider design phase that yields additional outputs beyond the research itself. **The key to assessing VfM in terms of effectiveness ultimately rests on the quality dimension, which is encapsulated in terms of the rigor of the exercise and, relatedly, the strength, validity and relevance of the findings.** From a VfM perspective, the ideal scenario is an approach that is high quality and low cost. We find that on balance, it is the 'general' qualitative and PEER approaches that achieve both higher quality and lower cost. By contrast, HCD is very high cost and tends to be lower quality. Both Sensemaker and RCA	<u>Practitioners:</u> Addressing the issues with respect to rigor above would alleviate many of the value for money concerns. Doing so might increase costs somewhat – hiring more experienced researchers, undertaking more involved research design and synthesis – but would also increase quality significantly. <u>Commissioners:</u> There is a need to be clear from the outset what the whole Theory of Change is for a research output, and there would be value in commissioning some follow-up/monitoring of impacts to be able to assess VfM against the full ToC. There should also be a greater understanding of the potential approaches that could be used for a particular research objective, and value for money considerations should enter at the commissioning stage. For example, rather than specifying that an exercise should necessarily use RCA, research could be tendered based on the research requirements, and then the approach offering the greatest VfM could be the one that is ultimately selected.

are in the middle, with a range of quality in terms of research output, and also reasonably expensive.

Based on our findings about RCA rigor and effectiveness, however, the working hypothesis would be that RCA impacts might be lower than some comparable methods.

References

External Documentation

ACFID. (2017). Principles and Guidelines for ethical research and evaluation in development.

Anderson, M.B., Brown, D., & Jean, I. (2012). Time to Listen: Hearing People on the Receiving End of International Aid. US: CDA Collaborative Learning Projects.

Andrews, M., Pritchett, L., & Woolclock, M. (2012). Escaping Capability Traps Through Problem Driven Iterative Adaptation (CID Working Paper No. 240). Cambrige, Massachusetts: Center of International Development at Harvard University.

Andrews, M. (2013). *The Limits of Institutional Reform in Development: Changing Rules for Realistic Solutions.* Cambridge: Cambridge University Press.

Arvidson, M. (2013). Ethics, Intimacy and Distance in Longitudinal, Qualitative Research: Experiences from Reality Check Bangladesh. *SAGE Publications.*

Bell, S., & Aggleton, P. (2016). Interpretive and ethnographic perspectives. Monitoring and Evaluation in Health and Social Development. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Bernard, H. R. (2011). *Research Methods in Anthropology: Qualitative and Quantitative Approaches* (5th ed.). Lanham, MD: Alta Mira Pr.

Bloom, P. (2010, June 4). Book Review - The Invisible Gorilla - By Christopher Chabris and Daniel Simons. *The New York Times*. Retrieved from https://www.nytimes.com/2010/06/06/books/review/Bloom-t.html

BPS – Statistics Indonesia (2016) *Consumption Expenditure of Population of Indonesia by Province 2016, Based on the March 2016 Susenas*

Calder, R., & Huda, K. (2014). *Adolescent Girls Economic Opportunities Study, Rwanda.* Development Pathways.

Carr, E. (2018). Open the Echo Chamber: A Blog About Development and Global Change. Available at: http://www.edwardrcarr.com/opentheechochamber/

Casella, D., Magara, P., Kumasi, T.C., Gujit, I., & Van Soest, A. (2014). *The Triple-S Project SenseMaker Experience. A Method Tested and Rejected (Working paper 9).* Triple-S.

Cernea, M. M. (1992). Re-tooling in applied social investigation for development. Planning: Some methodological issues. In N. S. Scrimshaw & G. R. Gleason (Eds.), *Rapid Assessment Procedures - Qualitative Methodologies for Planning and Evaluation of Health Related Programmes*. Boston: International Nutrition Foundation for

Developing Countries (INFDC). Retrieved from http://archive.unu.edu/unupress/food2/UIN08E/UIN08E00.HTM

Chambers, R. (1981). Rapid rural appraisal: Rationale and repertoire. *Public Administration and Development*, 1(2), 95–106. https://doi.org/10.1002/pad.4230010202

Chambers, R. (1994a). Participatory rural appraisal (PRA): Analysis of experience. *World Development*, 22(9), 1253–1268. https://doi.org/10.1016/0305-750X(94)90003-5

Chambers, R. (1994b). Participatory rural appraisal (PRA): Challenges, potentials and paradigm. *World Development*, 22(10), 1437–1454. https://doi.org/10.1016/0305-750X(94)90030-2

Chambers, R. (1994c). The Origins and Practice of Participatory Rural Appraisal. *World Development*, 22(7), 953–969.https://doi.org/10.1016/0305-750X(94)90141-4

Chambers, R. (2007). Overview – Immersions: Something is happening. *Participatory Learning and Action*, 57, 9–14.

Chambers, R. (2012, September 6). Robert Chambers - why don't all development organizations do immersions? [Blog]. Retrieved from https://oxfamblogs.org/fp2p/robert-chambers-why-dont-all-development-organizations-do-immersions/

Collins, D., Morduch, J., Rutherford, S., & Ruthven, O. (2009). *Portfolios of the Poor*. Princeton NJ: Princeton University Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

DDD. (2014). Doing Development Differently: The DDD Manifesto Community. Retrieved from http://doingdevelopmentdifferently.com/

Deprez, S., Huyghe, C., Van Gool Maldonado, C. (2012). *Using SenseMaker to Measure Learn and Communicate About Smallholder Farmer Inclusion*. Vredeseilanden/VECO.

Eyben, R. (2013). Uncovering the Politics of 'Evidence' and 'Results'. A Framing Paper for Development Practitioners. *Institute of Development Studies*. Retrieved from www.bigpushforward.net

Gentilini, U. (2016). Revisiting the "Cash versus Food" Debate: New Evidence for an Old Puzzle? *World Bank Research Observer*, 31(1), 135–167. https://doi.org/10.1093/wbro/lkv012

Gittelsohn, J., Pelto, P., Bentley, M. E., Bhattacharyya, K., & Jensen, J. L. (1998). *Rapid Assessment Procedures (RAP): ethnographic methods to investigate women's health*. United Nations University (UNU).

Given, L. M. (2008). *The SAGE Encyclopedia of Qualitative Research Methods*. Volume 2. L. M. Given (Ed.), SAGE Publications, Inc.

Graham, A., Powell, M., Taylor, N., Anderson, D. & Fitzgerald, R. (2013). *Ethical Research Involving Children*. Florence: UNICEF Office of Research – Innocenti.

Greene, J. (2009). Evidence and ‘‘proof” and evidence as ‘‘inkling” Chapter 8. In S. I. Donaldson, C. A. Christe, & M. M. Mark (Eds.), *What counts as credible evidence in applied research and evaluation practice?* (pp 153–167). London: Sage.

Guijt, I. (2016). Innovation in evaluation: Using SenseMaker to assess the inclusion of smallholder farmers in modern markets. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Handwerker, W. P. (2001). *Quick Ethnography: A Guide to Rapid Multi-Method Research*. Walnut Creek, CA: AltaMira Press.

Hasbullah, J. (2017). *Understanding Poverty and Poverty Data in Indonesia*. UNESCAP, (15), 2.

Itad. (2017). *Evaluation of The Hewlett Foundation's Strategy to Apply Human-Centered Design to Improve Family Planning and Reproductive Health Services in Sub-Saharan Africa.*

Jupp, D. (2007). Views of the Poor. *Participatory Learning and Action (PLA)*, 57.

Kidd, Steven and Emily Wylde (2011). *Targeting the Poorest: An Assessment of the Proxy Means Test Methodology. AusAID.*

Koleros, A., Jupp, D., Kirwan, S., et al. (2015). Methodological Considerations in Evaluating Long Term System Change: A Case Study from Eastern Nepal. *SAGE Publication.*

Leininger, M. M. (1985). Ethnography and Ethnonursing: Models and Modes of Qualitative Data Analysis. In M. M. Leininger (Ed.), *Qualitative research methods in nursing.* Orlando, FL: Grune & Stratton.

Lewis, D., Jupp, D., Arvidson, M. M., et.al. (2012). *Reality Check Reflection Report*. Sida.

Lewis, D. (2018). Peopling Policy Processes? Methodological Populism in the Bangladesh Health and Education Sectors. *World Development*, 108, 16–27. https://doi.org/10.1016/j.worlddev.2018.03.012

Mac an Ghaill, M. (1994). *The Making of Men: Masculinities, sexualities and schooling.* Buckingham: Open University Press.

Maina, R. N. (2017). *Gender–based violence: Assessing the applicability of the Participatory Ethnographic Evaluation Research methodology to enhance intervention Programmes.* (Dissertation is submitted in partial fulfilment of the requirements for the degree of MA Social Anthropology). School of Oriental and African Arts - University of London.

Masset, E., Jupp, D., Korboe, D., Dogbe, T., Barnett, C., Acharya, A., & Nelson, K. (2016). *Millennium Villages Evaluation: Midterm Summary Report*. DFID.

McCallum, L., Miller, J., Berry, S., & Hershey, C. (2016). Using social mapping techniques to guide programme redesign in the Tingim Laip HIV prevention and care project in Papua New Guinea. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Millen, D. (2000). Rapid Ethnography: Time Deepening Strategies for HCI Field Research. *In Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS* (pp. 280–286). https://doi.org/10.1145/347642.347763

Morse, J. M. (1994). Going in 'Blind'. *Qualitative Health Research*, 4(1), 3–5. https://doi.org/10.1177/104973239400400101

Morse, J. M. (2005). Insight, Inference, Evidence, and Verification: Creating a Legitimate Discipline Keynote Address for the II Congreso Iberoamericano de Investigación Cualitativa en Salud Madrid 22-25 de Junio de 2005. *International Journal of Qualitative Methods*, 5(1). Retrieved from https://sites.ualberta.ca/~iiqm/backissues/5_1/HTML/morse.htm

Mosse, D. & Kruckenberg, L. (2017). Beyond the Ivory Tower: Researching Development Practice. In Crawford, G., Kruckenberg, L., Loubere, N., & Morgan, R. *Understanding global development research* (pp. 193-212). 55 City Road, London: SAGE Publications Ltd. doi: 10.4135/9781473983236

Narayan, D., Chambers, R., Shah, M.K., & Petesch, P. (2000). *Voices of The Poor: Crying Out for Change*. Oxford University Press for The World Bank.

Nurbani, Rachma Indah (2015) "Food Price Volatility and the Worrying Trend of Children's Snacking in Indonesia" *IDS Bulletin* (46:6), November 2015.

OECD. & ADB. (2015). Education in Indonesia: Rising to the Challenge.

Office of Development Effectiveness (ODE). (2015). *Research for Better Aid: an Evaluation of DFAT's Investments*. Canberra: Department of Foreign Affairs and Trade.

Pain, A., Nycander, L., & Islam, K. (2014). *Evaluation of the Reality Check Approach in Bangladesh.*
Sida.

Pelto, G. H., Armar-Klemesu, M., Siekmann, J., & Schofield, D. (2013). The focused ethnographic study 'assessing the behavioral and local market environment for improving the diets of infants and young children 6 to 23 months old' and its use in three countries. *Maternal & Child Nutrition*, 9(S1), 35–46. https://doi.org/10.1111/j.1740-8709.2012.00451.x

Pink, S., & Morgan, J. (2013). Short-Term Ethnography: Intense Routes to Knowing. *Symbolic Interaction*, 36. https://doi.org/10.1002/symb.66

Poirrier, C. (2018). *Scoping Study: Evidence Translators' Role in Evidence-Informed Policymaking*. Results for Development.

Price, N., & Hawkins, K. (2002). *Researching Sexual and Reproductive Behavior: A Peer Ethnographic Approach.* University of Wales, UK: Centre for Development Studies.

Roger, P. (2017). Does evaluation need to be done differently to support adaptive management? [Blog].

Roller, M. R., & Lavrakas, P. J. (2015). *Applied Qualitative Research Design: A Total Quality Framework Approach (*1 edition). New York: The Guilford Press.

Ruby, J. (1980). Exposing yourself: Reflexivity, anthropology, and film. *Semiotica*, 30.

Seale, C., & Silverman, D. (1997). Ensuring rigour in qualitative research. *European Journal of Public Health*, 7(4), 379–384. https://doi.org/10.1093/eurpub/7.4.379

Seeberg, J., & Ray, R. K. (2016). Interpretation, context and time: An ethnographically inspired approach to strategy development for tuberculosis control in Odisha, India. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Shah, D. R. (2018). A Reality Check for Rapid Immersion in Development Research In search of rigour, ethics, and relevance Springfield Working Paper Series #5 (p. 40). *Springfield Centre.*

Spradley, J. P. (1980). *Participant Observation.* Holt, Rinehart and Winston.

Taplin, D. H., Scheld, S. & Low, S. M. (2002). Rapid ethnographic assessment in urban parks: A case study of Independence National Historical Park. *Human Organization,* 61(1). doi: 0018-7259/02/010080-14$1.90/1

Tierney, W. G., & Clemens, R. F. (2011). Qualitative Research and Public Policy: The Challenges of Relevance and Trustworthiness. In Higher Education: Handbook of Theory and Research (pp. 57–83). *Springer, Dordrecht*. https://doi.org/10.1007/978-94-007-0702-3_2

Tolley, E. (2017). *Traditional Socio-behavioral Research and Human Centered Design: Similarities, Unique Contributions and Synergies.* FHI 360.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review,* 80(5), 352–373. https://doi.org/10.1037/h0020071

TWP. (n.d.). Thinking and Working Politically Community of Practice: Discussion. Retrieved from https://twpcommunity.org/

USAID Learning Lab. (n.d.). Complexity-Aware Monitoring.

Walker, D., Samuels, F., Gathani, S., Stoelinga, D., & Deprez, S. (2014). *GirlHub Rwanda. 4,000 Voices- Stories of Rwandan Girls' Adolescence: A Nationally Representative Survey.* ODI.

Wild, L., Booth, D., & Valters, C. (2017). *Putting theory into practice.* London: ODI.

WFP. (2008). *Pilot Monitoring of High Food Price Impact at Household Level in Selected Vulnerable Areas.*

RCA Materials

RCA. (2008). *Bangladesh Reality Check Annual Report 2007.* Sida.

RCA. (2010). *Reality Check Bangladesh 2009 – Listening to Poor People's Realities about Primary Healthcare and Primary Education – Year 3.* Sida.

RCA. (2011). *Reality Check Bangladesh 2010 – Listening to Poor People's Realities about Primary Healthcare and Primary Education – Year 4.* Sida.

RCA. (2012). *Reality Check Bangladesh 2011 – Listening to Poor People's Realities about Primary Healthcare and Primary Education – Year 5.* Sida.

RCA. (2014). *Research Protocol SJ4P MEL Component.*

RCA. (2015a). *People Experiences of Security and Justice. A Reality Check Approach Report.* Palladium.

RCA. (2015b). *RAP Beneficiaries RCA. Pilot Study.*

RCA. (2015c). *RAP3 Nepal Rural Access Programme. Monitoring, Evaluation and Learning Component. Reality Check Approach Baseline Report.* Itad.

RCA. (2015d). *Reality Check Approach's Policy Influencing Potentials in Indonesia: A Reflection Note.*

RCA. (2016a) *Trainer's Guide Level 2.*

RCA. (2016b). *Adolescents and their Families Perspectives and Experiences on Nutrition and Physical Activities 'We Miss the Togetherness of the Past'.* UNICEF.

RCA. (2016c). *Perspectives of People Affected by Haze from Peatland and Forest Fires "Now no fire, no haze, but also no rice for people"*. UNICEF.

RCA. (2016d) *Trainer's Guide Level 1*.

RCA. (2016e) *RCA Study Design. Reality Check Approach Study to gather the perspectives of people affected by air pollution from peatland and forest fires*. UNICEF.

RCA. (2016f). *Study Brief: Perspectives of People Affected by Haze from Peatland and Forest Fires.*

RCA. (2016g). *RCA Study # 8: Insights into The Early Functioning of The Village Law from The Perspectives of Village Law from The Perspectives of Village Leaders and BPD.*

RCA. (2016h). RAP3 *DFID Nepal Rural Access Programme. Monitoring, Evaluation and Learning Component. Midline Report.* Itad.

RCA. (2016i). *RCA Study Brief: Household Finances: Insights from people living in poverty on their household finance management.*

RCA. (2016j). *"We never know how much cash we will earn today". Perspectives, Observations and Experiences of People Living in Poverty on Their Household Finance Management.* Indonesia.

RCA. (2017a). *RCA Transitional Report May 2017. RCA+ Project End of Phase 2 (2015-2017).*

RCA. (2017b). *Fieldwork Guide.*

RCA. (n.d.a). *Level 1 RCA Handout 9: Security Plan, Emergency Preparedness and Response.*

RCA. (n.d.b). *RCA Trainer's Guide Level 3.*

Annexes

Annex 1: Review Questions

Overarching RQ: To what extent, how and why has RCA been effective in generating and using poverty and social analysis in program and policy work?	
Framing RQs	**Sub-RQs**
What are the strengths and weaknesses of RCA relative to its objectives?	**What do stakeholders see as the primary (and secondary) features of RCA?**
	For what purposes is it intended and has it been used (e.g. to influence policy or management decisions)?
	How effective have they found it to be in addressing their objectives, either alone or in combination with other methods?
	What components/elements of RCA are seen as most valuable, for what uses, and why? In what contexts does RCA perform best?
	How robust is the RCA research methodology, research process (including information capture and synthesis), and research outputs?
	To what extent does it allow the host community or household to 'lead the agenda' or shape the interaction and findings?
	What does the host community identify as the positive impacts or impositions of the method, if any? Were there unexpected outcomes for participants or host communities?
	What is the perceived value of empathy?
How does RCA compare to other approaches to qualitative poverty and social assessment?	**What are the "hallmarks" of good qualitative and ethnographic research?**
	How does RCA "fit" with other qualitative design, monitoring and evaluation approaches? How does it work alongside a quantitative approaches to PSA?
	Does RCA address the short-comings of other approaches, or vice-versa?
	How does it compare to similar methodologies in terms of effectiveness and analytical power?
	What is the value for money of RCA compared to other approaches?
What is the potential relevance of RCA and similar poverty and social analysis approaches to the broader Australian aid program?	For what purposes is poverty and social assessment commissioned and in what ways is it used?
	Are RCA and other PSA quality assurance and review processes sufficient? Have ethics policies and practices been sufficient?
	What external and internal factors affect the uptake and use of poverty and social analysis?
	In what circumstances are RCA and other approaches to qualitative poverty and social analysis most effective?

Annex 2 – Bibliography

External Articles and Reports

ACFID. (2017). *Principles and Guidelines for ethical research and evaluation in development*.

Adolescents 360., IDEO., SFH., & PSI. (2016). *Emerging Insights for Design* [Powerpoint slides].

Adolescents 360., IDEO., & PSI. (2016). *A360 Emerging Insights for Design* [Powerpoint slides].

Adolescents 360., & PSI. (2016). *Adolescents 360 Formative Research Final Report: The Socio-cultural Drivers of Sexual and Reproductive Health for Adolescent Girls in Ethiopia*.

Adolescents 360., IDEO., & PSI. (n.d.). *Married Adolescents Insights. Insights for Design* [Powerpoint slides].

Addison, T., Hulme, D., & Kanbur, R. (2008). *Poverty Dynamics: Measurement and Understanding from an Interdisciplinary Perspective (Q-Squared Working Paper No. 52)*. Toronto: Q-Squared.

Anderson, M.B., Brown, D., & Jean, I. (2012). Time to Listen: Hearing People on the Receiving End of International Aid. US: CDA Collaborative Learning Projects.

Andrews, M., Pritchett, L., & Woolclock, M. (2012). Escaping Capability Traps Through Problem Driven Iterative Adaptation (CID Working Paper No. 240). Cambrige, Massachusetts: Center of International Development at Harvard University.

Andrews, M. (2013). *The Limits of Institutional Reform in Development: Changing Rules for Realistic Solutions.* Cambridge: Cambridge University Press.

Amelga, M. (1994). *A Review of Beneficiary Assessments Conducted by the Bank. ENVSP Consultant Report*. Washington D.C: The World Bank.

Arvidson, M. (2013). Ethics, Intimacy and Distance in Longitudinal, Qualitative Research: Experiences from Reality Check Bangladesh. *SAGE Publications.*

Asian Development Bank. (2012). *Handbook on Poverty and Social Analysis: A Working Document*. Asian Development Bank. Retrieved from https://www.adb.org/documents/handbook-poverty-and-social-analysis-working-document

Bell, S., & Aggleton, P. (2016). Interpretive and ethnographic perspectives. Monitoring and Evaluation in Health and Social Development. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Bernard, H. R. (2011). *Research Methods in Anthropology: Qualitative and Quantitative Approaches* (5th ed.). Lanham, MD: Alta Mira Pr.

Beta, P. (2015). *Design for Development & Humanitarian Innovation Playbook.* Wellington, NZ: Social Change Collective.

Bloom, P. (2010, June 4). Book Review - The Invisible Gorilla - By Christopher Chabris and Daniel Simons. *The New York Times*. Retrieved from https://www.nytimes.com/2010/06/06/books/review/Bloom-t.html

Brock, L. (2000). *An Introduction to Participatory Poverty Assessment*. UK: Institute of Development Studies.

Brown, E., Grellier, R., & Hawkins, K. (2016). The Use of the Rapid PEER Approach for the Evaluation of Sexual and Reproductive Health Programmes. In Stephen Bell and Peter Aggleton (Ed.) *Monitoring and Evaluation in Health and Social Development: Interpretive and Ethnographic Perspectives*. London and New York: Routledge.

BPS – Statistics Indonesia (2016) *Consumption Expenditure of Population of Indonesia by Province 2016, Based on the March 2016 Susenas*

Calder, R., & Huda, K. (2014). *Adolescent Girls Economic Opportunities Study, Rwanda*. Development Pathways.

CARE international. (2010). *Proposal for Participatory Ethnographic Evaluation and Research (PEER) Study Application for Ethical Approval.* Care International in Lao PDR.

Carr, E. (2018). Open the Echo Chamber: A Blog About Development and Global Change. Available at: http://www.edwardrcarr.com/opentheechochamber/

Casella, D., Magara, P., Kumasi, T.C., Gujit, I., & Van Soest, A. (2014). *The Triple-S Project SenseMaker Experience. A Method Tested and Rejected (Working paper 9)*. Triple-S.

Centre of Development Studies. (2004). *Ethics and the Use of Participatory Ethnographic Evaluation and Research (PEER)*. Options Consultancy Service Ltd, UK.

Cernea, M. M. (1992). Re-tooling in applied social investigation for development. Planning: Some methodological issues. In N. S. Scrimshaw & G. R. Gleason (Eds.), *Rapid Assessment Procedures - Qualitative Methodologies for Planning and Evaluation of Health Related Programmes*. Boston: International Nutrition Foundation for Developing Countries (INFDC). Retrieved from http://archive.unu.edu/unupress/food2/UIN08E/UIN08E00.HTM

CDA Collaborative Learning Projects. (2007). *Listening Project: Field Visit Report Kenya*. CDA Collaborative Learning Projects.

CDA Collaborative Learning Projects. (2007). *Listening Project: Report of the Listening Project Consultation.* CDA Collaborative Learning Projects.

Chambers, R. (1981). Rapid rural appraisal: Rationale and repertoire. *Public Administration and Development*, 1(2), 95–106. https://doi.org/10.1002/pad.4230010202

Chambers, R. (1994). Participatory rural appraisal (PRA): Analysis of experience. *World Development*, 22(9), 1253–1268. https://doi.org/10.1016/0305-750X(94)90003-5

Chambers, R. (1994). Participatory rural appraisal (PRA): Challenges, potentials and paradigm. *World Development*, 22(10), 1437–1454. https://doi.org/10.1016/0305-750X(94)90030-2

Chambers, R. (1994). The Origins and Practice of Participatory Rural Appraisal. *World Development*, 22(7), 953–969.https://doi.org/10.1016/0305-750X(94)90141-4

Chambers, R. (2006). *Poverty Unperceived: Traps, Biases and Agenda (Working Paper No. 270)*. UK: International Development Studies.

Chambers, R. (2007). Overview – Immersions: Something is happening. *Participatory Learning and Action*, 57, 9–14.

Chambers, R. (2009). Foreword. In N. Mukherjee (Ed.), *Speaking to Power: 27 Voice Tool* (pp. vii–viii). New Delhi: Concept Publishing.

Chambers, R. (2012, September 6). Robert Chambers - why don't all development organizations do immersions? [Blog]. Retrieved from https://oxfamblogs.org/fp2p/robert-chambers-why-dont-all-development-organizations-do-immersions/

Clear Horizon. (2009). *Quick Start Guide MSC Desig*n. Australia: Clear Horizon.

CMS communication. (2009). *Most Significant Change Stories: A Report.*

Collins, D., Morduch, J., Rutherford, S., & Ruthven, O. (2009). *Portfolios of the Poor*. Princeton NJ: Princeton University Press.

Cornwall, A., & Fujita, M. (2012). Ventriloquizing 'the Poor'? of Voices, Choices and the Politics of
'Participatory' Knowledge Production. *Third World Quarterly*, 33(9), 1751-1765. DOI: 10.1080/01436597.2012.721274. Routledge.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Dasandi, N., Marquette, H., & Robinson, M. (2016). *Thinking and Working Politically: From Theory Building to Building and Evidence Base (Research Paper No.37).* Birmingham, UK: Development Leadership Program.

Davies R., & Dart, J. (2005). *The Most Significant Change (MSC) Technique: A Guide to Its Use*.

DDD. (2014). Doing Development Differently: The DDD Manifesto Community. Retrieved from http://doingdevelopmentdifferently.com/

Deprez, S., Huyghe, C., Van Gool Maldonado, C. (2012). *Using SenseMaker to Measure Learn and Communicate About Smallholder Farmer Inclusion*. Vredeseilanden/VECO.

Doubek, J. (2016, April 17). Attention, Students: Put Your Laptops Away. *National Public Radio*. Retrieved from https://www.npr.org/2016/04/17/474525392/attention-students-put-your-laptops-away

Edmonds, R. (2016). Generating Local Knowledge: A Role for Ethnography in Evidence-Based Programme Design for Social Development. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Elmusharaf, K., Byrne, E., Manandhar, M., Hemmings, J., & O'Donovan, D. (2017). Participatory Ethnographic Evaluation and Research: Reflections on the Research Approach Used to Understand the Complexity of Maternal Health Issues in South Sudan. *Qualitative Health Research*, 27(9), 1345–1358. https://doi.org/10.1177/1049732316673975

Elmusharaf, K., Byrne, E., & O'Donovan, D. (2017). Social and traditional practices and their implications for family planning: a participatory ethnographic study in Renk, South Sudan. Reproductive Health, 14(1), 10. https://doi.org/10.1186/s12978-016-0273-2

Esselink, M. (2014, August). *The "Most Significant Change" Technique: A participatory method for M&E.* Presented at the National M&E Conference, Jakarta.

Eyben, R. (2013). Uncovering the Politics of 'Evidence' and 'Results'. A Framing Paper for Development Practitioners. *Institute of Development Studies*. Retrieved from www.bigpushforward.net

Fierro, R.S. (2014). *Using SenseMaker in Child-Centered Research*. Fierro Consulting LLC.

Gentilini, U. (2016). Revisiting the "Cash versus Food" Debate: New Evidence for an Old Puzzle? *World Bank Research Observer*, 31(1), 135–167. https://doi.org/10.1093/wbro/lkv012

Girl Hub. (2004). *Using SenseMaker to Understand Girls's Lives*.

Gittelsohn, J., Pelto, P., Bentley, M. E., Bhattacharyya, K., & Jensen, J. L. (1998). *Rapid Assessment Procedures (RAP): ethnographic methods to investigate women's health*. United Nations University (UNU).

Graham, A., Powell, M., Taylor, N., Anderson, D. & Fitzgerald, R. (2013). *Ethical Research Involving Children*. Florence: UNICEF Office of Research – Innocenti.

Greene, J. (2009). Evidence and ''proof" and evidence as ''inkling" Chapter 8. In S. I. Donaldson, C. A. Christe, & M. M. Mark (Eds.), *What counts as credible evidence in applied research and evaluation practice?* (pp 153–167). London: Sage.

Guijt, I., & Shah, M. K. (1998). *The Myth of Community: Gender Issues in Participatory Development*. London: Practical Action Publishers.

Guijt, I., & Hecklinger, J. (2010). Making Sense of Sense Maker: Evaluating Development Initiatives through Micro-Narrative Capture and Self-Tagging in Kenya [Powerpoint notes].

Guijt, I. (2016). Innovation in evaluation: Using SenseMaker to assess the inclusion of smallholder farmers in modern markets. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Handwerker, W. P. (2001). *Quick Ethnography: A Guide to Rapid Multi-Method Research*. Walnut Creek, CA: AltaMira Press.

Hanmer, L., & Hendrie, B. (2002). *Poverty and Social Impact Analysis (PSIA): common themes and issues arising from five pilot studies.* Prepared for the Joint World Bank/IMF/DFID PSIA Workshop, October 15-17, 2002. DFID.

Hasbullah, J. (2017). *Understanding Poverty and Poverty Data in Indonesia*. UNESCAP, (15), 2.

Heller, C. (2017). *Designing a Way to Measure the Impact of Design.* US: Design for Social Innovation.

Heltberg, R., Hossain, N., & Reva, A. (2012). *Living Through the Crisis: How the Food, Fuel and Financial Shocks Affect the Poor. New Frontiers of Social Policy.* Washington, DC: World Bank.

Higbie, M., Moore, M., & Scott, P. (2016*). Bringing Human-Centered Design to Teen Pregnancy Prevention.* US: PSI.

Hinds, R. (2013). *Tools for Participatory Analysis of Poverty, Social Exclusion and Vulnerability. Helpdesk Research Report.* Governance Social Development Humanitarian Conflict (GSDRC).

Hong, Y. A., Mitchell, S. G., Norman, K., Peterson, J. A., Latkin, C., & Tobin, K. (2016). Ethnographic process evaluation: A case study of an HIV prevention programme with injecting drug users in the USA. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Hurworth, R., Clark, E., Martin, J., & Thomsen, S. (2005). The Use of Photo- Interviewing: three Examples from Health Evaluation and Research. *Evaluation of Journal of Australasia,* Vol. 4 (new series).

IDEO. (2013). *MSI & IDEO.org. Research Outline* [Powerpoint slides].

IDEO. (2013, May). *New Opportunities for Youth and Reproductive Health* [Powerpoint slides].

IDEO. (2014, January). *MSZ-IDEO.org. Teens & Sexual Health in Zambia* [Powerpoint slides].

IDEO. (2015). *The Field Guide to Human-Centered Design.* US: IDEO.org

Indonesia Ministry of National Development Planning, & UNICEF. (2017). *SDG Baseline Report on Children in Indonesia.* Jakarta: BAPPENAS and UNICEF.

Institute for Research and Poverty. (2005). *Qualitative Approaches to the Study of Poverty and Welfare Reform: Current Challenges.* University of Wisconsin-Maddison.

Itad. (2013, April). *HCD in A360. How does it compare/differ to 'pure' HCD used in Hewlett?* [Powerpoint slides].

Itad. (2016). *Inception Report: Evaluation of the Hewlett-Packard Foundation's Strategy to Apply Human Centered-Design to Improve Family Planning and Reproductive Health Service in Sub-Saharan Africa.* Itad.

Itad. (2016). *Adolescents 360 Process Evaluation Inspiration Insight Deck.* Itad.

Itad. (2016). *RAP, KEPTA, MEL Programme Review of the Theory of Change* [Powerpoint slides].

Itad. (2016). *RAP, KEPTA, MEL Programme Review of the Theory of Change: Session 2* [Powerpoint slides].

Itad. (2017). *Adolescents 360 Process Evaluation. Inspiration Insight Deck* [Powerpoint slides].

Itad. (2017). *Evaluation of The Hewlett Foundation's Strategy to Apply Human-Centered Design to Improve Family Planning and Reproductive Health Services in Sub-Saharan Africa.*

Jackson, C. (2015). Case Study. Facilitating collaborative problem solving with human-centred design: the Making All Voices Count governance programme in 12 countries of Africa and Asia. *Knowledge Management for Development Journal* 11 (1): 91-106.

Jones, H., & Mendizabal, E. (2010). *Strengthening Learning from Research and Evaluation: going with the grain. Final Report*. London: IACDI.

Jupp, D. (2007). Views of the Poor. *Participatory Learning and Action (PLA)*, 57.

Kanbur, R., Chambers, R., Petesch, P., et.al. (2001). *Qualitative and Quantitative Poverty Appraisal: Complementarities, Tensions and the Way Forward.* Cornell University.

Kelly-Hanku, A. (2016). The Political Economy of Evidence: Personal Reflections on the Value of the Interpretive Tradition and its Methods. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Kidd, Steven and Emily Wylde (2011). *Targeting the Poorest: An Assessment of the Proxy Means Test Methodology. AusAID.*

Koleros, A., Jupp, D., Kirwan, S., et al. (2015). Methodological Considerations in Evaluating Long Term System Change: A Case Study from Eastern Nepal. *SAGE Publication.*

Laderchi, C.R. (2001). *Participatory Methods in the Analysis of Poverty: A Critical Review (Working Paper No. 62)*. QEH Working Paper Series.

Leavy, J., & Howard, J. (2013). *What Matters Most? Evidence from 84 Participatory Studies with Those Living with Extreme Poverty and Marginalization.* United Kingdom: Institute of Development Studies.

Leininger, M. M. (1985). Ethnography and Ethnonursing: Models and Modes of Qualitative Data Analysis. In M. M. Leininger (Ed.), *Qualitative research methods in nursing.* Orlando, FL: Grune & Stratton.

Lewis, D., Jupp, D., Arvidson, M. M., et.al. (2012). *Reality Check Reflection Report.* Sida.

Lewis, D. (2018). Peopling Policy Processes? Methodological Populism in the Bangladesh Health and Education Sectors. *World Development*, 108, 16–27. https://doi.org/10.1016/j.worlddev.2018.03.012

Longfield, K., Fortin, I., Wheeler, J., & Sievers, D. (2016). Permissions, Vacations and Periods of Self-reflection: Using Consumer Insight to Improve HIV Treatment Adherence in four Central American countries. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Maina, R. N. (2017). *Gender–based violence: Assessing the applicability of the Participatory Ethnographic Evaluation Research methodology to enhance intervention Programmes.* (Dissertation is submitted in partial fulfilment of the requirements for the degree of MA Social Anthropology). School of Oriental and African Arts - University of London.

Masset, E., Jupp, D., Korboe, D., Dogbe, T., Barnett, C., Acharya, A., & Nelson, K. (2016). *Millennium Villages Evaluation: Midterm Summary Report*. DFID.

McCallum, L., Miller, J., Berry, S., & Hershey, C. (2016). Using social mapping techniques to guide programme redesign in the Tingim Laip HIV prevention and care project in Papua New Guinea. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Millen, D. (2000). Rapid Ethnography: Time Deepening Strategies for HCI Field Research. *In Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS* (pp. 280–286). https://doi.org/10.1145/347642.347763

Mission Design Group. (2016). *Reinventing How the World's Toughest Challenges are Solved. Mission Design Group.*

Morse, J. M. (1994). Going in 'Blind'. *Qualitative Health Research*, 4(1), 3–5. https://doi.org/10.1177/104973239400400101

Morse, J. M. (2005). Insight, Inference, Evidence, and Verification: Creating a Legitimate Discipline Keynote Address for the II Congreso Iberoamericano de Investigación Cualitativa en Salud Madrid 22-25 de Junio de 2005. *International Journal of Qualitative Methods*, 5(1). Retrieved from https://sites.ualberta.ca/~iiqm/backissues/5_1/HTML/morse.htm

Mosse, D. & Kruckenberg, L. (2017). Beyond the Ivory Tower: Researching Development Practice. In Crawford, G., Kruckenberg, L., Loubere, N., & Morgan, R. *Understanding global development research* (pp. 193-212). 55 City Road, London: SAGE Publications Ltd. doi: 10.4135/9781473983236

Mulgan, G. (2014). *Design in Public and Social Innovation. What Works and What Could Work Better.* Nesta.

Musoke, D., Ekirapa-Kiracho, E., Ndejjo, R., & George, A. (2015). Using photovoice to examine community level barriers affecting maternal health in rural Wakiso district, Uganda. *Reproductive Health Matters*, 23(45), 136–147. https://doi.org/10.1016/j.rhm.2015.06.011

Narayan, D., Chambers, R., Shah, M.K., & Petesch, P. (2000). *Voices of The Poor: Crying Out for Change.* Oxford University Press for The World Bank.

Norton, A., Bella B., Brock, K., Kakande, M., & Turk, C. (2001). *A Rough Guide to PPAs: Participatory Poverty Assessment an Introduction to Theory and Practice.* UK: Overseas Development Institute.

Nurbani, Rachma Indah (2015) "Food Price Volatility and the Worrying Trend of Children's Snacking in Indonesia" *IDS Bulletin* (46:6), November 2015.

Overseas Development Institute (ODI). (2015). *Question guide: researching norms about early marriage and girls' education.* London.

Overseas Development Institute (ODI). (2016). *Gender and Adolescence Global Evidence (GAGE) Baseline Qualitative Instruments and Theory of Change Paper.*

Overseas Development Institute (ODI). (2016). Gender and Adolescence Global Evidence (GAGE) Learning and Capacity Strengthening, Chapter 5.

OECD. & ADB. (2015). Education in Indonesia: Rising to the Challenge.

Office of Development Effectiveness. (2015). *Research for Better Aid: an Evaluation of DFAT's Investments.* Canberra: Department of Foreign Affairs and Trade.

Oguntoye, S., Otoo-Oyortey, N., Hemmings, J., Norman, K., & Hussein, E. (2009). *FGM is With Us Everyday Women and Girls Speak Out About Female Genital Mutilation in the UK.* World Academy of Science, Engineering and Technology.

Otoo-Oyortey, N., Gezahegn King, E., & Norman, K. (2016). Designing Health and Leadership Programmes for Vulnerable Young Women Using Participatory Ethnographic Research in Freetown, Sierra Leone. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Options. (2004). Participatory Ethnographic Evaluation and Research (PEER) [Flyer]. Options Centre for Development Studies Swansea.

Options, CARE Myanmar, & Sex Workers in Myanmar Network (SWiM). (2005). *Living on the Edge Sex workers' lives in Myanmar*

.

Oxfam Novib. (2012). *Learning About Gender Equality. Testing the Ability of the Most Significant Change Methodology to Make Cultural changes Visible and Learn about Gender Equality.* The Netherlands: Oxfam Novib.

Pain, A., Nycander, L., & Islam, K. (2014). *Evaluation of the Reality Check Approach in Bangladesh.* Sida.

Palladium International Ptd Ltd (n.d.). *Child Protection Code of Conduct.*

Paull, M., Boudville, I., & Sitlington, H. (2013). *Using Sense making as a Diagnostic Tool in the Analysis of Qualitative Data (The Qualitative Report Volume 18, No. 27). TQR.*

Peace Development Fund. (1999). *The Listening Project: A National Dialogue on Progressive Movement Building: Problems, Prospects, Potentials.* US: Peace Development Fund.

PEER. (n.d.). *Follow up Information on PEER and Rapid PEER.*

Pelto, G. H., Armar-Klemesu, M., Siekmann, J., & Schofield, D. (2013). The focused ethnographic study 'assessing the behavioral and local market environment for improving the diets of infants and young children 6 to 23 months old' and its use in three countries. *Maternal & Child Nutrition,* 9(S1), 35–46. https://doi.org/10.1111/j.1740-8709.2012.00451.x

Pink, S., & Morgan, J. (2013). Short-Term Ethnography: Intense Routes to Knowing. *Symbolic Interaction*, 36. https://doi.org/10.1002/symb.66

Poirrier, C. (2018). *Scoping Study: Evidence Translators' Role in Evidence-Informed Policymaking*. Results for Development.

Potter, S., Khan, S., Watson, J., & Pederson, S. (n.d.). *Evaluations that Make a Difference.*

Price, N., & Hawkins, K. (2002). *Researching Sexual and Reproductive Behavior: A Peer Ethnographic Approach.* University of Wales, UK: Centre for Development Studies.

PSI. (2016). *A360 Innovation: Tanzania Human-Centered Design Concept Presentation* [Powerpoint slides]. PSI.

Ramalingam, B. (2016, February 19). What's Next in Design for Development? *Institute of Development Studies.* Retrieved from http://www.ids.ac.uk/opinion/what-s-next-in-design-for-development

Robb, C.M. (1999). *Can the Poor Influence the Policy? Participatory Poverty Assessment in the Developing World 2nd Edition.* World Bank.

Robb, C.M. (2003). *Poverty and Social Impact Analysis – Linking Macroeconomics Policies to Poverty Outcomes: Summary of Early Experiences*. IMF.

Roger, P. (2017). Does evaluation need to be done differently to support adaptive management? [Blog].

Roller, M. R., & Lavrakas, P. J. (2015). *Applied Qualitative Research Design: A Total Quality Framework Approach (*1 edition). New York: The Guilford Press.

RTI International. (2017). *Draft: Six Monthly Report. 1 July – 31 December 2016.*

Ruby, J. (1980). Exposing yourself: Reflexivity, anthropology, and film. *Semiotica*, 30.

Salmen, L.F. (2002). *Beneficiary Assessment: An Approach Described (Paper No.10*). Washington D.C: The World Bank.

Saville E., & Brown, E. (2005). *Living on The Edge Sex Workers' Live in Myanmar*. Options Consultancy Service Ltd, UK.

Schlingheider, A., Pellfolk, E., Maneo, G., & Desai, H. (2017). *Managing to Adapt: Analyzing Adaptive Management for Planning, Monitoring, Evaluation and Learning.* London, UK: The London School of Economics and Political Science.

Seale, C., & Silverman, D. (1997). Ensuring rigour in qualitative research. *European Journal of Public Health*, 7(4), 379–384. https://doi.org/10.1093/eurpub/7.4.379

Secure Nutrition. (2013). *Integrated Nutrition Project (Suaahara-Building Strong and Smart Families).* Retrieved from http://www.securenutrition.org/resources/integrated-nutrition-project-suaahara-building-strong-and-smart-families

Seeberg, J., & Ray, R. K. (2016). Interpretation, context and time: An ethnographically inspired approach to strategy development for tuberculosis control in Odisha, India. In S. Bell & P. Aggleton (Eds.), *Monitoring and Evaluation in Health and Social Development Interpretive and Ethnographic Perspectives* (1st Edition). London: Routledge.

Seelos, C., & Mair, J. (2016). When Innovation Goes Wrong (SSIR). *Stanford Social Innovation Review.* Retrieved from https://ssir.org/articles/entry/when_innovation_goes_wrong

Shah, D. R. (2018). A Reality Check for Rapid Immersion in Development Research In search of rigour, ethics, and relevance Springfield Working Paper Series #5 (p. 40). *Springfield Centre.*

Shutt, C., & Ruedin, L. (2013). *SDC How to Note Beneficiary Assessment.* Switzerland: Swiss Agency for Development and Cooperation.

Sjogren, S., & Scharbatke-Church, C. (2017, September 21). Reflections on Using Most Significant Change in An Anti-Corruption Program [Blog]. Retrieved from https://www.cdacollaborative.org/blog/reflections-using-significant-change-anti-corruption-program/

Spradley, J. P. (1980). *Participant Observation.* Holt, Rinehart and Winston.

Taplin, D. H., Scheld, S. & Low, S. M. (2002). Rapid ethnographic assessment in urban parks: A case study of Independence National Historical Park. *Human Organization*, 61(1). doi: 0018-7259/02/010080-14$1.90/1

Taylor, L., & Ngoc-Thuy, N. (2010). *CARE Mid Term Review of Legal Awareness & Life Skills for*
Women Project. CARE Laos.

Tierney, W. G., & Clemens, R. F. (2011). Qualitative Research and Public Policy: The Challenges of Relevance and Trustworthiness. In Higher Education: Handbook of Theory and Research (pp. 57–83). *Springer, Dordrecht*. https://doi.org/10.1007/978-94-007-0702-3_2

Tolley, E. (2017). *Traditional Socio-behavioral Research and Human Centered Design: Similarities, Unique Contributions and Synergies.* FHI 360.

Tsui, J., Hearn, S., & Young, J. (2014). *Monitoring and Evaluation of Policy Influence and Advocacy* (Working Paper No. 395). ODI.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 352–373. https://doi.org/10.1037/h0020071

TWP. (n.d.). Thinking and Working Politically Community of Practice: Discussion. Retrieved from https://twpcommunity.org/

USAID Learning Lab. (n.d.). Complexity-Aware Monitoring.

Walker, D., Samuels, F., Gathani, S., Stoelinga, D., & Deprez, S. (2014). *GirlHub Rwanda. 4,000 Voices- Stories of Rwandan Girls' Adolescence: A Nationally Representative Survey.* ODI.

Walovitch, J. (n.d.). Time to Listen: A Discussion Guide - CDA. Retrieved 31 July 2018, from https://www.cdacollaborative.org/publication/time-to-listen-a-discussion-guide/

Wild, L., Booth, D., & Valters, C. (2017). *Putting theory into practice.* London: ODI.

Wiles, R., Crow, G., & Pain, H. (2011). *Innovation in Qualitative Methods: A Narrative Review.* UK:
Sage Publications.

Wisor, S., Bessell, S., Castillo, F., Crawford, J., et al. (2016). *The Individual Deprivation Measure: A*
Gender-Sensitive Approach to Poverty Measurement. Melbourne, Australia: International Women's Development Agency Inc.

World Bank. Poverty Reduction Group. (2003). *A User's Guide to Poverty and Social Impact Analysis.*

WFP. (2008). *Pilot Monitoring of High Food Price Impact at Household Level in Selected Vulnerable Areas.*

Young, J., Davis, T., Sherlock, S., Pasanen, T., & Shaxson, L. (2015). Research for Better Aid: An Evaluation of DFAT's Investments. Canberra: Department of Foreign Affairs and Trade.

RCA Reports

RCA. (2008). *Bangladesh Reality Check Annual Report 2007.* Sida.

RCA. (2010). *Reality Check Bangladesh 2009 – Listening to Poor People's Realities about Primary Healthcare and Primary Education – Year 3.* Sida.

RCA. (2011). *Reality Check Bangladesh 2010 – Listening to Poor People's Realities about Primary Healthcare and Primary Education – Year 4.* Sida.

RCA. (2012). *Reality Check Bangladesh 2011 – Listening to Poor People's Realities about Primary Healthcare and Primary Education – Year 5.* Sida.

RCA. (2013). *Notes from the RCA Scoping Study.* Itad.

RCA. (2014). *Education Study in Tanah Papua.* Indonesia.

RCA. (2014). *Reality Check Approach Baseline Report. Contribution to the Monitoring, Evaluation and Learning Component supporting RAP3 Mid and Far West Nepal.* Itad.

RCA. (2014). *Research Protocol SJ4P MEL Component.*

RCA. (2014). *Understanding Poverty from the Perspectives of People Living in Poverty.*

RCA. (2014). *Understanding Social Assistance Programs from the Perspectives of People Living in Poverty. Indonesia.*

RCA. (2015). *Experiences and Perspectives of Direct Beneficiaries RCA Study Undertaken for the Rural Access Programme 3 Monitoring, Evaluation and Learning Component Mid and Far West Nepal.* Itad.

RCA. (2015). *People Experiences of Security and Justice. A Reality Check Approach Report*. Palladium.

RCA. (2015). *People Views and Experiences of the National Social Assistance Programs.* Indonesia.

RCA. (2015). *Perspectives of People Living in Poverty in Nepal: A background paper for the Mid-term review of Swiss Country Strategy (2013-2017).*

RCA. (2015). *RAP Beneficiaries RCA. Pilot Study.*

RCA. (2015). *RAP3 Nepal Rural Access Programme. Monitoring, Evaluation and Learning Component. Baseline Report.* Itad.

RCA. (2015). *RAP3 Nepal Rural Access Programme. Monitoring, Evaluation and Learning Component. Reality Check Approach Baseline Report*. Itad.

RCA. (2015). *Understanding Poverty from the Perspectives of People Living in Poverty.*

RCA. (2015). *Understanding Social Assistance Programs from the Perspectives of People Living in Poverty: Indonesia. Reality Check Approach Sub-report 1.* Indonesia.

RCA. (2015). *Understanding Social Assistance Programs from the Perspectives of People Living in Poverty. Reality Check Approach Sub-report 2.* Indonesia.

RCA. (2015). *"We Are Healthy, Why Change?". Perspectives, Observations, Experiences of People Living in Poverty on Their Hygiene and Nutrition.* Indonesia.

RCA. (2016). *Adolescents and their Families Perspectives and Experiences on Nutrition and Physical Activities 'We Miss the Togetherness of the Past'.* UNICEF.

RCA. (2016). *DFID Nepal Rural Access Programme. Monitoring, Evaluation and Learning Component. Midline Impact Assessment Report.*

RCA. (2016). *DFID Nepal Rural Access Programme. Monitoring, Evaluation and Learning Component. Midline Report.* Itad.

RCA. (2016). *Monitoring, Evaluation and Learning (MEL), Rural Access Programme 3 (RAP-3). Review of the RAP-3 Theory of Change.* RAP-3 KEPTA UKAID.

RCA. (2016). *People's Experiences and Perspectives on Recovery from the 2015 Earthquakes in Nepal*. Kathmandu: Palladium.

RCA. (2016). *Perspectives of People Affected by Haze from Peatland and Forest Fires "Now no fire, no haze, but also no rice for people"*. UNICEF.

RCA. (2016). *RCA Study # 8: Insights into The Early Functioning of The Village Law from The Perspectives of Village Law from The Perspectives of Village Leaders and BPD.*

RCA. (2016). *RCA Study Design #10. Gathering the Perspectives of Adolescent Girls and Boys and Their Influencers on Eating, Drinking and Physical Activity Behaviors.*

RCA. (2016) *RCA Study Design. Reality Check Approach Study to gather the perspectives of people affected by air pollution from peatland and forest fires*. UNICEF.

RCA. (2016). *Reality Check Approach Report: Local Perspectives and Experiences of the Village Law in Indonesia.* KOMPAK.

RCA. (2016). *Reality Check Approach Study Proposal: Perspectives and Experience Families Living in Poverty on Their Finances.*

RCA. (2016). *Reality Check Approach Study Brief. Household Finances. Insights from People Living in Poverty on Their Household Finance Management.*

RCA. (2016). *Research Brief Reality Check Approach: Early Implementation of The Village Law as Perceived by Village Officials and Local Communities*. KOMPAK.

RCA. (2016). *Study Brief: Perspectives of People Affected by Haze from Peatland and Forest Fires.*

RCA. (2016). *'We are learning by working' Perspectives, Experiences and Motivations of Youth on Employment and Vocational Training in Bangladesh.* Palladium.

RCA. (2016). *"We never know how much cash we will earn today". Perspectives, Observations and Experiences of People Living in Poverty on Their Household Finance Management.* Indonesia.

RCA. (2017). *Children and Their Families Perspectives and Experiences on Poverty and Social Protection.*

RCA. (2017). *Choice and Voice: People's Perspectives on Health Service in Pakistan*. Jakarta: Palladium.

Other RCA Resources

RCA. (2014). *Activity Note: Promoting Use and Uptake of the Reality Check Approach.*

RCA. (2015). *Education Study in Tanah Papua*.

RCA. (2015). *Prep Package Village Leader Study #8.* Indonesia.

RCA. (2015). *Reality Check Approach's Policy Influencing Potentials in Indonesia: A Reflection Note.*

RCA. (2015). *RCA Study: The Education Stream Draft Concept Paper.*

RCA. (2015). *RCA Study Brief: People's View and Experience of Education in Papua.*

RCA. (2015). *RCA Study Brief: People Views and Experiences of the National Social Assistance Programs.*

RCA. (2015). *RCA Study Brief: International Migration for Work: migrants contrasting perspectives*.

RCA. (2015). *RCA Study Brief: Perspectives, Observations, Experiences of People Living in Poverty on Their Hygiene and Nutrition.*

RCA. (2015). *RCA Study Brief: Perspectives and Experiences of Frontline Health Service Providers.*

RCA. (2016). *16-JAK2: Entering the community.*

RCA. (2016). *EVA - BHN RCA Concept Note*. Palladium.

RCA. (2016). *RAP, KEPTA, MEL Programme Review of Theory of Change.*

RCA. (2016). *RAP, KEPTA, MEL Programme Review of Theory of Change Session 2.*

RCA. (2016). *RCA Study Brief: Household Finances: Insights from people living in poverty on their household finance management.*

RCA. (2016) *Trainer's Guide Level 1.*

RCA. (2016) *Trainer's Guide Level 2.*

RCA. (2017). *Child Poverty and Social Protection.*

RCA. (2017). *Child Poverty and Social Protection* [PowerPoint slides].

RCA. (2017). *Embedding RCA supply and demand: Draft log frame vs 3 (2015 – 2017).*

RCA+. (2017). *Fieldwork Guide.*

RCA. (2017). *Findings to Report: How to Translate Findings into Public Documents.*

RCA+. (2017). *Operation Guidelines: RCA Study Personnel Input Guidelines.*

RCA. (11.17). *RCA Studies Financial Information.*

RCA+. (2017). *RCA Transitional Report May 2017. RCA+ Project End of Phase 2 (2015-2017).*

RCA. (n.d.). *Briefing Study 8v1.* Indonesia.

RCA. (n.d.). *Briefing Study 8.* Indonesia.

RCA. (n.d.). *Draft Risk & Mitigation Matrix of Study #8: Village Leader.* Indonesia.

RCA. (n.d.). *Draft Risk & Mitigation Matrix of Study #11: Understanding Air Pollution Impact from Peatland and Forest Fires.*

RCA. (n.d.). *Level 1 RCA Handout 9: Security Plan, Emergency Preparedness and Response.*

RCA. (n.d.). *Level 2 Training for Sub-Team Leader.*

RCA. (n.d.). *Level 3 Exercise: Reviewing Box Stories.*

RCA. (n.d.). *Level 3 Training: Finding Meaning from De-briefing Notes.*

RCA. (n.d.). *Level 3 Training: Guidance Notes for Managing Preliminary Findings.*

RCA. (n.d.). *Mixed Method Training: Groupwork.*

RCA. (n.d.). *Mixed Method Training: Handouts.*

RCA+. (n.d.). *Operation Guidelines: A Guidance on finding the 'right' people for RCA.*

RCA+. (n.d.). *Operation Guidelines: Archiving Guidance.*

RCA+. (n.d.). *Operation Guidelines: Draft Risk and Mitigation Matrix of Study.*

RCA+. (n.d.). *Operation Guidelines: Feedback on RCA study team member.*

RCA. (n.d.). *Operating Guidelines for RCA Local Support Partners.*

RCA+. (n.d.). *Operation Guidelines: Guidelines for the RCA Study Reference Group.*

RCA+. (n.d.). *Operation Guidelines: Level 1: Handout 19; Child Protection Policy.*

RCA+. (n.d.). *Operation Guidelines: Photo Consent Form.*

RCA+. (n.d.). *Operation Guidelines: Report template.*

RCA+. (n.d.). *Operation Guidelines: Study design template RCA studies.*

RCA+. (n.d.). *Operation Guidelines: Template ToR for commissioning RCA studies.*

RCA+. (n.d.). *Project Confidentiality and Research Data Protections.*

RCA. (n.d.). *RCA Study #8: Insights into the early functioning of the Village Law from the perspectives of village leaders and BPD.*

RCA. (n.d.). *Researcher Recruitment and Assessment Guide.*

RCA. (n.d.). *Team Member Evaluating English.*

RCA+. (n.d.). *Training Materials: Handout 1: Why research?*

RCA. (n.d.). *Groupwork instructions: Survey Data Collection Tools.*

RCA. (n.d.). *Village Head Relations.*

Testimonies from various people on RCA
https://www.youtube.com/watch?v=8Oz7d6Sv5QE

RCA for Policy Influence by Drew Koleros, Palladium
https://www.youtube.com/watch?v=V6yRkK66avY

Origins of RCA by Dee Jupp
https://www.youtube.com/watch?v=zgurzMEL0Ho

Rigor in RCA by Deep Jupp
http://www.reality-check-approach.com/related-resources.html

Annex 3 - RCA Case Studies: Aims and Objectives[66]

The Integrated Programme for Strengthening Security and Justice (IP-SSJ), DFID, Nepal

The Integrated Programme for Strengthening Security and Justice (IP-SSJ) is a DFID-funded, five-year program aiming to provide "Improved security and access to justice, particularly for traditionally excluded people across Nepal". The program aims to contribute primarily to three results areas, including:

- Change to social barriers that impede access to justice
- Increased engagement among communities and security and justice institutions for better prevention of crime and insecurity
- Improved responses to criminal violence through higher capabilities and performance of security and justice institutions"

Palladium leads a consortium of Nepali institutions in the Monitoring, Evaluation and Learning (MEL) component of IP-SSJ. The aim of the MEL component is to work with IP-SSJ implementers to provide high quality evidence to feed back into project implementation as part of a 'structured experimental learning' approach. This component aims to improve the quality and availability of evidence related to security and justice based on the specific data needs and decision timelines facing the relevant Ministries and decision-making bodies. It employs a "multiple method", rather than "mixed-method", approach.

An RCA study was conducted at baseline, and was to be the first of 3 RCA studies for this program, the other two being planned for mid-line (2016) and end-line (2018). It should be noted that neither of these other studies have yet been conducted, and RCA core team key informants for this case study were uncertain as to whether the RCA would be used longitudinally.

The primary motivation of the baseline study to which the RCA contributed is to improve the evidence base on how people experience and understand security and justice (S&J) in Nepal, with particular reference to poor and marginalized groups. As such, the RCA intended to fill a number of enduring evidence gaps related to security and justice in Nepal which remain despite the growing attention of both researchers and program implementers to this topic. Main gaps include: 1. little is known about how people experience violence and insecurity on a daily basis. Dominant understandings of 'violence' and 'crime' may not conform to how the individuals most likely to experience these events understand these issues, and require further understanding to accurately assess. 2. Second, data is particularly lacking with relation to people's experience and access to formal security and justice institutions, particularly the Nepal Police and Court system. True dynamics behind both crime and reporting remain largely unclear. As such, an in-depth qualitative study is particularly important to ensure that the baseline study for this program engages with the most pressing yet least understood factors affecting these issues. Framing these security and justice issues in ways consistent with how they are understood by communities will also be an essential factor for success behind this study.

Haze, UNICEF, Indonesia

This study is a part of a series of qualitative studies planned under the first year of UNICEF and the Government of Indonesia's (GOI) new Country Program (CPAP 2016-2020). This Reality Check Approach study aimed to complement secondary source review and provide an understanding of the

[66] Information from this Annex comes from official RCA documentation, supplemented by interviews with the RCA core team and others connected closely with the particular studies.

problem from the perspectives of those people affected by the haze. In addition to this, the study aimed to provide insights on the current land clearing practices, risk awareness and coping strategies and accountability measures for the people, in particular children, in the haze-affected areas. These insights were expected to inform the UNICEF and related GOI bodies including the Peatland Restoration Agency (Badan Restorasi Gambut, BRG) in order to influence ongoing policy dialogue for a community-based programming and resource allocation for health, environment and hazard risk management.

The study was also a component of a longer term partnership between UNICEF, RCA, UN Pulse Lab and UNOPs to reduce the risks of air pollution from wild fires to children's health and wellbeing. This study, in particular, was designed to understand the impact of peatland wildfires on the health, education, wellbeing and family situation of Indonesian children. This study aimed to provide a nuanced understanding of the complex social problems behind the haze crisis, specifically to understand:

1. Risk awareness and knowledge of air pollution and its health, education, and psychosocial impacts
2. Coping strategies of the community regarding health and education impacts of haze
3. Concepts of accountability in terms of the child rights (duty bearer and rights holder)
4. Real and perceived capacity to affect change

Village Law, DFAT, Indonesia

This Reality Check Approach (RCA) study was responding to the Government of Indonesia's need for evidence on the progress of implementation of the Village Law. The study's intention was to provide insights into current perceptions and experience of the implementation of Village Law from the perspective of village leaders and the Badan Permusyawaratan Desa (BPD). The concept of this study was first discussed during 2014 with DFAT and others - including TNP2K and World Bank - and was followed up in August 2015 with a series of meetings with various stakeholders including Bappenas, KOMPAK and the World Bank. The following issues were highlighted as potentially benefitting from further exploration through the RCA study:

- Context; village size, history, location, socio and political economy
- Who are village leaders?: their profiles, education levels, livelihoods, motivations for public service, affiliations and connections, basis of selection/election, their perceived roles and responsibilities, local relationships (with others in village government, BPD , service providers, community), accountability and support.
- Specific to implementation of the Village Law:
    - Understanding/perception of the intentions of Village Law at village level.
    - The changing role and responsibilities of Village Leaders under the mandates of the Village Law
    - The changing role and responsibilities of the BPD under the mandates of the Village Law
    - Preparedness of Village Leaders and BPD to manage these new responsibilities. Skills and information. Capacity and confidence. Absorptive capacity. Role and relationship of Facilitators. Experience, appropriateness and adequacy of cascade training and other capacity support.
    - Early experience of prioritizing, planning and use of new fund flows to villages.
    - Incentives and disincentives
    - Community participation and engagement throughout the planning and implementation cycle and BPD and Village leaders role in this.
    - Social accountability

This RCA study intended to explore these issues primarily from the perspectives of Village Leaders and members of the BPD by living with them and their families over several days and nights. The study will also include living with ordinary villagers and interacting with other community members

to understand the dynamics between the community and community leadership. In this way it is hoped that consideration will be given to context and relevance in designing and providing further support to the roll out of the Village Law, in particular to inform Ministry level discussions on plans for 2016 capacity support, targeting of training as well as further 'socialization' on the Village Law.

Adolescent Nutrition and Physical Activity, UNICEF, Indonesia

UNICEF has commissioned the Reality Check Approach (RCA) team to gather adolescent girls' and boys' perspectives on factors influencing their choices and behaviors in eating, drinking and physical activity. The was done in order to contribute to a new program to support the Government of Indonesia (GoI) to test public health interventions and identify policy options supporting good nutrition and healthy lifestyle.

This RCA study is part of a series of baseline survey and assessments that will be conducted in the two districts to inform the design of policy and program approaches, and to provide baseline information against which to assess change, though it is not part of a "mixed method" approach, but rather a "multiple method" one. The aim of the study was not only to further the insights gained from a recent formative research conducted by GAIN (Global Alliance for Improved Nutrition) in East Java focusing on adolescent girls and young women aged 15-25 years, prior to and after marriage, but to expand the scope of study to include differences between genders (sic), different geographical contexts (coastal with inland, rural with peri-urban) and physical activity as an equally important focus of the study.

The intention was that the findings of this study would be used to provide contextual understanding about the lives of adolescent girls and boys to UNICEF's nutrition-related teams and programs and to support the development of an approach to reaching the adolescent cohort to promote healthy eating habits and lifestyle. The RCA study findings and insights were also intended to be used in the design of the quantitative study UNICEF also commissioned under the same thematic focus and to provide an interpretive lens for quantitative data gathered post execution.

Household Financial Management, DFAT, Indonesia

This study was commissioned by the DFAT/RTI Knowledge Sector Initiative at the request of Bappenas in order to complement their understanding of household financing and comping mechanism and the role of cash transfers, loans and other programs designed to support disadvantaged families. The findings were intended to provide inputs into GoI's financial inclusion initiatives currently being designed and implemented, and fill a massive gap in knowledge concerning: 1. The role of practices such as informal borrowing and pawning, seasonal risks and seasonal demands for cash and how people manage with variable and multiple incomes; and 2. how social protection cash transfers are absorbed and used by households, including how people use new infusions of cash (immediate spend, to pay debt, so-called temptation spending', saving etc.).

It was hoped that the RCA study would assist in optimize the usefulness for social assistance by providing findings relevant to, for example, timing of payments, appropriate cash amounts and appropriate channels to increase the likelihood that cash would be used for intended purposes. The RCA also aimed to provide more understanding around people's willingness and ability to pay e.g. for utilities, health insurance..

Rural Access Programme (RAP3), DFID, Nepal

RAP3 Monitoring, Evaluation and Learning is led Palladium. The MEL contract is to deliver a monitoring and evaluation system that is primarily geared towards continuous learning and programme improvement, but will also ensure accountability on programme delivery. Implemented over the lifetime of the programme, the twin approaches of monitoring and an end line evaluation are to be supplemented by a number of thematic studies, starting with the RCA. RAP RCA is described as a longitudinal evaluation as a means to providing a theory of change based on people's perspectives and insights to better inform the design of survey instruments.

RCA involvement in RAP3 has included a number of studies. A scoping RCA study was conducted in December 2013 and used to help inform the design of the household survey. This was followed by a baseline RCA study in 2014 and a midline RCA study in 2016. The midline RCA study sought to better understand what changes had happened in the interim two year period since the baseline, how people living in communities in the project areas see change, and how they generally perceive the relevance of the RAP interventions. This was a mixed method exercise, with joint analysis being conducted by the RCA team and quantitative survey analysts from Statistics for Sustainable Development.

Millennium Villages Evaluation, DFID, Ghana

The RCA is part of a 3rd party mixed method MEL programme, similar to RAP, being led by Itad. The purpose of the RCA is to ground truth, qualify and provide an interpretive lens on the findings of the quantitative survey, in particular to increase understandings of people's attitudes and behaviors.

The Bangladesh RCA, Sida, Bangladesh

The first RCA was established by the Swedish Embassy in Dhaka and by Sida headquarters in Stockholm. It was conducted was in Bangladesh starting in 2007 as a 5 year project, with the same researchers returning to the same households every year over the course of 5 years. The aim was to gain a “fleeting glimpse” (KII, Qualitative Research Expert) into how people were accessing and experiencing health and education services, to supplement formal sectoral monitoring and evaluation data (used in both SWaPs) with “people-centred data” and thus to close the gap a little between people and policy makers by illuminating the lived experiences of poor people.

Annex 4 - Key Informants[67]

	Name	Position/Role	Organization	Relationship to RCA
1	01 Anonymous	Researcher	Reality Check Approach	Internal
2	02 Anonymous	Researcher	Reality Check Approach	Internal
3	Abisaputra, Iqbal	Researcher	Reality Check Approach	Internal
4	Adhima, Rizqan	Researcher	Reality Check Approach	Internal
5	Andrian, Lody	Lead Researcher	UN Pulse Lab	Connected
6	Antlov, Hans	Technical Adviser	Knowledge Sector Initiative	Connected
7	Ario Bismo, Pandu	Researcher	Reality Check Approach	Internal
8	Ayuandini, Sherria	Senior Lead Researcher for Urban Poverty Study	Reality Check Approach	Internal
9	Barahona, Carlos	Managing Director	Statistics for Sustainable Development	Connected
10	Barnett, Chris	Director, Technical Excellence	Itad	Connected
11	Bayley, Scott	Principal Specialist, Performance and Quality	Department of Foreign Affairs and Trade	External
12	Birhanu, Kiros	Researcher	Reality Check Approach	Internal
13	Boddington, Sarah	Director, Governance Section	Department of Foreign Affairs and Trade	Connected
14	Cameron, Lisa	Professorial Research Fellow	Melbourne Institute of Applied Economic and Social Research University of Melbourne	External
15	Chambers, Robert	Research Associate	IDS	Connected
16	Crawford, Joanne	Research & Policy Lead, Individual Deprivation Measure	International Women's Development Agency (IWDA)	External
17	Davis, Tom	Assistant Director, Sri Lanka and Maldives Section	Department of Foreign Affairs and Trade	External
18	Denny	Researcher	Reality Check Approach	Internal
19	Devitt, Rebecca	First Secretary, Governance	Department of Foreign Affairs and Trade, Australian Embassy Jakarta	External
20	Dogbe, Tony	Managing Director	Participatory Development Associate	Internal
21	Dumble, Sam	Statistician	Statistics for Sustainable Development	Connected
22	Flomersfeld, Sarah	SPM Strategy and Coordination manager	Department of Foreign Affairs and Trade, Australian Embassy Jakarta	Connected
23	Gadhavi, Vishal	RCA Lead on the Nepal RAP RCA M&E	Itad	Connected
24	Goulding, Sarah	Assistant Secretary, Global Development Branch	Department of Foreign Affairs and Trade	External

[67] This list does not include all of the RCA researchers who took part in the 2 focus groups in Jakarta

25	Grayman, Jesse Hession	Senior Lecturer and Discipline Convener, Development Studies	School of Social Sciences, Faculty of Arts University of Auckland	External
26	Greillier, Rachel	Senior Technical Specialist - Gender and Social Inclusion	Options Consultant UK	External
27	Guggenheim, Scott	Independent Social Development Specialist	Former adviser to AusAID/DFAT Indonesia	Connected
28	Hannigan, Lisa	Director, Poverty and Social Transfers Section	Department of Foreign Affairs and Trade	External
29	Hawkins, Kirstan	Strategic Advisor	Options Consultant UK	External
30	Indra, Yeni	Researcher	Reality Check Approach	Internal
31	Jupp, Dee	Technical Adviser	Reality Check Approach	Internal
32	Karetji, Petrarca	Team Leader	Knowledge Sector Initiative	Connected
33	Kartika, Astrid	Senior Program Manger	Department of Foreign Affairs and Trade, Australian Embassy Jakarta	Connected
34	Kelly, Linda	Co-Director	Institute for Human Security and Social Change	External
35	Kidd, Stephen	Independent consultant, Poverty and Social Analysis	Development Pathways	External
36	Koirala, Neha	Researcher	Reality Check Approach	Internal
37	Koleros, Andrew	Director of Research, Monitoring and Evaluation	Palladium (UK)	Connected
38	Lang, Hazel	Assistant Director, Regional Engagement Section	Department of Foreign Affairs and Trade	External
39	Lawe-Davies, Frankie	Assistant Director, Poverty and Social Transfers Section	Department of Foreign Affairs and Trade	External
40	Lewis, David	Professor	Department of Social Policy, London School of Economics & Political Science	Connected
41	Loar, Rebecca	HCD Researcher and designer	Independent	External
42	Mulhern, Emma	HCD Evaluator	Itad	External
43	Nixon, Nicola	Counsellor, Poverty and Social Development	Department of Foreign Affairs and Trade, Australian Embassy Jakarta	Connected
44	Norup, Stewart	M&E Adviser	MAMPU	Connected
45	Nugroho, Karishma	Independent consultant and Team Leader	Winrock, Thailand	External
46	Palayukan, Grace	Consultant	KOMPAK	Connected
47	Patel, Arthi	Social Development Consultant	Independent (formerly Head of Poverty and Social Transfers Team, DFAT)	External
48	Pinney, Andrew	Director	Statistics for Sustainable Development	Connected

49	Powis, Ben	Social Development Adviser	DFID Myanmar	Connected
50	Prawiradinata, Rudy	Deputy Minister PPN/ Head of Bappenas for Regional Development	Ministry of National Planning Development (GoI)	Connected
51	Purnamasari, Ririn	Senior Economist	World Bank	Connected
52	Rahwidiati, Diastika	Deputy Head Office	UN Pulse Lab	Connected
53	Riddle-Carre, Peter	RCA+ Team Leader	Reality Check Approach	Internal
54	Roche, Chris	Consultant		External
55	Roshita, Airin	Nutrition Specialist	UNICEF	Connected
56	Satria, Sentot	Senior Social Development Adviser	KOMPAK	Connected
57	Satriawan, Elan	Head of Policy Working Group	National Team for Accelerating Poverty Reduction (TNP2K)	Connected
58	Shanahan, Kate	Team Leader	MAMPU	Connected
59	Sharpe, Joanna	First Secretary- Poverty and Social Development,	Department of Foreign Affairs and Trade, Australian Embassy Jakarta	Connected
60	Shwe Yi Win, Ei	Managed PEER research used to explore insights about female sex workers in Myanmar in 2014	CARE Myanmar	External
61	Slattery, David	Director, Office of Development Effectiveness	Department of Foreign Affairs and Trade	External
62	Smales, Phillipa	RDI Network and Partnerships Manager	RDI/ACFID	External
63	Stein, Danielle	Technical Manager, MEL Component, IP-SSJ	Palladium (UK)	Internal
64	Sumarto, Sudarno	Senior Policy Adviser	National Team for Accelerating Poverty Reduction (TNP2K)	Connected
65	Taylor, Ben	Director	Springfield Centre	Internal
66	Taylor, Laura	Assistant Director, Australian Volunteers Section	Department of Foreign Affairs and Trade, Australia	External
67	Thomas, Dan	Second Secretary	Department of Foreign Affairs and Trade Australia	Connected
68	Tnunay, Damaris	Partnership Manager	MAMPU	Connected
69	Tobing, Deb	Researcher	Reality Check Approach	Internal
70	Uguy, Samy Leroy	Head of Research and Development Unit	Ministry of Village, Disadvantage Regions and Transmigrations (GoI)	Connected
71	Walker, David	Senior Consultant	Itad	External
72	Willcox, Matthew	Founding Partner	The Business of Choice	External
73	Winoto, Anna	Program Director	KOMPAK	Connected
74	Wrecker, Richard	Risk Reduction Specialist	UNICEF	Connected
75	Wyler, Bernie	Team Leader	Maxwell Stamp, Uganda	Connected

76	Yulaswati, Vivi	Director of Poverty Reduction and Social Welfare	Ministry of National Planning Development (GoI)	Connected
77	Yulia	Researcher	Reality Check Approach	Internal

Annex 5 – Building trust in the longitudinal Bangladesh RCA

A previous review of the Bangladesh RCA highlights four relevant findings concerning building trust. First, almost all host households felt uncomfortable initially to host unknown outsiders. Second, households were not exactly clear as to the purpose of the exercise or what would necessarily come out of it. Third, despite this, the experience of having an audience which was interested in hearing what they had to say was clearly a positive experience for many.

> *This mode of working eventually began to create strong bonds that helped to build trust and create high quality conversations. People told us that they were appreciative of simply being listened to by outsiders, and this made it more likely that they would talk openly about their situations: 'We have many complaints and suggestions to make, but nobody ever listens*' (Lewis, 2018:22).

While Lewis et al. (2012) suggests that there was no evidence that people had been empowered by the interactions, they appreciated the fact that they might be able to pass on their views to those in power, and that this might results in positive change for Bangladesh.

Fourth, a relationship of trust and friendship was built up over time because, unlike other outsiders, RCA researchers returned: "'*firstly we did not believe that they would come back. But at the end of the first visit, there was a relationship that we shared with each other and we thought she could not lie to us'*" (Lewis et al., 2012: 41). Arvidson[68] (2013) writes that the longitudinal aspect of the study played to their advantage: during the fifth year, for example, they were told by some people, 'I actually lied to you last time we spoke'.

As noted by Lewis (2018), revisits incurred trust that researchers had a genuine interest in understanding practice on the ground and that talking to researchers would not result in negative repercussions. This raises real questions about whether pulse taking studies are able to establish the intimacy needed for more open and honest interactions to take place, and cautions against "over-reaching", as participant responses are likely to be less valid early on in an immersion or as part of a one-off research exercise (see Pain, Nycander and Islam, 2014 and Shah, 2018).

[68] Arvidson participated in this RCA study as a researcher.

Annex 6 - Utilizing a transect walk in the RAP RCA

The below is an excerpt from the RAP Beneficiaries Study Report (RCA, 2015b). It described how a PRA method has been used to enhance understanding in the RCA study.

The team walked from before the start of the RAP road. This was to understand its relevance in terms of connectivity first hand but also to avoid being automatically connected to the RAP project (had we been driven to the start point). In order to get a sense of the road's 'zone of influence' the team walked several kilometres along the road and off the road. Walking the road provided the opportunity to meet many people and chat generally about the community and the road…As walkers (with back packs) it felt very normal for the team and acceptable to those we engaged with to fall into conversation with the contractors and their workers and then use these experiences to have further conversations with villagers and we were not connected to RAP (thereby maintaining the important objectivity and independence…The road walk also enabled the team to experience the state/condition of the road. For example, there were numerous cuts made across the road to allow water to drain from the fields above the road… 'wheat needs less water so we drain it away'. But many of these had resulted in muddy trenches which were very difficult to negotiate 'the ambulance cannot get through now because of these'. A tractor driver said about these cuts in the road 'yes we tell them not to do it but talking to these people is like talking to an ox'. There was also a water point on the leeward side of the road which was continuously running and also cut a channel across the road. Asked if the contractors were going to do anything about the illegal cutting or this water point trouble spot, we were told it was not in their contracts so 'no'. Walking the road also enables the team to spot project sign boards and engage people in conversations about what the projects are. The fact that people often do not know or are unclear provided insights in themselves.

Annex 7 - Memory recall: what does the evidence tell us?

The typical cognitive psychologist's definition of short term memory is 15-30 seconds — enough to remember the beginning of a sentence so that the end of the sentence makes contextual sense. There are many factors that allow (or prevent) something going from short-term to longer term memory — the vividness of it, how well it fits with existing mental models, "noise" at the time, what has immediately preceded it or follows it. Sequence is also important — we have a tendency to remember best the last thing in a sequence, next best the first thing, and worst of all the stuff in the middle.

Recall is better when the environments are similar in both the learning (encoding) and recall phases, suggesting that context cues are important and therefore writing notes needs to happen at the time of the event or very shortly thereafter.

Memory recall is also a function of the depth of mental processing, which is in turn determined by connections with: i. pre-existing memory (we discuss this in point 4., below) and ii. time spent processing the stimulus, amongst other things, Thus "shallow processing" (such as, typically, that based on sound or writing) leads to a relatively fragile memory trace that is susceptible to rapid decay, whereas "deep processing" (such as that based on semantics and meanings) results in a more durable memory trace. This suggests that even when items are written down, it is important that these be processed in a way that makes meaning of them. This is typically done in the field by researchers having a way to note early insights and reflections and possible hypotheses separately from more verbatim and factual note taking, for example in a column down the right-hand side of the page in a notebook, or highlighted in some way. This could be supplemented with daily de-briefs with the research team. Research has found that we are better able to remember information if we encode it in a meaningful way. When we engage in elaborative encoding we process new information in ways that make it more relevant or meaningful.

The efficiency of memory recall can be increased to some extent by having prior knowledge of a subject, as well as an organized cognitive map or mental framework of the issue. Such schemata are also applied to recalled memories, so that we can often flesh out details of a memory from just a skeleton memory of a central event or object. This suggests that a prior conceptual framework can aid with memory recall; however care does need to be exercised as the use of schemata may also lead to memory errors as assumed or expected associated events are added that did not actually occur.

Vivid, outlying information is more likely to be remembered, because it is easier to recall. This is partly down to something called the availability heuristic. If there is something in our memory that is easier to recall, then we will give that more weight to that information than we rationally should That said, outliers tend to be noticed if they connect to things we know in some way. Things that are completely outside our expectations of an experience may not be noticed because we have no way to place them in our cognitive maps.

Sources: Craik, Fergus I.M., and Robert S. Lockhart (1972)" Levels of Processing: A Framework for Memory Research" in Journal of Verbal Learning and Verbal Behavior 11. 671-84; Tulving, Endel and Donald M. Thomson (1973) "Encoding Specificity and Retrieval Processes in Episodic Memory" in Psychological Review, Vol. 80, No. 5, 352-373; What We Miss, New York Times review of The Invisible Gorilla, accessed 15 October 2018; https://www.npr.org/2016/04/17/474525392/attention-students-put-your-laptops-away, accessed 15 October, 2018; Matthew Willcox, Institute for Decision Making, pers. comm., 3 October 2018.

Annex 8 – Presentation of data in the Haze RCA Study Report

The table below is taken from the Haze RCA study report (p. 20). It is referenced in the report thus: "Table 4 shows what people think about the changing intensity of the haze". It is not clear whether this reflects the general impressions of a number of researchers or if this reflects a number of exercises carried out in communities. Nor is it clear how many people's views this represents, and which people, whether there were diverging opinions, and who held these. This is not atypical of the way that data is presented in reports. It reflects the lack of precision in data collection and analysis.

Table 4: Is the haze getting worse?

Location	Is the haze getting worse?	How long have companies been in the area?
Peri-urban S Sumatra	Lighter haze before 2015	No companies near the village
Rural S Sumatra	Same as before	1995
Peri-urban C Kalimantan	Worst haze in 1997 and 2014/15	2002
Rural C Kalimantan	Worst haze in 1997 and 2014/15	2009
Peri-urban W Kalimantan	Worst in 1997, 2015 haze associated with less trees/low water	2005
Rural W Kalimantan	Same as before	No companies
Jambi	Mostly same as before	Early 2000 and 2009 (one company)
Riau	Few people saying yes	2000

In the IP-SSJ study, Table 6 presents what residual effects of the Maoist insurgency are experienced by "people". It does not present information on which people – what caste or ethnicity, male or female, young or old - experience these residual effects. As a result, the data is not helpful.

It appears that many of tables and other visuals that present rankings and scorings were actually produced by researchers after fieldwork based on their recollections, not by participants during research, as would be the case were participatory tools used more systematically. This data would be much richer and more reliable if captured in the field.

Annex 9 - Assessment of analytical rigor in regard gender and social inclusion issues in the RAP3 and IP-SSJ RCA studies

While the RAP3 baseline report notes some caste differences, it is largely gender blind. For example, "*The reliance on remittances in all the study villages except Village C (Humla) means people emphasise the special importance of networks and connectedness in securing good jobs, particularly in India*" (RCA, 2015c: 14). One wonders whether this applies to men and women equally, and to all caste and ethnic groups equally. The report discusses choices around which children to educate when there is resource scarcity, and reports that people told them they do not decide based on gender [*sic*] but we are not told whether male and female parents, teachers, male and female children all reported this, and there appears to have been no observation of whether there were differences between what people reported and what happened in reality, nor asking for examples (possible case studies) of where this had happened.

While the RAP3 midline RCA does somewhat better in highlighting gender differences, but still, in our view, falls short of what one would expect. For example, there is very little intra-household disaggregation or information on intra-household dynamics, something that the RCA claims is – and we would expect would be - a strength of the approach because researchers live within households. Often information is attributed to "households" or "families" (with an assumption of consensus). For example, "*Out of the twenty five host households with whom the RCA team has close interactions, nine felt they are better off than in 2014. Of these three families attribute this, at least in part, to RAP roads*" (RCA, 2016h: 43).

We asked one of the Team Leaders why, though women's and men's views were presented, there was no substantive gender analysis in the IP-SSJ study. The Team Leader replied, "*We were not asked to do a gender analysis. We'll only look at gender issues if they are raised by local people. If it's not raised, we'll step back and not push it…Focusing on gender can lead to unconscious bias*" (KII, RCA core team member). We are concerned that an exploration of gender issues, which should be part of all poverty and social analysis, would be considered, in and of itself, unconscious bias. However, even when these issues were raised by study participants, there appears to have been a lack of probing or analysis, which to us signals either complete ignorance of the importance of exploring gender issues, and/or unconscious bias against exploring these issues. For example, "*In some cases, women began discussing domestic violence only to stop themselves mid-way, or be told to stop by others sitting nearby or participating in the conversation. However, in other villages, women who had been physically abused by their husbands spoke freely to our researchers in the presence of other community members*" (RCA, 2015a: 19). One is left wondering what the reason for these potentially normative differences were as they are not explored in the report.

Understanding poverty, gender, caste and ethnicity issues in relation to security and justice were central aims of the study: "*The primary motivation for this study is to improve the evidence base on how people experience and understand security and justice (S&J) in Nepal, with particular reference to poor and marginalised groups*" (RCA, 2015a: 12). When we asked the other Team Leader for the study why the social analysis for the study was so weak, the response was that, "*In this study, we were not trying to get into caste issues all that much. Maybe RCA is not the best approach for uncovering caste relations, but it is good for getting intra-household issues, gender issues*" (KII, RCA core team member).

It is our assessment that there were a number of findings in the report that were likely a result of relying solely on what local people reported, without understanding the gender or caste issues in the local context, or the positionality of the study participant. An example of this is the statement in the IP-SSJ RCA report that, "*No tensions, prejudice or disrespect was apparent in any location between Muslims and Hindus. On the contrary people talked about 'all being brothers' and 'we are all the same poor people'*" (RCA, 2015a: 5). These are very typical statements made by Nepalis to outsiders (including other Nepalis). In point of fact, there have been a number of major incidents of sectarian

violence in the districts where research was carried out. Reports of religious prejudice were played down in the report, with the report statin the following:

> *We feel that this was largely to test the situation and attempts not to cause offence to outsiders rather than a reflection of prejudice within the community. Though the Muslim family acknowledged some prejudice in the community, they said that this was largely related to Hindus not accepting food from Muslims, which they understood as ritual rather than ill will* (p. 24).
>
> *In Kapilvastu (1) we also stayed in a mixed community that included 10-12 Muslim households. Here, our HHH told us they lived together without conflict, though there were some incidences when Hindus put tikka on Muslims* (p. 24).
>
> *It can be argued that safety and security issues do exist for people which do not surface because they have become normalised or people prefer to keep silent. The intimate and informal interaction created in the RCA study enables people to talk about issues they face in the home more freely than in public spaces such as focus group discussions. Under these conditions, the RCA team felt that there was less a sense that people keep quiet than the issues are normalised* (p. 44).
>
> *Concerns were only expressed to our team in terms of unwillingness to give offence around norms of custom and were not manifestations of prejudice* (p. 5).

This last quote reveals little understanding that such prejudice is customary, and suggests potentially ill-informed etic overlays.

Annex 10 - RCA's contribution to RAP3 Quantitative Baseline questionnaire development and analysis

The RCA baseline fieldwork and analysis were done in the months leading up to the enumeration of the baseline questionnaire. This planned sequencing allowed the findings of the RCA baseline analysis to be incorporated into the baseline questionnaire. The following list represent the contributions to the quantitative questionnaire from the RCA baseline fieldwork:

1. Changes to the order and wording of certain questions, including more nuanced list of choices, preventing large numbers of "other" specify responses, that are so difficult to handle in the quantitative analysis.
2. Inclusion of new questions (see below for specific new questions added as a result of the RCA analysis). Particularly the RCA analysis bought into the questionnaire the potential downside risk of increased road access through questions on changing crime and insecurity, and perceptions around the reasons for the changes.
3. Respondents were given an opportunity to identify other aspects of their lives that changed for the worse over the past year.
4. Change in quality of service delivery over the past year and perceived reasons for these changes. Until the RCA analysis, the questionnaire had not embraced the idea that improved road access might bring some negative changes as well as the hypothesised positive that permeated the logframe.
5. Questions on increasing citizen engagement, by asking respondents if they had ever felt like complaining about poor service delivery and if not why not.
6. Question ordering to reduce social desirability bias and/or respondent resistance. The question asking the ethnicity/caste of the respondent was moved from very early on in the sequence to much later, to prevent respondent resistance or social desirability bias, from having shared their caste status with the enumerator.
7. The timing of the quantitative enumeration was informed by results from the RCA in terms of the time of the day and in relation to the agricultural calendar, to ensure that enumeration timing minimised clashes with peak labour periods.

There were also a number of specific questions added to the baseline quantitative survey emanating from RCA baseline analysis

1. Do you think that crime and/or in-security overall in your area has increased, decreased or stayed about the same this year compared to the previous 1 year?
2. What type of crimes or insecurity have increased over the last 1 year? (Do not read aloud, but code when mentioned.)
    a) Why do you think this type of crime or insecurity has increased?
3. Other than crime or insecurity, are there other things that have changed for the worse in your community in the last one year?
    b) What are these things other than crime that have changed for the worse in your community in the last one year? (Do not read out options, but code when mentioned.)
    c) What do you think has caused these changes?
4. Have you had a complaint /concern about any services in this community during the last year?

d) With which service? (Do not read out options but code when mentioned.)
e) Did you do anything about your concerns about the services?
f) If you did something about your concerns, what was it? (Do not read aloud, code when mentioned.)
g) Were you satisfied with the outcome after raising the issue?
h) If you did not do anything about your concerns, why not? (Do not read aloud, but code when mentioned.)

- Pers. Comm., Quantitative evaluation expert

Annex 11 – VfM summary tables

Table 4 – Summary of comparators' approach to key research stages

	RCA	PEER	HCD	Sensemaker	General Qual
Preparation					
Background research	Very little to none	Analysis of existing literature, develop conceptual framework for research	Limited background research	Background research using existing literature	Analysis of existing literature, develop conceptual framework for research
Instrument design	"Areas of conversation" are identified based on discussions with commissioners	Question guides are designed to provide peer researchers with a set of topics to explore	Range of participatory and ethnographic techniques, designed by multi-disciplinary team and implemented in a flexible manner.	Framing of 'prompting question' and signifying questions	Generally based on research questions identified by commissioner, and on secondary literature review Range of participatory and ethnographic techniques identified to investigate research questions. Flexibility ranges, but is generally less flexible where comparability across sites is important.
Fieldworker selection/Team composition	Researchers trained in RCA but no particular background in qualitative work or content area	Peer researchers are recruited. They do not need to be literate, but should represent the community of interest as best as possible.	Depends on qualitative approach. Examples here focused on adolescent RH, where adult researcher paired with youth for actual data gathering. Full field team often quite large, including a range of backgrounds.	Small team trained in capturing stories and guiding participants through process of answering signifying questions	Skilled qualitative researchers, generally with content area expertise, with specific training on particular instruments and contexts

	RCA	PEER	HCD	Sensemaker	General Qual
Fieldwork					
Training	3-5 days of training, including an immersion	Peer researchers are trained, normally over 3-5 days	7-day training boot camp (Ethiopia A360)	Training on story capture/interpretation, either paper-based or using tablets and associated software	Varies, generally between 3 days and 1 week for researchers with existing qualitative experience, depending on the complexity of the issue to be researched and the tools to be used
Sampling	Geographic differences, some loose criteria to meet 'range' of participants including poorest but not always implemented in practice	Peers identify a handful of their network to interview (normally 3-5 each)	Not representative, in specific areas for delivery of design. Around 300 included in Ethiopia A360 formative research. Adolescents identified by health workers (therefore limited range of different types/backgrounds of girls' perspectives included)	Depends on the design. Can be as representative as desired, from random sampling to purposive.	Depends on the method, but normally some kind of purposive sampling depending on the research framework, to provide balance across a range of criteria
Data collection	No (or limited) note-taking	Peer researchers undertake a series of interviews (presented as more informal 'chats'), normally doing several interviews with each participant. No notes are taken during the interviews, instead findings are downloaded to lead researchers who debrief the peer researchers, ideally frequently.	2 week process. Research teams paired youth with adult researcher, plus another for note taking.	Fairly quick, many stories can be captured in a single da.	Varies – depends on specific tools used, but documentation in the field normally rigorous.
Analysis					

	RCA	PEER	HCD	Sensemaker	General Qual
Synthesis and analysis	No synthesis during fieldwork Lead researcher debriefing with each researcher, then sense-making workshops including all researchers. Analysis primarily undertaken by Lead Researcher and other Senior Researchers who have not been involved in the field research	Synthesis/analysis undertaken with peer researchers initially as lead researchers debrief peer researchers, then continued iteration by lead researchers using qualitative techniques (coding, sorting, analyzing, etc).	Daily downloads' of key observations, emerging themes, initial coding of information. Across researchers, documentation analysed in an iterative way for key themes and insights to emerge, and to start testing ideas for potential design solutions.	Only synthesis done in the field itself is that by participants themselves as part of signification questions. Analysis is iterative, in that data initially analysed by lead researchers using software, iterating between 'big picture' and story packs. Once an initial set of observations/findings is reached, ideally the process then includes a 'human sensemaking' stage to take results back to participants for further insights and interpretation.	Iteration across researchers in the field as well as during analysis phase. Iterative process of coding, interpretation, categorisation, hypothesis testing, generally lead by Lead Researcher, with input from other researchers.
Documentation of analysis process	Minimal documentation of fieldwork, synthesis or analysis	Notes from debriefings are the main documentation	Researchers given kit for documentation and consent process in A360, but this is not always standard. Documentation of analysis generally poor	All participant stories and answers to signifying questions recorded	Documentation should be rigorous

Table 5 - Summary of rigor in research preparation and design by comparator

	PEER	HCD	Sensemaker	General Qual
Preparation and design				
Context/understanding	Lead researchers develop framework based on existing literature and evidence base	Secondary research phase should inform design BUT, in practice, weak use of other research beyond basic DHS data and researchers are not subject or geographic experts	Secondary research phase should inform the design of the prompt and signifiers	Secondary research phase should inform the design of the prompt and signifiers
Framework to guide enquiry	Theory/framework used based on existing understanding of context/evidence	Draws on frameworks and techniques in other qualitative/participatory work	Frameworks are used to inform the understanding of the issue being investigated, questions to ask, etc	Frameworks are used to inform the understanding of the issue being investigated, questions to ask, etc
Experienced, reflexive and well-trained researchers	Peer researchers are by nature inexperienced, but they are provided	Lead researchers should be well trained, paired with local researchers with less	Researchers need to have specific capacity to implement Sensemaker In practice,	Researchers are well-trained in qualitative approaches In practice Depth of experience

	PEER	HCD	Sensemaker	General Qual
	training. Lead researchers are experienced social scientists.	training. But in practice, HCD lead researchers not experts in qualitative research. Local (youth) researchers have limited training	Adequate investments in capacity within organisations can be under-estimated, and there can be a reliance on a small number of external consultants	can depend on the context ; training then needs to supplement lack of experience, quality depends on lead researcher ability to train and coach researchers in the field and monitor quality in real time, correcting issues as they arise
Multi-disciplinarity	Researchers tend to be anthropologists.	Not explicitly incorporated although some applications have been done in combination with many disciplines (e.g. A360)	Not an explicit aspect of the approach, but certainly can be incorporated in the framework although Not evident in available case studies	Not an explicit aspect of the approach, but certainly can be incorporated in the framework Not evident in available case studies

	PEER	HCD	Sensemaker	General Qual
Understanding position	Training focuses on ensuring peer researchers understand their position in asking questions (framing, probing, etc) and interpretation.	In theory, Training of researchers BUT, in practice, Design research often conducted by very young non-locals who have a limited background in developing country contexts and formal research approaches and can bring significant (unrecognised) bias to research. International design team members are not trained in qualitative methods and are not generally aware of issues or how to mitigate.	Power and relational dynamics between the researcher and participant are minimized by the 'self signification' process. But in practice, the position of the researchers is still important (needing to understand the context in which narratives are provided), and is not always adequately considered.	Power and relational dynamics between the researcher and participant are recognised. Instruments, techniques, and training designed to acknowledge these dynamics and limit bias.

Table 6 - Summary of rigor in fieldwork by comparator

	PEER	HCD	Sensemaker	General Qual
Fieldwork				

	PEER	HCD	Sensemaker	General Qual
Being unobtrusive	Researchers drawn from community of interest in order to ensure deep contextual understanding and trust, and are taught not to be 'judgemental' or to lead respondents into particular answers. Interviews can be presented as 'chatting' rather than formal. Asking questions in the third person to avoid respondents having to divulge private information.	Often includes researchers from intended population, includes efforts at immersion in experiences related to product being designed, but emphasis less on reducing bias from being observed and more on allowing designers to gain experience of user's context.. In practice, youth/local researchers integrated only partially; although recognition of need to incorporate them throughout process more. Research approach not geared to build trust so deep issues are not always surfaced.	The research is not entirely unobtrusive, since researchers interview participants in a very structured format, but there is an acknowledgement of the need to ensure participants are made to feel comfortable. In practice, there is not always adequate reflexivity about the ways in which the obtrusiveness of the interview might affect responses.	The research is not entirely unobtrusive - there is normally a clear acknowledgement of the research as a process - but there is an acknowledgement of the need to ensure participants are made to feel comfortable.
Triangulation (method, time, persons)	Each peer researcher interviews 3-4 community members, sometimes over multiple sessions. Researchers can be actively chosen to represent a range of community experiences.	Interviewing different actors; Using different techniques (participatory, interview) to explore from multiple angles But in practice, recruitment process may lead to similar population/voices and hence early saturation	Triangulation across participants implicit in that they are all asked the same prompting question and signifiers, but no triangulation across methods. Triangulation across time can be done if implemented in a longitudinal context but this is very rare in practice.	Triangulation between data sources (q-squared), across instruments/techniques, and across researchers.
Respondent validation	Peer researchers actively involved in interpretation of findings	Core part of philosophy of ensuring understanding of participant's perspectives	Respondents are explicitly used to validate their own narratives, but are also often brought in again for a 'human sensemaking' phase, where they interpret the patterns that have emerged from the data.	Respondents can be brought in again for a 'human sensemaking' phase, although this not often done in practice

	PEER	HCD	Sensemaker	General Qual
Faithful and accurate recording	No notes are taken during interviews (to maintain intimacy and trust) but field notes should be recorded within 48 hours of each interview by lead researchers, although in practice there is often a larger lag	In theory, documentation as per qualitative techniques being used but in practice is generally much less rigorous than standard qualitative research.	Data is captured in the software	Detailed record-keeping, whether recordings/transcriptions/translations of individual interviews and focus groups; capture of participatory approach data; field notes; etc

Table 7 - Summary of rigor in fieldwork by comparator

PEER	HCD	Sensemaker	PEER	HCD
Synthesis, Analysis, and Reporting				
Iterative analysis	Synthesis/analysis undertaken with peer researchers initially, then continued iteration by lead researchers using qualitative techniques (coding, sorting, analyzing, etc) to build and test hypotheses and explore representativeness of cases.	Multiple points for synthesis, including 'daily downloads' with initial coding and synthesis of emerging findings in situ, as well as drawing together across researchers at the end, although in practice synthesis can be too reductive. It can also be spread out geographically limiting depth of overall synthesis and extent of hypothesis testing in practice.	The analysis is inherently iterative, including inductive and abductive approaches. But in practice the extent of iteration possible is limited by the quality of the prompting and signifier questions; poor instrument design cuts of potential for meaningful iteration	The analysis is iterative, normally involving textual analysis of data in a process of sorting, indexing/ coding, summarising, abstracting/interpretation.
Building and testing hypotheses in data analysis		Not rigorously done, researchers may not have skills to do this		
Ensuring representative-ness of cases		Not done in practice	Depends on the quality of the sampling strategy used. Best practice has taken a thoughtful approach to sampling, so results are as representative as desired. Poor examples have faltered on this aspect, leading to a lack of confidence in the results.	Depends on the nature of the research, but generally done well, based on sampling strategy and detailed background research on the context and content
Transparency	Thorough documentation of interviews and synthesis/analysis process	Not explicitly incorporated into the methodology. In practice transparency is low,	Transparency is provided by the fact that data is captured in the software, and analysis	Process well described and evidenced using data gathered

		with a lack of clarity over why certain design options preferred by IDEO.org team	could then be replicated later if desired	