Power, data and knowledge in development sector impact evaluation activities: critically engaging with the impact iceberg

Paul Kelly

HighWire Centre for Doctorial Training

Lancaster University, UK

This thesis is submitted for the degree of Doctor of Philosophy

15th of September 2019

Dedicated to

Sirinart, Benjamin, Jeorge and Jemma

"My point is not that everything is bad, but that everything is dangerous, which is not exactly the same as bad … If everything is dangerous, then we always have something to do. So my position leads not to apathy but to a hyper- and pessimistic activism."

Michael Foucault (1983: 231-232)

Declaration

I declare that the content of this thesis is my own work and that this has not been submitted in any form for the award of a higher degree elsewhere.

I also confirm that any quotation or paraphrase from the published or unpublished work of another person has been duly acknowledged in this work which I present for examination.

Excerpts

Excerpts from this thesis have been used in the following publications, presentations, conference manuscripts and academic publications.

Simeonova, B., Karanasios, S., Galliers, R. D., Kelly, P., & Mishra, J. (2018b). Where is Power in Information Systems Research? Towards a Framework. *International Conference on Information Systems (ICIS)*, 13-16 December 2018, San Francisco, US. https://aisel.aisnet.org/icis2018/behavior/Presentations/10/

Simeonova, B., Karanasios, S., Galliers, R. D., Kelly, P., & Mishra, J. (2018a). New ways of organising in collaborative knowledge sharing: examining the effects of power. *European Group for Organizational Studies Colloquium (EGOS)*, 5-7 July 2018, Tallinn, Estonia.

Kelly, P. (2018). An activity theory study of data, knowledge, and power in the design of an international development NGO impact evaluation. *Information Systems Journal, 28*(3), 465-488. https://doi.org/10.1111/isj.12187

Hayes, N., Introna, L. D., & Kelly, P. (2017). Institutionalizing Inequality: Calculative practices and regimes of inequality in international development. *Organizational Studies, 39*(9), 1203-1226. https://doi.org/10.1177%2F0170840617694067

Kelly, P. & Larner, J. (2016). Using activity theory to understand the design patterns visible in circulations of knowledge and data in the 3rd sector. *Invited presentation at the Centre for Research, Activity & Learning (CRADLE)*, 21 April 2016, University of Helsinki, Finland.

Kelly, P. (2016). Impact knowledge supply chains and activity theory: A story of power, knowledge, a philanthropy and a small NGO. *European Group for Organizational Studies Colloquium (EGOS),* 7-9 July 2016, Naples, Italy.

Kelly, P. (2015). Our need to talk about power & knowledge in evaluation & policy cultures. *Politics & Ideas*. http://politicsandideas.org/our-need-to-talk-about-power-knowledge-in-evaluation-policy-cultures/

Secker, E. & Kelly, P. (2015). Quick and Dirty M&E: Co-designing and co-evolving for engaged impact. *Youth | Participation | Impact Conference*, 2-4 September 2015, University of Manchester, UK.

Kelly, P. (2015). Evaluating impact, or the impact of evaluations? NGOs, donors and ontological politics in an activity theory study of power dynamics in international development. *European Group for Organizational Studies Colloquium (EGOS).* 2-4 July 2015, Athens, Greece.

Kelly, P. & Cowen, M. (2014). CHAT & Governmentality: The uniform in the management and activities of migrant workers. *European Group for Organisational Studies Colloquium (EGOS),* 3-5 July 2014, Rotterdam, Netherlands.

Hayes, N. & Kelly, P. (2014). Impact Assessment in the Development Sector and Sociomaterial Rankings. *Invited presentation at the Information Systems and Innovation Group*, London School of Economics, London, UK.

Hayes, N. & Kelly, P. (2014). Performing Impact Assessment in the Development Sector: The Case of Imagine Foundation. *Latin American and European Management and Organisation Studies Group (LAEMOS)*, 31 March – 4 April 2014, Havana, Cuba.

Abstract

The study opens with a simple observation: International Development is now a digitally data/knowledge intensive sector. This mundane remark suggests development has also become a rich domain for studying *how power, data and knowledge are related*. In data/knowledge intensive development, our ideas about impacts may appear empirical and objective, based on data and expertise, and often digitally immaterial. However, these ideas involve power and politics which shape data desires and data cells, knowledge products, packaging and pitching, governing processes and digital systems. Such dynamics place demands upon agents like non-government organisations (NGOs), transforming them from an old emphasis on tangible and physical relief, such as providing medicine, food, or clothes to those in need, to constructing and exchanging digital data and expert impact products in diverse development markets and bureaucracies. Processes, like impact evaluation are caught up in this shift to data/knowledge intensity and the power relations, inequalities and silences that accompany the shift. This study elucidates these power/data/knowledge relations, specifically with regard to impact evaluation itself, but more broadly regarding data/knowledge intensive development, or what some observers have christened "development 2.0".

Two bodies of literature are relevant to the study. Firstly, the normative discourse on *how-to-do evaluation*, and secondly, critical perspectives which acknowledge power and politics, and *the impacts of (the doing of) evaluation*. The prescriptive discourse is found to rely on old and problematic, yet still pervasive and implicit, models of data/knowledge construction which ignore or elide practice and power. Therefore, an alternative view of data/knowledge construction is proposed, and a research approach developed to explore how critical insights might live in the wilds of development sector evaluation practice. The approach draws on Cultural Historical Activity Theory (CHAT) to scaffold engagement with research partners and to critically analyse evaluation activities. The research design features two qualitative and collaborative cases, where the author offers consulting advice and simultaneously researches evaluation activities with two small NGOs, one expert at evaluation (based in India), and one learning about evaluation (based in Thailand). The CHAT analysis articulates evaluation activities, systems, contradictions, and the notion of temporal activity chains where impact data and knowledge are step by step, incrementally, edited over time, bifurcated, and packaged into

exchangeable, legitimate impact products. The expert NGO case shows how this is effectively and instrumentally done, but the novice NGO case reveals the struggles and transformations required to become expert evaluators in development 2.0. Both cases reveal a range of silences, inequalities, assumptions, omissions, priorities, and opportunities which are part of contemporary impact evaluation mechanics.

Three contributions are offered in response to impact power/data/knowledge problems. The first contribution, to evaluation literature, is the "impact iceberg". This is a diagnostic, big picture tool which describes contemporary legitimate impacts, and amplifies the illegitimate edited-out impacts, those silenced or submerged under the iceberg waterline. The second contribution is a set of four provocative conceptual devices which audit evaluation practice and power, namely: "data/knowledge networks", the "6P sensitivities", the "impact spectrum"; and the "datamentality/datamateriality" oppositions. These concepts contest conventional audit cultures which align evaluation power and control towards managerial, commercial, digital, and expert domains. The final contribution is "critical engagement", a methodological proposition, which tests and brokers critiques in engagements with research partners to see if and how they are accepted, rejected, or adapted in NGO evaluation practices.

In sum, the contributions are advocated as components of an alternative "evaluation-as-practice" view. This view diverges from existing evaluation views, particularly scientific, business-pragmatic, participatory, and technology-centric perspectives. As such, the study speaks to the concerns of power/data/knowledge relations in the practices of evaluators, NGOs and donors whose work is part of development 2.0 today, as well as to information and knowledge management professionals and researchers who wish to not hide the power relations generated in the mundane data/knowledge intensive worlds they inhabit.

Acknowledgements

Contents

List of Tables

List of Figures

Chapter 5

Chapter 6

Chapter 7

Chapter 8

List of thesis terms and abbreviations

Audit 1.0 / Audit Culture – a form of auditing that has spread in managerialist approaches to development work, which classifies and monitors organisational inputs and outputs to re-distribute control towards managerial stakeholders, often using statistical metrics.

Audit 2.0 – a form of auditing, novel to this thesis, which refocuses auditing processes on multi-stakeholder dialogue, power relations and closer understandings of practices, away from distal standardisation, decision-making, data capture and surveillance frames.

Critical Configurations – a term used to analyse evaluation processes from a broad socio-political perspective. The term "configurations" speaks to patterns and relations rather than linear causes and effects. The term "critical" highlights concerns about power and inequality.

CDS – Critical Development Studies, an academic sub-discipline, highlighting mainstream development omissions, particularly around power, post-colonialism and inequality.

CHAT – Cultural Historical Activity Theory (also Activity Theory), a broad range of related approaches to social science research and development, with shared roots in 1920-1930s Soviet social psychology, particularly work by Lev Vygotsky and Aleksei Leontyev.

Critical engagement – a research approach adopted in this thesis, which focuses on how critical perspectives are interpreted, adopted, adapted or rejected by research partner practitioners as part of their own work practices.

Development 2.0 – a phrase originally denoting data intensive and social media impacts in development, and here explicitly extended to indicate increasing data/knowledge intensity, digital representations, and globally scaled circulations of development data/knowledge products in the sector.

DIKW – the Data, Information, Knowledge and Wisdom Hierarchy, or Pyramid, a widespread and implicit model in contemporary data/knowledge work. Identified in the late 1980s, the model foregrounds digital data as a basis for knowledge, and knowledge as instrumental for organisational decision-makers.

Data/knowledge – a term used to denote data/knowledge intensity, and to foreground the demand for empirical and digital data sets and professional technical knowledge as key components of these intensities.

DS – Development Studies, an academic discipline focused on development in global and local contexts; emerged in 1950s-1960s after the Second World War and during postcolonial shifts.

E-development – a term used in academic and practitioner communities to denote the use of information systems in managing development interventions, programs, and projects.

HTSG Thailand – pseudonym used to label the NGO in the second case study, chapter 6.

ICT4D – Information and Communications Technology for Development, a relatively new discipline with parentage in Information Systems and Development Studies.

ICT4D 1.0 – term used by observers to denote 1990s approaches to using ICTs in development, to make development more effective, or efficient, but not to radically change development.

ICT4D 2.0 – term used by observers to denote uses of ICTs from the 2000s, which support participatory, grassroots and locally owned development processes using digital technologies.

Impact Governing – term used to denote the application of technical evaluation methods and results in framing or explicitly controlling the work of development agents, such as NGOs.

Impact Iceberg – a term and contribution to evaluation literature, which outlines which kind of impact concerns and products are rendered audible, above the waterline in contemporary impact evaluation, and which are rendered silent, under the waterline

Impact Marketing – term used to denote the use of impact data/knowledge, in reports, arguments, visualisations and other media, to highlight success, win funding and improve reputational status

IS – Information Systems an academic discipline focused on the design, use and impact of diverse information systems in the contemporary world, involving technical and social concerns.

INGO - International Non-Government Organisation, specifically large NGOs which have professionalised and grown in scale since the 1990s.

Input/output meta-model – terms used to focus on how technical impact evaluation methods have come to highlight managerial and expert definitions and intervention inputs and outputs, which frame evaluation concerns, results, relationships and practices.

Knowledge Management / KM – academic and practitioner field which focuses on models and ways of capturing, storing, sharing and using knowledge, predominantly within contemporary organisations.

KM4D – Knowledge Management for Development, an emerging sub-field of largely development sector professionals, consultants and academics doing Knowledge Management in or focusing on development contexts.

NGO – Non-Government Organisation, part of the third sector, an organisation not privately or publicly owned.

Power/data/knowledge – a term used to indicate the interrelations between power, data and knowledge during their construction, application and circulation in the development sector.

Rural India – pseudonym used to label the Indian NGO in the first case study, chapter 5.

TIEK – Technical Impact Evaluation Knowledge, a term used to denote technical, professional and expert, often scientific or participatory, definitions, methods and prescriptions on how to do evaluation. A largely normative discourse.

Note: Terms are capitalized when showing they are unique labels, abbreviations or names of academic disciplinary fields (e.g. "Knowledge Management" as a discipline). In the thesis, lower case is used for more general use or to indicate practitioner activities (e.g. "knowledge management" as a general work activity).

List of Appendices

1 Introduction: Power in data/knowledge intensive impact evaluation

1.1 Development 2.0 impact evaluation

This study begins with a perhaps unsurprising observation: that International Development is now a data/knowledge intensive sector. This observation means it has also become a rich domain for studying power, data and knowledge relations. This thesis will argue that in addition to development's continuities with historical empire and its popular image as a moral and caring sector where tangible help for others is the priority, it is international development's data/knowledge intensity, its focus on the conceptual, intangible and digital, that is now most pivotal in explaining its power dynamics (Hayes et al, 2017; Kelly, 2018). Non-government organisations (NGOs) in the sector face demands that transform them from an old emphasis many had on tangible and physical relief, such as providing medicine, food, or clothes to those in need, to today constructing and providing digital data and expert knowledge products for diverse markets, funders, partners and bureaucratic needs. Never before have such travelling "*ideas*", "*theory*" or "*policy*" as representations been so important in poverty elimination (Mosse, 2007). Such intangible ideas and concepts not only shape but actually constitute

practice and power in development. They shape what data is captured, edited and discarded, and what knowledge is packaged, ignored, exchanged or pitched, in pervasive development processes, such as impact evaluation.

Impact evaluation is one of many pervasive development processes that illustrates these power/data/knowledge relations well. Development evaluation discourse, like development finance, is dominated by relatively few large knowledge producers such as The World Bank, the United Nations, USAID, DIFD, large international NGOs (INGOs), and universities. This contrasts with the thousands, or millions, of small NGOs working around the globe (Lewis & Kanji, 2009), who come to consume aid finance, data/knowledge prescriptions, and demands for impact results. They exchange impact results for funding and reputation opportunities (Wallace et al, 2006). Development impact evaluation is particularly problematic because of these uneven data/knowledge edits and flows, and the evaluation discourses that regulate what impacts can or should be, from far away.

Yet mainstream discourses, such as technical impact evaluation knowledge (hereafter "TIEK"), ignore such inequalities, and instead continue to prescribe rigorous methods and robust result frames for NGOs to adopt. Such prescriptions presume access to digital technologies and the often ideological benefits of embracing TIEK. Normative technical evaluation discourse has become part of the wider data/knowledge[1] shift, evident in calls for a new, modern and emancipatory view of development based on digital technologies. This call has been characterised as "development 2.0". It is founded on the integration of diverse information and communications technologies (ICTs) into development processes at the level of state funding and specific initiatives (Thompson, 2004: 2; Thompson, 2008). With these technology foundations, development 2.0 appears to widen our knowledge of poverty and marginalisation, making development faster, more efficient, more impactful, and somehow digitally immaterial.

Diverse benefits are claimed in this modernist, technology-centric discourse. Development 2.0 increases connectivity, participation, co-creation, collaboration, pluralism, involvement in governance, and policy-making (Leadbetter & Cottam, 2007; Thompson, 2008), and

---

[1] The terms "data and knowledge" or "data/knowledge" are used regularly to highlight digitisation and professional expertise. They feature heavily in development terminology and debates around evidence, results, inputs, outputs, big data for development (BD4D), local knowledge, and knowledge for development (K4D).

challenges old development (Thompson, 2008: 825-826). Heeks (2008; 2010) highlights how the assumptions, structures and processes of development are changing. Development 2.0 will connect the excluded, support collective power and bring new opportunities for work, entrepreneurs and businesses (ibid, 2010). Quaggiotto (2007) claims traditional aid i.e. "development 1.0", is in crisis, in need of renewal, new intelligence, and aid equivalents of digital organisations like Amazon or eBay. Quaggiotto advocates freeing data, failing quickly (2009), and overcoming the already apparent development 2.0 inertia of "*cultural resistance*" to modern technology (2010).

Transitions towards such data/knowledge intensity and development 2.0 are paralleled in the promise of e-development and consulting models, focused on more efficient, effective understanding of aid problems and solutions because of technologies, private sector solutions and digital data (Brigham & Hayes, 2013: 112). In these often uncritical views, modern technologies, data analytics, and knowledge management have become instrumental for progressive impact. They fix where old development failed. Such promises are part of the latest wave of development modernisation discourse, following shifts from colonial to technical discourses in the 1950s and 1960s (Kothari, 2005a), and NGO shifts from grassroots innovation to global service providers in the 1980s and 1990s (Wallace et al, 2006). Now is the turn of digitisation and data/knowledge intensity.

Critical voices have already warned against neglecting the power relations and inequalities generated in shifts towards increasing dependence on digital data, information and knowledge (Avgerou, 2002: 55; Hayes & Westrup, 2014: 20; Walsham, 2001: 56; Feldman and March, 1981; Orlikowski and Baroudi, 1991). But these voices have often been marginal in academic fields, and all but ignored in practitioner-oriented literatures. These concerns span Development Studies (DS), Information Systems (IS), Information and Communications Technology for Development (ICT4D), Knowledge Management (KM), and Knowledge Management for Development (KM4D). This thesis, in an effort to critique the technical or modernist promise, elucidates how data/knowledge-intensive practices can generate power relations, inequalities, omissions, and assumptions, and how they can marginalise different kinds of knowledge, knowing, and knowers (Blackler, 1995).

The vehicle used in the study to illustrate these seemingly silent, mundane problems and inequalities, is development NGO impact evaluation. Today, a plethora of data/knowledge intensive approaches, techniques, resources and prescriptions bombard development 2.0's

agents, NGOs, and evaluators, from big data to crowdsourcing, e-development to mobile for development (M4D), each involving complex architectures, specialisms, and new practices. Many of the specialisms are new, but some reach back into previous decades. The history of evaluation dates back to the 1950s (Barrow, 2000), and today's varieties focus heavily on the impact of policies, programs, projects and interventions (Roche, 1999; Wallace et al, 2006; Mohr, 1995). Impact evaluation is data/knowledge intensive, one process among many that contributes to the broader development 2.0 shift. It features diverse kinds of data - empirical, raw, processed, and cleaned. It also features information, knowledge, and evidence of diverse kinds: statistics, cells, formulas, methods, principles, graphs, stories, vignettes, narratives, voices, claims and representations of all kinds. Robust representations are demanded in an often vicious (Roche, 1999: 2) or volatile (Picciotto, 2012: 213) aid sector. And competing paradigms prescribe diverse practices from scientific to participatory, democratic, neoliberal, bureaucratic or autocratic evaluations (Picciotto, 2015; Duflo and Kremer, 2005; Macdonald, 1993; Norris, 2015).

Addressing the power/data/knowledge winners and losers, silences and sounds, assumptions and omissions in development impact evaluation is the focus of this study. Given that established institutional voices, from the World Bank, to the IMF, large International NGOs and so forth, are already loud, there is a need to find quieter voices that enunciate the realities of diverse evaluation practices, rather than global or institutional impact agendas. This need led to a study focus on small, local NGOs who were doing evaluation work. It turns out, that small NGOs doing impact evaluation, sometimes succeeding, sometimes struggling, tell us much about power, data, knowledge dynamics in development 2.0.

1.2   Motivation

Why is it important to acknowledge power in the everyday practices of professional data/knowledge work, in impact evaluation, or development 2.0 more broadly? One pragmatic justification is the well-documented waste, ineffectiveness and inefficiencies that accompany the spending of billions of dollars on aid each year (Moyo, 2009; Easterly and Easterly, 2006; Wallace et al, 2006; Yanguas, 2018). Another more critical justification is that one aid project always and instrumentally reproduces another in the same mold (Ferguson, 1990; Escobar, 1995/2011; Li, 2007) to keep the business of aid going (De Haan, 2009). Although important,

a third and fitting justification for this study concerns how development evaluation work is itself made up of impactful practices and governing relations, regardless of documented representations or published results. This study looks at the *impact of (doing) evaluation, rather than the evaluation of impact,* at what evaluations do, not what they say. The everyday practices of impact evaluation are not traditionally the subject of documentation, representations, reports, immaterial digital data or expert knowledge exchanges. However, this everyday evaluation work is where power relations and inequalities are generated.

This justification for the study implicates the critical importance *data/knowledge intensive work practices*. For example, how new information systems can generate organisational conflicts (Markus, 1983: 438), or how information behaviours, of individuals and organisations, are not limited to decision-making alone (Feldman & March, 1981: 175). Although such decision-making may be seen as intellectual, rational or immaterial, it requires diverse performances and practices, people carrying, hiding or showing information, organisations hoarding gluts of information to demonstrate capabilities, and many uses of information, such as justifying actions in retrospect. In ICT4D, Walsham (2001: 57–58) highlighted the importance of studying power in these diverse practices. Brigham and Hayes (2013: 27) described how a mix of technologies and conceptual models in e-development shape evaluation and perceptions of NGOs from a distance. They argue that what appear to be professional values and efficiencies can function in practice as Trojan horses for neoliberal exploitation. Such works suggest one cannot fully understand power in data/knowledge intensive work without looking at work practices.

In fact, ignoring power and practice obfuscates evaluation issues around politics, inclusion, effectiveness, efficiency and equality. Thus, it is vital to incorporate power and practice in understanding processes like impact evaluation. Data and knowledge in this view are particular, physical, unequal, social, political and material parts of processes, embedded in databases, heads, laptops and papers, circulating in aid organisations and networks. They are not immaterial, purely digital, apolitical or universal. Therefore, this study looks at the ways in which data and knowledge are material and narrow, as opposed to broad or immaterial in popular techno-determinist views and development 2.0 promises.

Furthermore, unequal configurations do not come from nowhere. On the contrary, they are related to wider issues often beyond the typical concerns of TIEK or development 2.0. The study incorporates such critical views of wider data/knowledge problems and practices. It

borrows from Foucauldian inspired literatures on governmentality (Foucault, 1979/2000; Ferguson, 1990; Escobar, 1995/2011; Li, 2007) and developmentality (Lie, 2015a; Ilcan and Philips, 2010; Hayes et al, 2017), related concerns around audit cultures (Strathern, 2000; Harper, 1998; Porter, 1996; Power, 1996; Townley, 1995), new public management and new development management (Hood, 1995; Kerr, 2008; Dar & Cooke. 2008), and critiques of information, data and knowledge models (Blackler, 1995; Tuomi, 1999; Lambe, 2011; Fricke, 2009; Beer, 2018; O'Neill, 2016). These literatures are important because they help us shed light on the politics of knowledge, evidence, evaluation and development today (Eyben et al, 2015; Gardner & Lewis, 2015: 179; Parkhurst, 2017). The study goes on to highlight the assumptions and omissions often excluded in TIEK literature, particularly normative scientific or participatory literatures. As we shall see, evaluation machines generate not just legible results showing scientific causes of impact, but also evaluation silences, changes to agents like small NGOs, and diffused power relations in evaluation networks.

The justification is further founded on three significant evaluation concerns. Firstly, there are many evaluation omissions, assumptions, inequalities and silences. This study highlights some of these, specifically how evaluation practices are changing, how agents such as small NGOs are being transformed, and how governing processes and power relations are also being re-configured (Hayes et al, 2017). Secondly the legible, audible evaluation results that feed development management, managerial decision-making, results-based management (e.g. OECD-DAC, 2002) and evidence-based policy (Parkhurst, 2017; Hammersely, 2005), are found not to be grounded, as routinely claimed, in linear and logical processes, nor objective, raw, empirical data, but rather in complex sets of responses to evaluation prescriptions, demands, histories and contingencies (Kelly, 2018). Thirdly, moves by global institutions such as the World Bank to set up as a Knowledge Bank (Thompson, 2004: 3), or by large state donors to fund knowledge for development (K4D) platforms and programs are predicated on the view that more evaluation, more impact analysis, more digital data, and more expert knowledge are undiluted goods that help to fix development problems. However, experts, platforms and technologies developed in the north risk arriving in the south as packages, "*loaded with an embedded virtual value system*" (Danowitz, et al, 1995: 28, in Thompson 2004: 2). Therefore, one must ask: what governing dynamics or audit cultures do these platforms and processes, particularly around impact evaluation, bring with them, in their wrapping? How do agents such as the small NGOs around the world respond to such imported packages? The study sheds light on such concerns.

Thus, the motivation and justification for the study is that during the shift to data/knowledge intensive process like evaluation, and development 2.0 more broadly, important literatures (e.g. ICT4D, KM4D, Evaluation and Development Studies) have not fully addressed the unequal, silent, power dynamics of such work. Why should researchers, managers, evaluators, NGOs, donors, beneficiaries, or activists be concerned about such lacunas? The answer is because our work contributes to the silences, politically and actively, our mundane, everyday, professional activities, from selling impacts in reports or on websites, to analysing data sets. The study advocates an alternative view of evaluation, which contrasts with expert led TIEK and widespread data/knowledge management models. This is because such prescriptions and models fail to identify or respond to the diffused power dynamics inherent in networks of evaluation practices.

1.3 Research approach and themes

The approach taken involved understanding what small NGOs did when they performed evaluations, describing their activities and analysing how their actions generated unequal data, knowledge and power relations. Cultural Historical Activity Theory (CHAT) (Engeström, 1987; Blackler, 1993; Karanasios and Allen, 2013) was selected as an approach that could support both engagement with NGOs, and critical analysis of their activities. The approach stemmed from a need to both consult and research, via an exchange of evaluation advice and research access between the partner NGOs and the author. This approach shares some similarities with action research. It is termed Critical Engagement in the study, outlined in chapter 4 and reviewed in chapter 7.

CHAT was used to collaborate with two NGOs in particular case settings (Mukute & Lotz-Sisitka, 2012; Kontinen, 2007). One NGO was an expert evaluator, and one novice. The first case involved distant interactions over telephone, conference calls and document exchanges between the author in the UK, the long-term philanthropic funder in London, and the NGO based in central India. The empirical work ran from August 2013 to April 2014, with interactions including six group conference calls, two phone calls, three co-designed evaluation tools, and ten evaluation data spreadsheets. The second NGO was based in northern Thailand and had multiple short-term funders. Empirical data included email exchanges, evaluation plans and analysis, telephone calls, skype conference calls, two field trips to Thailand for a

total of fifteen days which included six workshops. Workshops were led by the author to train the novice NGO on impact evaluation. This case ran from May 2014 to December 2016.

One early question guided the start of the first NGO case collaboration in 2013. This early question was: *How is development NGO impact evaluation shaped and performed in practice?* By early 2014, as the research matured in terms of the literature, empirical cases, and the theoretical and methodological frame, a more specific view emerged concerning NGO impact data/knowledge construction and the simultaneous generation of power dynamics. This resulted in a more precise problem space, a single overarching question, and three more specific research questions. These were:

- *Overarching question:* Does development impact evaluation include power relations?

Under this, the three specific questions were:

1. How is impact evaluation data/knowledge constructed at small development NGOs, in practice?

2. Are / how are power relations generated during impact evaluation data/knowledge construction at small development NGOs?

3. Can / how can power relations be addressed and responded to in impact evaluation data/knowledge construction practices at small development NGOs?

The approach, questions and first case NGO engagement shaped a number of core themes in the thesis. These were: the adoption of a socio-political perspective on evaluation; practice as the unit of analysis; and power as the object of study.

A first theme is the socio-political perspective on development impact evaluation, necessary for studying power relations. This contrasts with technical evaluation discourse which prescribes methods and ways to do evaluation better, with more rigour or robust data (e.g. Duflo and Kremer, 2005; Mohr, 1995). It also contrasts with more participatory perspectives on evaluation which foreground the emancipation of target participants (e.g. Chambers, 1994; O'Sullivan, 2007). In contrast to scientific truth, or participation intent, the diffused power, practice and politics of development evaluation (e.g. Eyben et al, 2015; Nauta, 2004; Wallace et al, 2006) are central to the study.

A second theme is the foregrounding of social or organisation practices. In order to understand the relationships between power, data and knowledge as they are constructed or generated in situ with study participants, the practice-oriented view was essential. Practice-based views focus on rich descriptions of everyday activities, of participant lifeworlds, actions, routines or performances, not logical deductions, economic models, or sociological categories (Reckwitz, 2002: 244-247). In this view, knowledge and data in cases are not considered as verbal, textual or statistical representations alone (Miettinen et al, 2009: 1312), but as elements within people's activities, as tools to be used to do something. This meant the research paid attention to what NGOs do with evaluation knowledge and data, rather than whether this data is scientifically valid, or that knowledge empowering.

Finally, a third theme is power. Power can be understood in diverse ways, as hierarchical, as restrictions or *power over*, as productive or *power to* (Clegg, 1989; Clegg et al, 2006), or as relational (Foucault, 1980; Escobar, 2011; Avgerou, 2002; Walsham, 2000: Blackler, 2011; Simeonova et al, 2018b). In this study, power is viewed as relational, as potentially productive and destructive, as relations that are enacted in specific practices, rather than as economic or cultural resources, held by some individuals, but lacking in others. Power in this study is also seen as inextricably related to data, knowledge, prescriptions and results. This view of power as part of professional evaluation practices means power is often silent or unobtrusive (Blackler, 2011), not recognised as power by participants. This contrasts with views of power as resistance, or part of open conflicts, (Korpelo, et al, 2004), or development battlefields (Long & Long, 1992).

As such, there is a mix of conceptual themes drawing on two schools of thought in the thesis. There are sector wide critiques of development drawing on governmentality and developmentality studies (e.g. Foucault, 1980; Ferguson, 1990; Li, 2007; Lie, 2015a; Ilcan & Philips, 2010; Hayes et al, 2017) significant in the literature and discussion chapters. And there are specific, descriptions of in situ NGO practices, agency, exchanges, transformations, and learning at the level of micro-activities, drawing on Cultural Historical Activity Theory (Engeström, 1987; Blackler, 1995; Kontinen, 2007; Karanasios, 2014; Kelly, 2018) particularly significant in the approach and NGO case chapters. These two schools of thought, one distant and one local, were key thesis components, bringing diffused power relations and local agencies into a productive conversation. These two vantage points, and their frictions, seeded the three thesis contributions.

1.4 Contributions

There are three principal contributions in the study. These are the "Impact Iceberg" diagnostic tool, "Audit 2.0" devices for understanding power and practice in data/knowledge intensive work, and "Critical Engagement" as an approach to exploring critical insights with research partners.

The Impact Iceberg is a big picture, diagnostic tool, which conveys sector wide impact evaluation challenges. The iceberg positions evaluation within wider sector data/knowledge flows in order to make explicit how current evaluations elevate some kinds of knowledge (e.g. efficiencies, cause/effect attributions, policy influence), yet submerge or silence other kinds of impact knowledge e.g. how NGOs are being transformed, how evaluation is becoming more technical, digital and market oriented, and how professional norms enact governing controls over evaluations. The iceberg is a contribution to the body of evaluation literature.

The second contribution comprises four novel auditing concepts, or Audit 2.0 devices. These devices contrast with "Audit 1.0", defined in the discussion chapter as audit culture perspectives which exert control and surveillance functions over those audited. In contrast, audit 2.0 focuses squarely on power and practice, to make these controlling functions visible and actionable. The four audit 2.0 devices are: data/knowledge chains and networks; the 6P data/knowledge sensitivities; the impact spectrum; and datamentality/datamateriality. These are conceptual contributions to evaluation and data/knowledge management theory and practice. They contrast with TIEK and widespread data/knowledge management legacy models, discussed in the theory chapter, that submerge power and practice. Tools, checklists and discussions can draw on the audit 2.0 devices to inform and scaffold more reflective, critical or sensitive evaluation processes.

The third contribution is Critical Engagement. This is a methodological contribution. Critical engagement borrows from CHAT (Engeström, 1987; Blackler, 1995), activist anthropology (Lewis, 2005), participatory development and critical studies, in asking how critical insights may be adopted, adapted, ignored or rejected in the wilds of development evaluation practice. The contribution describes key lessons from the study to inform tactics for brokering critical insights into power and practice with development sectors actors.

The three contributions collectively speak to a view of "Evaluation-as-Practice". Evaluation-as-practice contrasts with mainstream views of evaluation as scientific endeavours or participatory intent. Evaluation-as-practice offers alternative ways of thinking about impacts and data/knowledge products that circulate in development 2.0, seeing them not just as results or products, but as process and practices in which power and politics are key features and concerns. The literature, theory and approach chapters outline the foundations for each of the contributions respectively – the impact iceberg, audit 2.0 devices, and critical engagement. Each contribution is then fully articulated in the discussion chapter. The thesis contributions elevate power and practice back above the iceberg waterline as legitimate and uncomfortable concerns for those involved with development impact evaluation work specifically, and data/knowledge intensive work more broadly.

In terms of key thesis literatures and audiences, Thompson's (2008: 833) research agenda for development 2.0 is insightful because it highlights the need for interdisciplinary collaboration across Information Systems and Development Studies communities. Following Thompson, key audiences for the thesis and contributions are: Development Studies and Evaluation researchers and practitioners; Information Systems and Knowledge Management researchers and practitioners, particularly Information and Communications Technology for Development (ICT4D) and Knowledge Management for Development (KM4D); Cultural Historical Activity Theory researchers and others using approaches to social practice and power relations; Development sector practitioners, observers, or activists who wish to collaborate with, develop alternatives to, or contest evaluation processes and claims.

1.5   Study context: history and demands

The historical context of increasing data/knowledge intensity in development, NGO work, and impact evaluation, as well as examples of the contemporary demands for impact knowledge, are sketched below. These are indicative accounts, not exhaustive, but they show the growing intensities and set in motion the thesis focus on how evaluations are configured by more than scientific searches for impact causes or participatory zeal, as claimed in much TIEK.

Firstly, development data/knowledge intensity is accelerating in knowledge waves or fashions. A good example of this is from Leal (2007) who sees a knowledge landscape where successive

parades of new knowledge arise, often attempting to de-centre an apolitical economic narrative and promising to open up issues of power and participation. Leal's waves feature 1960s post-colonial community development, Cold War modernisation discourse, 1970s integrated rural development, 1990s participatory development, capacity building and human rights, poverty reduction in the 2000s, and today's sustainability discourse[2]. Leal concludes that the "*five glorious decades of development*" since the world wars illustrate how "*development has failed to fix underdevelopment and has maintained the status quo*" (Leal, 2007: 539-540). A key problem is that such knowledge fashions don't recognise one of their recent ancestors, empire (Kothari, 2005a). The fashions all feature a need-to-know the other, performed in distinct ways. For example, in Malinowski's colonial policy advice (Lewis, 2005); British government financed anthropology in East Africa (Peacock & Dosser, 1958)[3]; the establishment post World War Two of Development Studies courses and journals[4] (Sumner, 2006: 649); modern management's roots in imperial command and control (Prasad, 2003); how national economies have origins in colonial economics (Mitchell, 2002: 6); how research methods like action research can be tainted with colonizing intents (Cook, 2005); and how even technology design can reinforce colonial relations (Irani et al, 2010). However, as new fashions emerge and data/knowledge intensities spread, recognising such continuities means dealing with uncomfortable relations of power, data and knowledge.

Secondly, NGO data/knowledge intensity is historically growing, unequal, and concentrated in large, northern NGOs (De Haan, 2008: 69)[5]. During the period of World Bank structural

---

[2] Other examples include De Haan (2009: 69-88) who describes a 1950-1960s focus on "kickstarting economies", 1970s resource redistribution, 1980s structural adjustment, 1990s governance and poverty reduction, and a broadening results agenda since 2000; or Moyo (2009) who depicts the rebuilding of European infrastructure in the aftermath of World War Two, 1960s industrialisation, 1970s poverty reduction, 1980s stabilisation and structural adjustment, a 1990s buttress for democracy, and the "*present day obsession with aid as the only solution to Africa's myriad of problems*" (ibid: 10).

[3] See the British funded Colonial Development and Welfare Act 1925 that influenced East Africa operations (Peacock & Dosser, 1958); also National Income Accounting (Sumner, 2006: 645, footnote 6)

[4] Titles: Economic Development and Cultural Change (1952), Development (1957), Journal of Development Studies (1965), Development and Change (1970), World Development (1973) (Sumner, 2006: 649).

[5] There are over 47 types of NGO (Najam, 1996: 206; Lewis & Kanji, 2009: 9) and 85% of funding goes to large NGOs, not the invisible (Srinivas, 2009: 617) millions of small NGOs (Tvedt, 2006: 679; Els & Cartensen, 2015: 1-2; Richmond & Carey, 2005: 2; Lewis & Kanji, 2009: 2; Humanrights.gov, 2016; The Indian Express, 2010).

adjustment, the Washington Consensus, and the opening up of global markets in the 1980s (Williamson, 1990), NGOs were donor's "*favoured child*" because of their grassroots innovations (Bebbington et al, 2008). However, in the mid 1990s development refocused on state service provision, and governance (Banks & Hulme, 2010: 6). By the mid 2000s, criticisms grew concerning NGO roles in wider institutional structures (Gill, 1997; Bebbington, 2004: 732), how aims bent towards donors, compliance, reporting, service models, and away from community interests. Successful NGOs expanded, hiring professionals and specialists, applying business, marketing and capacity building models, often at the expense of participatory goals[6] (Flyvie & Ager, 1999: 1395; Korten, 1990). Such shifts towards compliance, professionalisation and business orientation mean that NGOs, particularly successful INGOs, have become data/knowledge intensive, through increased reporting, capacity building, advocacy, business management, global networks, communications and evaluation. Yet the data/knowledge inequalities and power relations within such historical shifts remain marginal and uncomfortable to deal with.

Thirdly, similar data/knowledge intensities have accelerated in impact evaluation itself. The discipline dates back to 1950s predictive techniques which evolved into formal approaches, such as Environmental Impact Assessment (EIA), Social Impact Assessment (SIA) and Social Cost Benefit Analysis (SCBA) (Barrow, 2001; Roche, 1999: 18). By the 2000s, there was an explosion in new approaches, from inductive or process accounts, to community dynamics, utilisation-focused evaluations and impact pathway analysis (e.g. Patton, 1997; Coffman, 2002; Estrella, 2000; Springer-Heinze et al, 2002)[7]. These featured alternatives to pre-set indicators, such as network analysis, outcome mapping (Earl et al, 2001), or mixed-methods (Picciotto,

---

[6] One example in terms of data/knowledge intensity and business orientation is the Red Cross, founded in 1863, originally to send rations, medicine and tobacco to soldiers (Bugnion, 2009: 3). It is now a $3 billion plus global federation. Similarly, CARE originally sent millions of packages containing "*staples such as butter, dried milk, canned meat, and sometimes even chocolate and chewing gum*" (CARE, 2016) to Europe after World War Two. Then, CARE saw poverty as a need for "*basic goods, services and healthcare*". Today, their understanding is much more conceptual: "*... poverty is often caused by the absence of rights, opportunities and assets, largely due to social exclusion, marginalization, and discrimination*" (ibid).

[7] Between the 1970s and the 2000s, a great many evaluation models evolved, for example: the "logframe" as managerial in the 1970s; participatory approaches in the 1980s (Chambers, 1997; Marsden & Oakley, 1991; Guba & Lincoln, 1989); a 1990s move away from older predictive or cost benefit approaches (Pingali, 2001) towards qualitative evaluations (Norrish & Sayce, 2006: 146).

2012). Importantly however, donors remained largely unaware of the benefits of participatory or qualitative evaluation (Norrish & Sayce, 2004: 152-153), instead favouring experimental designs, such as randomised control trials (RCTs) (Picciotto, 2012: 215).

Many have applauded the increasing method diversity and increasing use of ICTs and digital data, as part of "*new routes*" for evaluation (Norrish & Sayce, 2006: 146-7; also, Lefebve and Lefebve, 1996; Heeks, 2002a, 2002b; Stoll et al, 2002). However, Jerven (2014) estimated the Sustainable Development Goals (SDGs) will cost $15 billion to evaluate, and the Millennium Development Goals (MDGs) would have cost $27 billion. Such figures led Jerven (ibid: 16) to question the need for more and more digital data, and UN's data revolution claims (e.g. UN IEAG, 2014).

Accompanying these historic data/knowledge intensity trajectories are diverse contemporary demands for development impact data/knowledge. These are rarely seen as problematic in technical evaluation discourse. Seven examples are indicated below, and each suggests impact evaluation involves more dynamics than scientific or participatory claims alone.

A first set of demands is direct results. Since the 1990s, NGOs have had to show increasing results, outputs, outcomes and impacts (Roche 1999; Wallace et al, 2006; Norrish & Sayce, 2006). Such results focus on "*accounting for funds and highlighting achievements*" (Wallace et al, 2006: 40). Governments also mandate evaluations (e.g. DFID, 2003: 139-140; 2002).

Second, demands follow the emergence of technical development discourse in the 1950s-1960s (Escobar, 2011; Kothari, 2005a; 92), and increasing subspecialisms today, from UN cluster systems[8] to specific sub-field evaluation forms, such as evaluation in Communications for Development (C4D) (Lennie & Tacchi, 2013), Gender Responsive evaluation (UN Women, 2015; UNDP, 2006; Bustelo & Espinoza, 2015) or ICT4D (Heeks & Molla, 2009).

Third, demands concern how impacts must align with specific project and programme aims and goals. Project data and knowledge are routinely benchmarked to specific funder niches, global standards such as the 17 Sustainable Development Goals (SDGs) and their

---

[8] UN cluster systems prescribe fine grained specialisations, separating emergency management tasks into recovery, education, telecommunications, food security, health, logistics, nutrition, protection, shelter and water, sanitation and hygiene (IASC, 2006; UN OCHA, 2016; Puri, 2015: 22-23)

accompanying 169 subtargets (United Nations, 2016a, 2016b), or to any of the thousands of predefined, often highly complex, data indicators available (e.g. World Bank, 2016a, 2016b). Such goals shape impacts, despite the lack of evidence on the ability of targets to guarantee progressive impacts (Stiglitz, 1999: 587-8; Sandbrook, 2011; 421; Puri, Sahay & Lewis 2009; Walsham, Robey & Sahay, 2007).

Fourth, demands to manage risk. UK government aid budgets have been used to manage national risks such as refugees (BBC News, 2015) or to deliver lucrative trade agreements (Swinford, 2016). Leading aid agencies also fight global health risks, such as Ebola (USAID, 2016). Further risks include terrorism, corruption, climate threats, and military conflicts. One problem that results is that more data, information and knowledge collected to manage risk can counter-intuitively lead to new risks and less control (e.g. Ciborra, 2006: 1354)[9].

Fifth, technological demands relate to the increasing importance and visibility of information systems in development (Avgerou, 2002; Walsham & Sahay, 2006: 7). This involves the need to adopt new software systems (e.g. Kenny, 2014: 12), e-development (Brigham & Hayes, 2013), mobile for development (M4D), big data for development (BD4D), or other diverse ICT innovations in development (Karanasios, 2014; Walsham, 2001), and thus in evaluations.

Sixth, narratives demonstrating participatory impacts have also become more important in aid communications. Participation conveys moral concern and democratic engagement, and adds value to positive media messages (Kontinen, 2007; Wallace et al, 2006: 31). Even those who call for qualitative, contextual data, human centred design or community empowerment, are therefore demanding more evaluation data/knowledge[10].

Seventh, and most significantly, are demands to show successes in aid funding markets. After the turn to market liberalism to deliver development in the 1980s (Mohan & Stokke, 2000:

---

[9] A further risk problem is how professionals and resources are mobilised to fix "over there" challenges when risks are identified as belonging to developing nations. For Dean (2010: 202), such operations leads to reconfiguring societies as "*quasi-markets*" using risk management, mitigation, and expertise. The role of such expertise may not be to empathise, cure or care, but to measure and reduce risk (Miller & Rose, 2008: 107-108).

[10] There is little consensus on defining participation (Long, 2001), however a key concern is whether it is authentic and puts the poor first (Chambers, 1994), or is faux, hiding a soft imperial control (Schuurman, 2009: 845; Cooke, 2003). Either way, NGOs face demands to include participation data, evidence and knowledge in impact representations.

248), NGOs had to "*learn the new paradigm*" (Wallace et al, 2006: 163), even though observers suggested that NGOs had become "*too close for comfort*" with funders (Banks, Hulme & Edwards, 2015: 707; Hulme & Edwards, 1996). This means NGOs remain under pressure to demonstrate impacts to funders in order to survive, often on overtly commercial terms (Davies, 1997a: 616; 1997b: 5).

Demands on NGOs
for impact evaluation data / knowledge in the international development sector
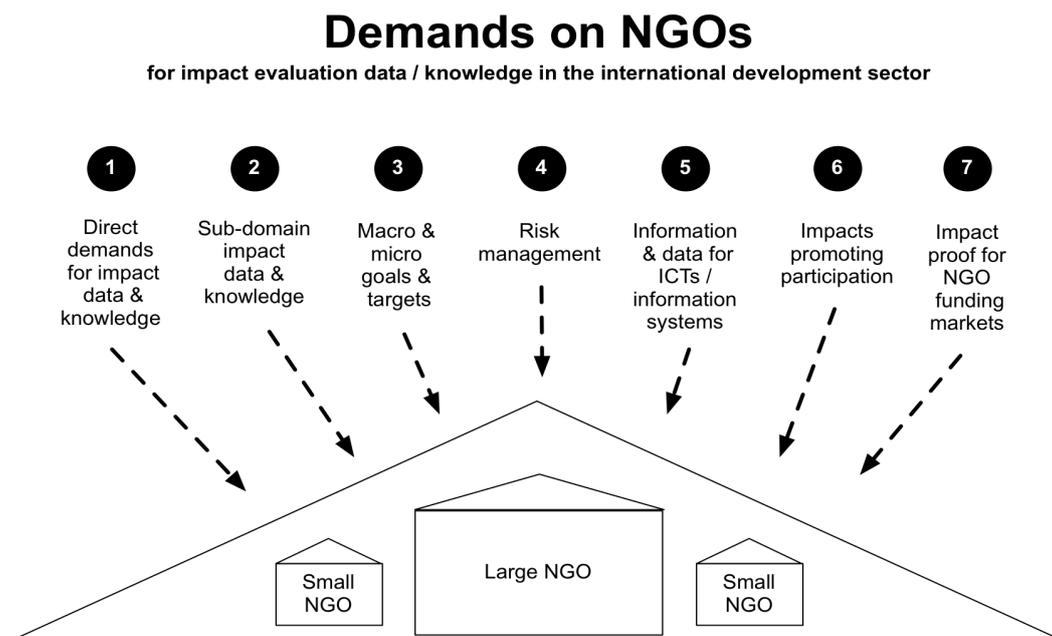
Figure 1.1: Development 2.0 demands for impact data/knowledge

These diverse sets of sector demands all contribute to increasing development 2.0 data/knowledge intensity. However, they do not foreground or provide responses to the unequal power relations that are generated by the emergence and acceleration of such data/knowledge intensities.

1.6   Structure

This section orients the reader to the structure and main points of each of the thesis chapters.

Chapter 1: Introduction - presents the main focus of the thesis on power/data/knowledge relations in NGO impact evaluation, as part of today's development 2.0 landscape. The thesis

motivation, approach, core themes and research questions are described. Key audiences and final contributions are outlined, and the historical and contemporary context is sketched, before this thesis structure preview.

Chapter 2: Literature review - describes two contrasting areas of development impact evaluation literature. The first literature domain is normative "technical impact evaluation knowledge" or TIEK. This literature concerns goals, methods and prescriptive discourse to support aid organisations in performing impact evaluations. The second domain concerns critical perceptions on development evaluation. This literature contrasts with TIEK, and deals with problems related to prescriptive discourse, how evaluations, NGOs and power relations unfold in practice and change over time. These issues are termed "critical configurations".

Chapter 3: Theory - builds on problems identified in the literature around the lack of theoretical grounding for power, data and knowledge relations in TIEK. The "DIKW Pyramid", a widespread and implicit model of data, information, knowledge and wisdom is found to implicitly underpin contemporary evaluation models. An alternative view of impact data/knowledge relations involving products, processes, power, political participation and practice, the "6P Sensitivities", is proposed because DIKW and TIEK elide the diffusion of power and practice in evaluation networks, aid chains and critical configurations.

Chapter 4: Approach and Design - outlines a critical and engaged research stance, explains the rationale, and justification for using Cultural Historical Activity Theory (CHAT) to operationalise critical engagement in the study. CHAT's framework for engaging research partners and for critiquing power and practice via activity systems, contradictions and temporal chains of activities is described. The chapter explains the research design, why the author selected two NGO cases for the study, access to research partners, and how data was analysed. The result is a comparative and qualitative CHAT study of two NGO case collaborations between 2013 – 2016, and a framework for exploring the ground between scholarly critique and pragmatic engagement.

Chapter 5: NGO Case 1 Rural India - features the first NGO, who are expert evaluators with a long-term philanthropic funder, and who run female farmer livelihood programs. The sequence of evaluation activities is described in the empirical section, and the analysis section articulates the CHAT activity systems, contradictions and temporal activity chains. Two kinds of impact are evident in the case, one local, unclear, and changeable kind of impact closely

associated with farmer lifeworlds (Impact-1), and one digitally documented and represented kind of impact deployed in marketing efforts (Impact-2). Power relations are generated as impact data/knowledge is incrementally edited along a temporal chain of evaluation activities, with illegible local voices and contexts edited out or submerged, and managerial, marketing and data management concerns expertly elevated by the NGO and philanthropy.

Chapter 6: NGO Case 2 HTSG Thailand - features the second NGO as a novice impact evaluator, lacking guidance and resources. The empirical section describes the NGOs learning, struggling and transformations. The analysis section articulates the activity systems, contradictions and temporal chains. However, in this case two contradictions are pertinent, between two kinds of impact as in the first case, and two forms of the NGO. HTSG are transforming from an old version of themselves into a new data/knowledge intensive NGO. The temporal chain illustrates the changes, elevations and submerging of different notions of impact, and the transformation towards, but not reaching yet, impact expertise. HTSG's confusion and lack of data/knowledge capacity are key case results.

Chapter 7: Discussion – this chapter builds on the CHAT analysis and results concerning how power is generated in small NGO evaluation data/knowledge construction activities, to arrive at three contributions. The first contribution is to the evaluation body of literature and concerns the big picture narrative of the Impact Iceberg, where certain impact results become legitimate, and others illegitimate, silent, submerged under the waterline. The governmental and developmental processes in todays' development 2.0 landscape identified in the literature are revisited, and seen to configure the iceberg. The second contribution is conceptual and pragmatic, labelled as Audit 2.0 devices, and challenges TIEK models and the DIKW legacy. The devices are: firstly, data/knowledge chains and networks; secondly, the 6P Sensitivities; thirdly, the Impact Spectrum; and fourthly, the contrasting notions of Datamentality and Datamateriality. The third contribution is Critical Engagement. This is reviewed as a way of taking critiques of power into the wilds of evaluation practice. Together, the contributions are advocated are elements in seeing evaluation-as-practice, which contrasts with scientific, business pragmatic, technology-centric, or participatory views of evaluation.

Chapter 8: Conclusion - reviews the key goals and core arguments of the study, describes a series of limitations, and lists future research options. It recaps how the study has articulated *the impact of evaluations* in data/knowledge intensive development 2.0, rather than the normative "how to" foregrounded in technical evaluation discourse. Implications for

evaluation, data/knowledge models and development policy and practice are re-iterated at the end.

2 Literature: From technical to critical impacts

2.1 Introduction to the review

The review explores two literature areas around development impact evaluation. The first body of work is technical impact evaluation knowledge, termed "TIEK". The second is broader socio-political perspectives, termed "critical configurations". These two literature domains view impact evaluation differently, meaning that surveying each is important in understanding impact power/data/knowledge relations. These relations span the politics of evidence (Eyben et al, 2016), the politics of knowledge (Gardner & Lewis, 2015: 179; Escobar, 2012: xvi) and questions by ex-World Bank analysts about vested interests and evaluation capture (Picciotto, 2015: 152) or about the vast distances between evaluation stakeholders (Bamberger et al, 2010: 2). They also cover the concerns of seminal evaluation scholars, such as Weiss and Macdonald on opposite sides of the Atlantic, who argue that evaluation is always "*enmeshed in politics and inevitably political*" (Norris, 2015: 135). Such broad concerns dictate a literature review that covers not just the *how to* of technical evaluation capability, but also the critical views which elucidate more clearly the power/data/knowledge relations. These two literature sections show not just *how to do evaluation,* but also *what doing evaluation does*.

Firstly, section 2.2 describes TIEK through illustrative examples of prescriptions, specifications, evaluation methods, and impact results as part of organisational information flows. This literature concerns how to evaluate impact, how to evaluate well, and how to define what impacts are or can be (e.g. Mohr, 1995; Roche, 1999; Earl et al, 2001; Picciotto; 2012; 2015; Duflo & Kramer, 2005; Davies & Dart, 2005; Chambers, 1994; Pattern, 2014; Heeks & Molla, 2009). TIEK is a technical and professional discourse which incorporates a wide range of normative or prescriptive claims. It concerns approaches, methods and models, on how to ensure methods are rigorous and data is robust. This literature is a resource from which development sector organisations draw on, in order to know how to make legitimate impact data/knowledge, when to do it, who can do it, and how to define it. The section concerns what works in impact evaluation, for impact evaluators.

Secondly, section 2.3 explores how development impact evaluation is more broadly configured (e.g. Wallace et al, 2006; Norrish & Sayce, 2004; Anderson et al, 2004; Mosse, 2004a; Eyben et al, 2014; Gardner & Lewis, 2015; Banks & Hulme, 2012; Sumner, 2006; Ilcan & Phillips, 2008; Long & Long, 1992; Lie, 2015; Ferguson, 1994). These critical configurations speak not to the normative prescriptions of TIEK, but its circulations, the affects and drivers, the consequences of using, adopting, applying and mobilizing TIEK in development. The configurations concern not how best to evaluate, but what happens when we do evaluate. It is about the impact of evaluation in contrast to TIEK's focus on the evaluation of impact. The section covers configurations related to evaluation itself, to NGOs, and to the political governing processes of impact evaluation.

Finally, section 2.4 sharpens the study's focus on evaluation practices. It firstly juxtaposes the historical data/knowledge intensive context, with the many impact demands, and the TIEK prescriptions, which together constitute a generic information centric meta-model of evaluation, as inputs and outputs aimed at key development decision-makers. It then presents implications from this observation, which speak to the omissions of TIEK, how we view evaluation as scientific, participatory, critical, or something else, and why the study of impact evaluation power/data/knowledge requires not just scholarly critique, but engagement in evaluation practices.

2.2 Technical impact evaluation knowledge (TIEK)

The word impact may convey a brutal image of a "*hammer hitting a surface*" (PARC, 2004; Norrish & Sayce, 2006: 1), but impact evaluation is much more subtly and technically shaped through academic discourse and industry publications. Roche's important work with development NGOs in the late 1990s, led him to consider different definitions of impact assessment, foregrounding systemic analysis and the attribution of causes as central concerns: "*Impact assessment is the systematic analysis of the lasting or significant changes - positive, negative, intended or not - in people's lives brought about by a given action or series of actions" (Roche, 1999: 21).*

In contrast, Earl's Outcome Mapping, published by the Canadian International Development Research Centre (IDRC), highlighted learning and accountability as central (Earl et al, 2001). The International Food Policy Research Institute (IFPRI, 2002) based in Washington D.C. prescribed eight considerations: *a) scale, attribution and time, b) supply versus demand approaches, c) importance of surprise, d) indicators, e) methodological issues, f) time lags, g) ex ante assessments, and h) ex post assessments*

The Organisation for Economic Co-operation and Development (OECD), Development Assistance Committee (DAC), a global body promoting development evaluations[11], published the influential "*Glossary of Key Terms in Results Based Management*" in 2002, and the "*Principles for Evaluation of Development Assistance*" in 1991. Their definitions became the "*most widely shared*" and the "*best*" according to Stern (Stern et al, 2012: 5; Stern, 2015: 4). DAC defined impact as: "*positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended" (OECD DAC Glossary, 2002).* The principles stated:

> "*An evaluation is an assessment, as systematic and objective as possible, of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability. An evaluation*

---

[11] DAC became part of OECD on 23 July 1961. It is an international forum of large aid funders, with 29 members (http://www.oecd.org/dac/dacmembers.htm). The World Bank, IMF and UNDP act as observers. See http://www.oecd.org/dac/developmentassistancecommitteedac.htm

should provide information that is credible and useful, enabling the incorporation
of lessons learned into the decision-making process of both recipients and donors."
(OECD DAC, 1991: 5)

The principles also stipulated that evaluations should: "improve future aid policy, programmes and projects through feedback of lessons learned and provide a basis for accountability, including the provision of information to the public" (OECD DAC, 1991: 5)

The International Initiative for Impact Evaluation (3ie) is another prominent voice in development impact evaluation, founded in 2008 and funded by the Bill & Melinda Gates Foundation, the William & Flora Hewlett Foundation and UKAID (3ie, 2016). 3ie have traditionally focused their definition on the use of experimental research designs to attribute changes to program interventions. They state impact evaluations constitute: *"A study of the attribution of changes in the outcome to the intervention. Impact evaluations have either an experimental or quasi-experimental design"* (3ie Impact Evaluation Glossary, 2012: 5).

These formations are primarily concerned with two key issues: firstly, the need to produce impact data and knowledge, and secondly how to primarily use impact data and knowledge to inform organisational decision-makers. Underlying these definitions is a view of evaluations as delivering knowledge as "*products*" (Mosse, 2004a: 81), as granulated chunks of data, information or knowledge, passed on to decision-makers such as policy makers, programme managers, and sometimes the general public.

Along with defining impacts, prescribing evaluation specifications is another way of framing how to plan, design or conduct robust evaluations and collect empirical data. Specifications speak to the number of people in evaluation teams and required credible evaluator skills and professional experience (UNDP, 2002: 55; Roche, 1999: 61). Recent studies have attended to who is involved and to what extent, such as in how to include commissioners, managers or beneficiaries in evaluation processes (Stern, 2015; Groves, 2015).

Having the right team, with the right experience, and appropriate levels of independence is seen as crucial. For example, evaluators may be external consultants, interdepartmental evaluators, or internal assessors (Picciotto, 2013). O'Sullivan (2004) and Perloff (1979) see the independence of evaluators on a spectrum between at one end scientific objectivity and at the other more engaged and therefore less independent, but with deeper access and understanding of contexts and stakeholders. In this sense, people, teams, skills, and independence levels are

important evaluation specifications that frame what kind of approach to take, what data becomes available, and how legitimate the data may be.

Evaluation thinking and concepts are also targets of technical discourse. Publications prescribe requirements for making evaluations empirically robust, and for applying theories of change (Vogel, 2012: Anderson, 2006). TIEK includes guides for conducting formative research to understand context, establishing counterfactual scenarios, contrasting baselines with end line studies, defining beneficiaries, specifying outcome variables, and deploying quantitative methods (e.g. White, 2009: 7; Puri et al, 2014: 6).

Furthermore, impact evaluation includes temporal prescriptions too, with specific practices deployed at different times or in different sequences. Feasibility and predictive studies are done ex-ante, before interventions commence (Roche, 1999; Barrow, 2000). Baseline studies are done either before or at the beginning of projects to capture inception data for later comparisons (Roche, 1999; Puri et al, 2015). Monitoring is continual, through a project lifecycle (Puri et al, 2015). Mid-term impact evaluations are formative and enable reviews for changes to plans and indicators. Ex post facto evaluations are summative, done at the end, soon after, or even years after project completion.

It may appear that such technical specifics relate to a *methodological mythology* (Norrish & Sayce, 2004), however these complex specifications constitute real results, interactions and effects. For example, one temporal aspect is the sensitivity of indicators to often rapidly changing measures such as in humanitarian emergencies where dehydration or diarrhea can become a matter of life and death for a child (Puri et al, 2015: 21). Technically therefore, the right activity at the right time with the right data aligns with prescriptions and contributes to real impacts on real lives. As such, people, teams, experience, independence, concepts, thinking, approaches, and timing are all important technical elements that have consequences, results and influence evaluation configurations.

The boundaries of what is or is not considered to be legitimate impact evaluation are often re-iterated in technical discourse, in relation to organisational processes and to other modes of evaluation. Firstly, impact evaluation as an organisational process informs other activities, such as planning, program and project management Wallace et al, 2006), learning and change operations, funding decision-making, or policy development (Roche, 1999). Although these

practices overlap with impact evaluation, they are rarely technically analysed as part of evaluation prescriptions or guidelines. They are conceived of as related technical sub-domains.

Since the emergence of environmental impact assessment (EIA) and social impact assessment (SIA) in the 1950s-60s which preceded development impact evaluation today (Barrow, 2000; Roche, 1999), the field of evaluation has grown and fragmented into a variety of sub-domain technical expertises. Established forms of evaluation used in development today include: Environmental Impact Assessment (EIA), Social Impact Assessment (SIA), Impact Assessment (IA), Monitoring and Evaluation (M&E), Monitoring, Evaluation and Learning (MEL), Monitoring, Evaluation, Learning and Accountability (MELA or MEAL), Monitoring, Evaluation & Communication (M&E&C), Transparency & Accountability (T&A), and even Performance Management (individual, unit and organisational levels).

Each evaluation sub-field has definitions, technical specifics and normative methods. For example, many researchers would not consider monitoring as encompassing the same concerns as evaluation, such as program efficiency, effectiveness, impact or sustainability (Puri et al, 2015). Roche distinguished monitoring and evaluation from impact assessment according to timing, depth of analysis and specificity (Roche, 1999: 26). The World Bank separates monitoring, operational evaluation and impact evaluations wherein the former two are elements in continual feedback on project goals, and the latter is concerned with cause and attribution of outcomes and impacts (Khandker et al, 2015: 8-16). Hyman & Dearden (1998: 275) recommend internal monitoring is best performed by field staff, whereas impact evaluation is best reserved for external data collectors. Perrin (2012) separates monitoring, evaluation and impact evaluation into three processes, each differing in terms of timescales, objectives, data required, and knowledge produced. However, Perrin does acknowledge that monitoring and evaluation serve impact evaluation over the long-term, through a sharing of data and results.

In summary, the reason for detailing example technical specifications, definitions and referencing the boundaries of TIEK is to show how prescriptive discourse frames understandings of what impacts and evaluation technically can or should be. The extent to which prescriptions are in fact adhered to, or deliver what they promise, is a separate question concerning how evaluations might unfold in practice.

Perhaps the most important aspect of TIEK is the expanding range of approaches and methods used to plan and implement evaluations. Methodology is a hotly contested area for prescriptive

debate and knowledge production. As alluded to in chapter one, there has been a split between more experimental/quasi-experimental designs and more qualitative or participatory designs (Bamberger et al, 2010, Picciotto, 2012, Chambers, 1994). The use of quantitative methods grew extensively after the millennium:

> *"In the last decade there has been an explosion of quantitative impact evaluations of program interventions in international development. This has been driven both by trends within academia and pressure from international organizations like the World Bank, and has culminated in efforts to adopt the standards and methods of bio-medical clinical trials in making knowledge claims about the effectiveness of particular interventions." (Bamberger et al, 2010: 1)*

Picciotto (2012: 216) argued that the popularity of experimental designs and randomized control trials (RCTs) and their status as a "*gold standard*" promoted by evaluators such as Duflo & Kremer (2005), has actually been a short-term trend bucking a longer trend of increasingly qualitative data and contextual evaluation. In the 1990s, Chambers (1994: 953) pointed to this longer trend moving away from "*extractive survey methods*" towards participation, suggesting that participation's time had finally come. Proponents of experimental designs had even asserted the need for "*qualitative grounding of quantitative evaluation*" (Campbell, 1974) and seminal evaluation scholars contributed to the longer-term trend by foregrounding stakeholder engagement, theory driven evaluation and rejecting experimental designs for all but the simplest of discrete queries (Picciotto 2012: 219). Unsurprisingly then, there has been a great increase too in the variety and spread of contextual and participatory evaluation methods and approaches, although donors and commissioners often prefer experimental designs (Patton, 2014a; 2014c; Norrish & Sayce, 2004: 153).

The increasing breadth and depth of evaluation's methodological discourse is accompanied by confusion too. "*The evaluation tree is still growing*", according to Picciotto (2012: 220), and at a rapid rate, as evaluations are deployed by governments, international NGOs, local NGOs, corporations and philanthropic organisations (Cashmore, 2010). There are regular calls for more rigour and more appropriate methods (e.g. Cashmore et al, 2010; Puri et al, 2014; Stern et al, 2012), despite the ongoing methodological confusion. DFID's own publications acknowledge the confusion and complexity of contemporary impact evaluation discourse:

"This is a study report dealing with difficult methodological and theoretical challenges faced by those who wish to evaluate the impacts of international development policies. It is in parts difficult to read as it was also difficult to write!"
(Stern et al, 2012: 38)

Mathison (2004) considers technical confusion to be part of the search for agreement and consensus in the professional discipline. In the Evaluation Encyclopedia, she aims to clear the ground, listing 21 quantitative methods and 36 qualitative methods. Quantitative methods feature benchmarking, correlation, multitrait-multimethod analysis, regression analysis, pre-post design, time series analysis and surveys. Qualitative ones include unique case analysis, grounded theory, fieldwork, checklists, content analysis and ethnography (ibid).

It is important not to under-estimate the growing diversity of methodological discourse, whether formally labelled as methods, approaches or other terms such as frameworks, models or techniques. There are literally hundreds, if not thousands[12], of models, methods, techniques, approaches and other *how to* step by step guides and prescriptions for evaluation currently in circulation.[13].

Debates in the professional discipline concern which model, method or approach facilities which kinds of results for which decision-makers, what the validity, threats, strengths and weaknesses of each are in terms of scientific rational, practical application and participatory authenticity. The range of context rich methods includes systematic collection of anecdotes in order to compile "*anecdata*" (Davidson, 2014; Patton, 2014b) and case studies (Herron & Quinn, 2014). Mohr covers a number of statistical methods for evaluating quantitative data (Mohr, 1995), and on top of this is the gold standard of RCTs (e.g. Duflo & Kremer, 2005). Chamber's rural appraisal (1994), Holland's participatory statistics (2013), and O'Sullivan's collaborative evaluations (2004) present participatory options. Empowerment based evaluations also seek to capacity build participant skills although there is a trade off in terms

---

[12] On discussing the diversity of methods at one academic conference, a participant disagreed with the authors claim that there were hundreds of models and methods in circulation. Their point was that hundreds is an understatement and there are nowadays thousands of evaluation models, methods, techniques and prescriptions.

[13] Appendix 5 shows a table of methods and techniques, from statistical benchmarking, multivariate regression analysis, or binary probit regression, to contextual forms such as Appreciative Enquiry or Realist Evaluation.

of time and resources regarding how scientifically robust impact data, knowledge and narratives can be (O'Sullivan, 2000).

In summary, technical specifications, definitions, boundaries, approaches and methods frame a rapidly expanding technical discourse for mastery, prescribing ways of doing and ways of thinking about impacts and evaluation of those impacts. Yet, what is also important is how methodological knowledge is deployed in order to produce results as data, information and knowledge for certain kinds of stakeholders working in certain sector specific organisational functions and capacities.

Results-based management (RBM) was introduced into donor and UN agencies in the 1990s (van den Berg, 2004: 67) and became popular in the sector, contributing to the increasing demands for evaluations and impact evidence, data and knowledge for organisational decision-making. For example, the UNDP Handbook on Monitoring and Evaluating for Results (2002) describes the emphasis on evaluation results, and how results-based management helps organisations ensure that "*processes, products and services contribute to the achievement of clearly stated results*" (UNDP, 2002: 9). The handbook goes on to explain how aid organisations require "*evaluative evidence*" and the management, organising, packaging and "*disseminating of information and knowledge on time to the right decision makers*" (ibid: 77). Evaluative evidence is cited as important (Phillips & Phillips, 1997) to show "*what's working*" (ibid: 204), "*potential success*" (page 154), "*impact*" (197), "*value*" (450), "*program success*" (13), "*tangible benefits*" (434), "*improvement*" (230), "*knowledge acquisition*" (161), "*evaluation success*" (49), "*learning*" (169) and "*future needs*" (280). These and the many other benefits of evaluation, such as ensuring accountability, supporting funding strategies, improving effectiveness and efficiency or aiding reporting, have all been emphasized before in the literature (for example see Roche, 1999; Davies, 1997a; Norrish & Sayce, 2004; Duncombe, 2009; Banks & Hulme, 2012: 17-18; Barr et al, 2005; Burger & Owens, 2010). How organisations obtain these results depends largely on an information centric evaluation meta-model. How they are used depends on functional requirements inside organisations.

Organisational results, circulated as data, information and knowledge in reports, white papers, email messages, or other media, and utilized in project and management decision-making practices, such as in meetings or presentations, rely upon an evaluation input-output information model. Much impact evaluation literature has tended to view assessments or evaluations as classifiable in terms of an objective measure where inputs and outputs can be

compared and assessed (Duncombe, 2009; Heeks & Molla, 2009; Brigham & Hayes, 2013: 128). This view fits the rational management approaches used in aid bureaucracies, large donors and large INGOs (Wallace et al, 2006: 31; Dar and Cook, 2008).

The classical information-results perspective can be seen in Trochim's (2006) Planning-Evaluation Cycle for managerial and administrative processes, which features planning and evaluation phases. The evaluation phase consists of formulating evaluation questions and hypotheses, conceptualizing how to measure outcomes, designing the process and utilizing the results in management or decision-making (ibid). Roche describes impact evaluation as an ongoing set of processes that captures inputs such as beneficiary input, within a project lifecycle spanning design, assessment, implementation and evaluation (Roche, 1999: 31). The UN Women's handbook on How to Manage Gender Responsive Evaluation (2015: 7) describes a results-based approach to evaluation, embedding the input-output model of evaluation within the results-based management framework, to bring together planning, monitoring and evaluation, whilst "*ensuring evidence for decision-making, learning and accountability*" (Figure 2.1).

Figure 1. Steps in the results-based management cycle
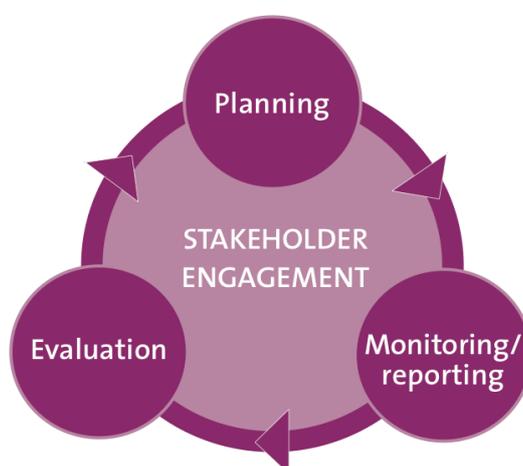
Figure 2.1: Evaluation in results-based management (UN Women, 2015: 7-8)

The Gender-Responsive Evaluation handbook states that results are "*outputs, outcomes and impacts*" (ibid: 8; see Figure 2.2). Activities are defined as technical inputs, outputs are changes

in participant skills or services, outcomes are mid-term changes, and finally impacts are long-term sustainable changes, which support UN MDGs, SDGs, or related national goals (ibid: 9).

Figure 2. Programmatic causal chain

Inputs → Activities → Outputs → Outcomes → Impacts
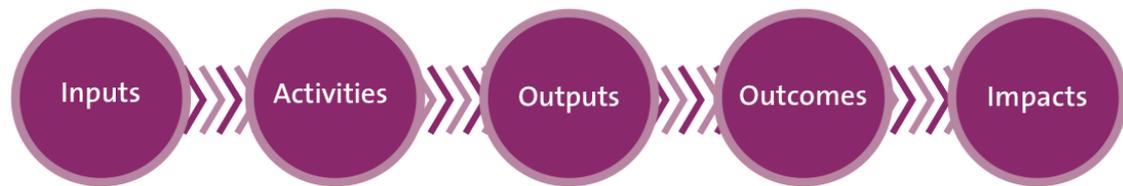
Figure 2.2: Impact chain from inputs to impacts (UN Women, 2015: 8)

In order to promote more evaluation results feedback and sharing with beneficiaries, Groves (2015) proposes an analytical model of evaluation (Figure 2.3). This model is the evaluator's implementation process. It shows a logical process of evidence gathering, data analysis and finally dissemination of results in knowledge products.

Design → Evidence Gathering → Data validation and analysis → Dissemination and Communication

Figure 2.3: Groves' model for impact evaluation results sharing (2015: 2)

Both the logical model of "*inputs-to-outputs-to-impacts*" and the rational process from "*design to analysis-to-communication*" are notable, as they illustrate the underlying linear, scientific and targeted rationality underpinning TIEK. Gardner & Lewis (2014: 38) refer to Wallace et al to sum up how such models envisage linear and logical target results.

> "*A culture of managerialism where change must first be envisaged, then detailed, described and planned for. Once implemented, projects must demonstrate the achievement of pre-set results, which must be measured and reported on in*

quantitative terms. Change is understood as linear, logical and controlled following theories of change based on cause –and-effect-models" (Wallace et al, 2013: 4).

Quantitative or qualitative, depending on the evaluation objectives, different methods are recommended to produce different target results in order to satisfy such evaluation logics and goals. For example, causal proof of impact through RCTs may serve issues of accountability and effectiveness (Duflo & Kremer, 2005). Local project learning, relationship building, and understanding may best be served by more contextual approaches, such as Most Significant Change (Davies & Dart, 2005). Whatever the specific objectives, approaches or methods deployed, there is an argument that results are the overriding deliverable. Results are seen as important to bridge "*vast distances*" between decision-makers and participants. This bridging requires scientific or participatory rigour to defend assertions of impact truth and to build decision consensus. Bamberger et al (2010) argue that there is a strong case for using mixed-methods to produce results that bridge this gap:

> *"However, given that a central challenge in international development is that the decision makers (development economists included) are in the business of studying people separated from themselves by vast distances – social, economic, political and geographic – there is a strong case for using mixed methods to both help close this distance and to more accurately discern how outcomes (positive, negative, or indifferent) are obtained, and how any such outcomes vary over time and space (context)." (Bamberger et al, 2010: 1)*

However, the concerns of data/knowledge intensive decision-makers, managers and policy developers, referred to earlier in definitions by the OECD DAC or 3ie and including the need to scale projects, define impact attribution, or perform systemic analysis (OECD DAC, 1991; 3ie, 2012; Earl et al, 2001; Norrish & Sayce, 2004; Roche, 1999: 21) are very different from the concerns of target participants and communities. Picciotto (2012) argues that the strong results-based agenda is due to the aid industry actually being in relative turmoil, faced with debt, skepticism and more recently economic austerity. All these risks contribute to an ongoing need for aid agencies to demonstrate rigour and scientific certainty in terms of impacts. In Picciotto's view, these dynamics keep quantitative methods afloat and in demand as they promise what appear on the surface to be robust results, through numbers and statistics. However, the larger take away point in the information centric framework is that results need

to be defensible within organisational decision-making practices, and to respond to impact data/knowledge sector demands, regardless of method.

To sum up, this literature has revealed a range of technical concerns: from impact definitions, to evaluation specifications, an evaluation linear logic, and how knowledge products are targeted at decision-makers in organisations. Yet, what is missing concerns what such a growing discourse does, what kinds of evaluation and evaluators it produces, and how methods and results alter power dynamics between aid stakeholders, bureaucracies, experts and communities. To understand these missing parts it is necessary to look outside TIEK, to other literatures which focus on socio-political perspectives, inequalities, and power relations.

2.3 Critical configurations

The historical trends, sector specific demands and TIEK discourse contribute to configurations in the sector that reach beyond normative issues and topics prioritised in the TIEK literature. Therefore, this section looks at such wider, critical issues and the socio-political concerns that surround technical evaluation. The section covers three particular areas. Firstly, critical configurations around how impact evaluations are practiced in the sector. Secondly, configurations that transform NGOs, particular small NGOs in this study. And thirdly, configurations that govern evaluation processes, generating power/data/knowledge relations and inequalities, and silencing dialogue and negotiations, in favour of technical solutions (e.g. Ferguson, 1994; Kothari, 2005a; Li, 2007; Lie, 2015a; 2015b; Hayes et al, 2017).

2.3.1 Critical configurations of evaluation

There are diverse methods and prescriptions for understanding impact evaluation as scientific, or participatory in nature, but not for seeing evaluation as organisational practices, unexpected and messy contingencies, or political conflicts. Professionals witness these aspects, but are under pressure to document results, successes, and rigour, not accounts of the back and forth of evaluation processes, its complex, political or messy practices (Law, 2004). This section outlines example observations from the literature concerning such configurations, which can

appear illegible, unclear or un-prescribed within TIEK. They are however significant in how impact evaluation is configured in practice and how it is changing today.

If evaluation is viewed as part of historical practices and political interactions, in contrast to being apolitical and outside such interactions (Picciotto, 2015: 152-153; Wallace et al, 2006; Eyben et a, 2015; Norris, 2015: 135) then TIEK results, methods, truths and participatory claims can be understood not as prescriptions, but as valuable elements of ongoing evaluation practices. Scientific and participatory discourses and methods become not opposed to views of evaluation practices, but aspects of them. This section describes evaluation configurations, firstly concerning evaluation beyond individuals or methods alone; secondly as expert problems instead of diverse voices; thirdly, as historically and politically situated; and finally, as having important aspects of reputational provenance within global aid networks.

First of all, evaluation independence and governance respond to demands exerted upon individual evaluators and asked of evaluation methods and results, but evaluation in practice is more complex than this frame. Scientific rigour or participatory commitment targeting interventions sites alone, cannot explain this wider frame and its distributed relations. Distortions occur not just because of methodological issues such as poor variable design, or poor statistics, but also because of personal evaluator prejudice, ambition, conviction, misinterpretation, distortion of evidence, or misleading results (Ioannidis, 2005; Picciotto, 2012: 221). Beyond individual ethics or target groups, evaluation processes shape their own stakeholders. Evaluation independence is often understood as a question of individual evaluator integrity, ethics and skills, rather than a wider set of overt and covert pressures, responses, and practices (Picciotto, 2013: 18-19). Framing evaluations via individual skills, ethics and behaviours constricts wider analysis. Consultants for instance, face real pressures to deliver timely and clear results in exchange for fees, and this configures evaluation practices and documented results:

> *"Most evaluators operate as consultants. They give primacy to the utilization of evaluation results and limit their ambition to the timely provision of evaluative knowledge to decision-makers" (Picciotto, 2013: 19)*

Such constraints limit more extensive evaluation depth or engagement. For example, Bamberger et al (2010: 2) argued that mixed-methods can reduce the over-reliance on experimental RCT methods and bridge the gap in understanding between distant managers and

field workers/participants. However, such gaps are not limited to RCTs alone, but to potentially any evaluation. Quantitative, qualitative and mixed-method evaluations are all susceptible to constricted frames, such as consultant timescales, fees and pressures for defensible results.

Criticism of digital technologies (O'Neill, 2014; Beer, 2018), statistical methods (Muller, 2018), and participatory designs (Cooke & Kothari, 2001; Nichols, 1999; Mama, 2000) support this idea that evaluations in practice have many often unequal consequences for different stakeholders. This means, that in trying to understand what evaluation does in an "*unstable, noisy and risky operating environment*" (Picciotto, 2012: 222) one must investigate evaluation data/knowledge practices across the aid chain, in a space wider than the target intervention sites alone, wider than a single method, and wider still than an individual evaluator's ethics or skill. A question this raises for the study is how small NGOs navigate such wider evaluation configurations in practice?

Secondly, TIEK prescriptions tend to work through expert problem frames which constrain the inclusion of alternative voices. Deniston describes how expert problem framing determines evaluation needs, methods and envisaged results.

> *"... there is little reason to have a program unless there is a problem. Accordingly, let us define a problem relative to a given program as some predicted condition that will be unsatisfactory without the intervention of the program and satisfactory, or at least more acceptable, given the program's intervention (Deniston 1972; 1974; Mohr, 1995: 10)*

Furthermore, Mohr (1995: 13) gives an example of how a statistical method choice is a response to a problem definition, in describing a police wiretapping program to catch drug pushers. Mohr's problem definition features a flowchart showing how wiretapping can lead to arrests, reduced drug consumption and even increased happiness, which all feature as program objectives for the evaluation to investigate. Logframes and theories of change, both widely adopted development tools, use a similar linear and causal information model to define impacts.

However, in contrast Edwards (1979) describes how evaluation problems, objectives and accompanying indicators are assigned by influential agents, in historically influential roles, using particular expert knowledges that enable data to be collected, organized, analysed and validated. Expertise, or influence in Edwards account, is applied both before and after the

choice of a desired method or arrival at results. In this way experts seek to shape the history and politics of actual evaluation processes, problems and perceptions of solutions.

However, even though experts normally define problems and impacts, they don't need to. Anderson et al (2012) report on interviews with 6,000 aid recipients, and two excerpts illustrate alternatives to the loudness of expertise. Firstly, a monk on the Thai-Burmese border called for "*foreigners to quiet their voices*" in determining what development impacts could or should be (ibid: 7). Secondly a Solomon Isles project officer yearned to hear people's feelings, not monitoring statistics (ibid: 8). Norrish & Sayce argue that evaluations do not just gaze upon "*natural systems with loads of feedback loops*", but social wholes and power relations (2004: 166). These examples suggest that TIEK is applied in contexts, involving histories, roles, and diverse voices belonging to people who remain at evaluation sites long after project staff, evaluators, and TIEK methods have left. It is they who witness long-term impacts. But how do NGO evaluations in practice balance such expertise and spaces for diverse voices?

Thirdly, observers have noted how evaluations have hidden politics, and how influential methods historically reshape future evaluation practices. These observations are rarely acknowledged in TIEK. Development projects and evaluations can result in reduced inequalities, but can also contribute to political manoeuvring or new inequalities. Two examples illustrate this point below.

Firstly, actual evaluations depend on contingent political negotiations. Edwards (1979) shows this through a description of an evaluation design for the California Board of Education racial desegregation plans, following the 1977 court ruling on the Mary Anne Crawford case. On the surface, the design was technically sound, a multi-attribute measurement utility with 150+ indicators. However, Edwards vividly narrates the behind the scenes political negotiations. Construction of the indicators required many hours of personal coaching between the evaluator, Edwards, and individual board members (judges, politicians and professors), in order to establish indicators that were acceptable and weighted for each expert's agreement. Indictors pertained to many issues, such as children's bus routes, teacher quality, and the importance of racial desegregation versus educational attainment. The negotiations were arduous because board members insisted upon their own agendas and refused to go on record regarding biases, such as expressing preferences for educational quality over racial-ethnic composition, or vice-versa. After initial design and negotiations, the indicator tree and every twig on the tree was reviewed by 42 further teams in school sub-districts. The final variable tree became "*too*

complex" (Edwards, 1979: 19), and the evaluation plan was eventually shelved, leaving Edwards to conclude that open political dialogue would have been more valuable than a complex technical evaluation. This case is insightful because it is rare to see descriptions of backstage politics in TIEK.

A second case illustrates how an evaluation method can unexpectedly evolve over time into universal standards that threaten diversity, voice and participation. Porter (1996) describes how Cost Benefit Analysis (CBA) was initially conceived of as a technical tool used in disputes over land management and budget allocations between different US government agencies in the nineteenth century. Over time CBA became standardized and used across the US public sector. In the 1960s economists adapted CBA, adding economic theory to underpin it and transform it into a method of auditing and objective evaluation founded on economic principles. Hammond (1960; Hammond et al, 1976) considered this entry of economics into CBA as a license to "*concoct imaginary data*" (see Porter, 1996: 187). Nevertheless, the new CBA evolved further, entering areas such as the military, government programs, the pricing of intangible assets, public health and education. Porter concluded that CBA began as a strategy for limiting the play of politics in public investment decisions, but became a universal calculating standard that protagonists believed was a mechanical way of objectively, empirically rendering accountability and expunging bias or manipulation. This resonates with normative views of CBA and auditing as contextless, ahistorical and apolitical, subject to technical expertise alone. In practice, CBA did not start out as such a system and its political journey to global discursive power, as traced by Porter, has significantly shaped many evaluation methods used today, smoothly submerging data/knowledge politics.

These two rare examples, of backstage politics and the genealogy of a method, reveal another question for the study rarely attended to in TIEK. Do NGO evaluations acknowledge hidden politics or how methods shape evaluation processes?

A fourth concern in how evaluations are configured is the reputational origins or "provenance" of methods and evaluators. Certain organisations influence the way that we evaluate, but see no problem with their own reputational influence. This applies to TIEK methods and prescriptions, reports, and evaluators themselves, coming from large well-known institutions, large INGOs, or small unknown local NGOs.

Particular methods and reports are frequently initiated, funded, matured and disseminated by already influential development agencies and experts. For example, this is true of the Critical Stories of Change approach (Cornwall, 2005; ActionAid, 2012) or the Do No Harm approach, born from CDA collaborative projects, funded by DFID, USAID and the Norwegian Ministry of Foreign Affairs (Wallace, 2015). Rarely are new technical innovations produced, disseminated by, or credited to small NGOs or beneficiaries themselves.

RCTs provide a good example case. They are understood as offering certain advantages or disadvantages within methodological debates. However, they can also be viewed as historically formed conceptual packages, which produce effects upon organisations and industries. Methodological prescriptions and expert commentators are key influences here, visible when evaluation pioneers such as Cronbach, Stake and Campbell promoted mathematical, positivist approaches and randomised tests as gold standards in the 1960s and 1970s (Picciotto, 2012: 219; Campbell & Stanley, 1963). RCTs, a method imported from medical trials, became the gold standard for development impact evaluation in the early 2000s (Picciotto, 2012: 215). A group of young Massachusetts Institute of Technology (MIT) researchers postulated that "*the effectiveness of development assistance was best ascertained through experimental methods*" (Picciotto, 2012: 18), and Esther Duflo, the charismatic founder of the MIT Poverty Action Lab, claimed at the 2003 World Bank Conference on development effectiveness that:

> *"Just as randomized evaluations revolutionized medicine in the 20th century, they have the potential to revolutionize social policy during the 21st" (Duflo & Kremer, 2005).*

The Bill & Melinda Gates Foundation and the William & Flora Hewlett Foundation, intent on entering the development sector, enthusiastically supported RCTs. In 2004, the two philanthropies convened the Evaluation Gap Working Group, hosted at the Center for Global Development (CDG, a US private think tank originally based in Washington D.C), and by 2006 published the "When Will We Ever Learn" report supporting RCTs as the definitive new standard in impact evaluation (Center for Global Development, 2006; Picciotto, 2012). Despite weaknesses with experimental RCTs, with Pattern (2014; 2015a) suggesting they were not a gold standard but a "*fool's gold*", and others conceding RCTs inappropriateness for evaluating foreign aid impact (Mohr, 1995: 71) or evaluating complexity (Picciotto, 2012), their appeal is still substantial in planning, fund management and administrative decision-making. Garbarino & Holland identify this appeal in measuring over large areas.

"... a country or region ... to achieve breadth in coverage and analysis. Typically, the random sample survey produces quantifiable data that can be statistically analysed with the main aim of measuring, aggregating, modelling and predicting behaviour and relations." (Garbarino & Holland, 2009: 7)

The RCT story illustrates how reputation or provenance is important. RCTs were migrated from medicine, garnered support at MIT, endorsed by CDG, a reputable think tank, convened and funded by two household name philanthropies entering the development sector, labelled a gold standard, and dispersed into development globally. Hamersley (2005: 95) even suggested there was an evidence-based "*movement*", seeking to build a global position and influence. Methods have provenance; they are not apolitical or ahistorical.

Brigham & Hayes (2013: 121) more recently showed how the Value Management Method (VMM) used in e-development and impact assessment, had origins at Harvard University, and grew through the university's "*network of consultancy firms, research centers, publishers and leading US universities who are global in reach*" (see also Thrift, 1997). This enabled VMM to accrue a high degree of legitimacy, authority and expertise, becoming a "*soft law*" of governance that could not be ignored, in Nye's terms (2005).

A small NGO may find it risky to decline a donor request to assess impact using a reputable UN, Harvard or MIT evaluation system. The NGO, and their communities, must accept the provenance, assumptions, and potential frames of such imported evaluation packages. Thus, the provenance question for the study concerns if and how local or globally reputable impact evaluation methods are preferred in NGO evaluation practices?

2.3.2   Critical configurations of NGOs

In the current climate of demands, TIEK prescriptions, and data/knowledge intensity the literature suggests NGOs are changing. Arguably this has already happened for large INGOs working in multiple countries, with numerous departments, employing professionals from different subfields with different specialisms Wallace et al, 2006). However, smaller national NGOs face similar pressures. Banks & Hulme (2012) acknowledge that NGOs have changed in relation to increased competition for institutional funding. Furthermore, such NGO transitions are normal in the sector (Lewis & Kanji, 2009: 13-15).

Wallace et al (2006: 112-113) noted issues regarding evaluation that affect NGOs, such as increased needs for accountability, costs of monitoring, identifying causes, participatory tensions with donor goals, and threats to community relations. Norris (2015: 139) warned about failing to see the impact of evaluation upon well-being, organisational performance or the systemic outcomes of evaluation processes. In such ways, the impacts of evaluation upon NGOs go beyond efficiency, effectiveness and robust results. A significant yet mundane transformation that must be addressed by the literature is towards data/knowledge intensity, through activities of evaluating, monitoring, data collecting, analysing, planning, reporting, marketing, and communications. Below, the literature on four of these areas is reviewed: the technical restructuring of NGOs as outsourced service providers; the communications and marketing of impact; confusion as a commodifiable capacity deficit; and how field-work is displaced by conceptual data/knowledge work.

A first observation from the literature concerns the restructuring of NGOs into contemporary outsourced service providers. Technical impact evaluation functions within broader managerial drives towards audit and measurement, which increased in the 1990s, and following 9/11 in the US (Gardner & Lewis, 2015: 35-36). Managerial strategies such as short-term project life cycles, a results focus, use of ICTs for reporting, the screening of goals, and documenting impacts, all feature in making NGOs accountable to donors and management. Logframes, regular evaluations, and rigid indicators remain widely used and required by donors (Wallace et al, 2006: 91-93) in compliance frameworks. They are used to break up short-term projects into discrete goals, inputs, activities and outputs, often over a 1-3 year timeframe. Such tools shape outcomes and evaluations, involving considerable time and effort from contracted NGOs.

Short-term projects are also difficult to evaluate in terms of sustainable impact. Baseline data is not always captured and is normally of rudimentary quality (Norrish & Sayce, 2004; Roche, 1999), and it is unrealistic to imagine that contract short-termism can "*fix*" historically formed socio-economic problems (Norrish & Sayce, 2004). This is especially difficult when there are "*huge amounts of interest in the planning stage*" of a project, but staff "*start leaving the ship like rats at the end of a project*" (2004: 3). If impact sustainability is uncertain in short time frames, then short-term service contracts and compliance frameworks using logframes, indicators, and technical evaluations risk enacting control relations rather than ensuring long-term impacts.

Moves towards managerial control through project goals, indicators, logframes and associated evaluations, threatens the ability of NGOs to innovate, experiment, negotiate with grass roots organisations, and engage closely with communities, all attributes of NGOs lauded in the 1970s-1980s (Drabek, 1987; Bebbington et al, 2008; Banks & Hulme, 2012: 21). Such shifts resonate with the earlier moves from "*small-d development*" to "*big-D development*" where the former was politically engaged, innovative, and community led, and the latter foregrounded rational plans, managerial controls, target setting, technical measuring and expert interventions (Banks & Hulme, 2012: 13; see also Power et al 2002; Bebbington et al 2008).

> *"The replacement of broader goals of empowerment by measurable outputs overlooks or ignores the systems, processes and institutions that perpetuate poverty. This depoliticises strategies open to NGOs for promoting 'little-d' development, leading them away from relationships with social movements and towards narrow and specific targeted programmes for 'big-D' development (Banks & Hulme, 2012: 13).*

Critical perspectives on managerial outsourcing positions TIEK, with its emphasis on expert methods and decision-maker results, as a component of NGO service frameworks, rather than as tools for engagement and political negotiation. In managerial models of development, participation is not political action, since it makes no attempt to change the "*systems, processes and institutions that perpetuate poverty*", implying that local NGOs "*flutter*" around the bottom rungs of the participation ladder, where communities have little control over development (Banks & Hulme, 2012: 13).

Advocates of participatory development argue that external determination of local goals and objectives in service models erodes grassroots participation and empowerment, and escalate a professionalism encouraged by donors in a rush to achieve smart, tangible and quantifiable measures of impact (see Hailey, 2000; Power et al, 2002; Elbers & Arts, 2011).

It is important to note that development 2.0 and digital information systems are also constitutive in such compliance architectures (Brigham & Hayes, 2013). Thus, another consequence of the need to capture and prove impacts is an increasing reliance on ICTs for evaluation data, information and knowledge management. Walsham (2001) discusses how technology in global development leads to local adaptations, improvisations and questions of surveillance and control regarding for example how data is stored, analysed or transmitted. Thompson (2002)

questions the methods and tools used to capture such data, arguing for locally sensitive data regimes in rural healthcare in South Africa. This suggests the demands for using ICTs in impact evaluation and managerial control can produce insensitive processes and inequalities, where NGOs align to donor technologies for capturing, storing, transmitting impacts.

The problem with managerial restructuring, impact measures, distant planning, and use of digital systems is that the transformation of NGOs may signal how technical efforts are in place, but not the desired socio-economic changes, towards equality, poverty alleviation, and empowerment (Gardner & Lewis, 2015: 36). The overriding, perhaps unintended effect, is that NGOs and particularly smaller ones around the globe, become outsourced service providers for donors and larger NGOs who manage sector finances, data and knowledge. Many authors have critiqued this NGO contracting model (Banks & Hulme, 2012: 13; Edwards & Hulme, 1996; Bebbington, 1997; Elbers & Arts, 2011), but the shift to data/knowledge intensity, to needing, capturing, to knowing and transmitting impact, is a significant and growing concern. Thus, a question that emerges for the study concerns if and how distal managerial plans and compliance targets shape local needs and contexts for NGO evaluations.

The literature suggests that a further critical configuration that emerges from the need for NGOs to demonstrate results, impacts and effectiveness is an increasing focus upon communications, marketing, public relations and branding in order to support fundraising, public reputation, and niche expertise promotion. This trajectory was noted in the late 1990s when Roche observed that fundraising, assessing impact, maintaining a high profile, and gaining press coverage were all elements in a "*vicious circle*" (Roche, 1999: 2).

Public profiles and branding contribute to trust and reputation. In the 1990s, many NGOs began to expand in size (Banks & Hulme, 2012: 4) with communications becoming a key part of their work. However, smaller NGOs around the world may not have the resources or capabilities to expand, hire PR people or compete in the vicious circle. DFID UK prioritized funding to just five large NGOs between 1997 and 2004 (Wallace et al, 2006) and the vast majority of global humanitarian aid went to just 28 large agencies and NGOs in 2013 (Els & Cartensen, 2015: 1-2). Estimates of around 35,000 large NGOs in the world, and over a million smaller formal and informal ones (Lewis & Kanji, 2009), suggest that communications, marketing and branding have become crucial for survival. In fact, it may be difficult for small organisations to compete, establish their niche, market their brand, and demonstrate their impact and effectiveness without communications expertise and resources. In a competitive climate, NGOs are under

pressure to exaggerate successes for donor and public consumption, which influences funding flows, market reputations, and aid policies (Eyben et al, 2015).

NGO communications are important during a time when new actors are entering development, such as international consulting firms, philanthropies, celebrities, private businesses, diasporas, and advocacy groups (Richey & Ponte, 2014; Ponte & Richey, 2014). Branding and marketing means focusing on a public who consume development messages through campaigns and marketing produced by and with these actors, such as through a celebrity's popularity, so that the primary function of the communications is to "*create a community that 'care' for development*" (ibid: 9). Demonstrating successful impact on poverty, hunger, disease and discrimination are key elements in such communications. Richey & Ponte characterize the place of impact in communications through a celebrity-focused trope:

> *"The impact is also based in images – pictures of the world. Northern shoppers are animated by a confident aid celebrity Bono, who speaks from American television screens claiming: 'We can be the generation that ends extreme poverty. This is our moment to show what we're about.' " (Richey & Ponte, 2014: 15)*

Richey and Ponte draw attention to how the instrumental use of images and representations in aid marketing and campaigns influences development policies, discourses and practices (ibid: 9). Such branding of impacts has now become a component in the ethical consumption of aid, and the strategic "*cause marketing*" used by NGOs in the sector (see Kothari, 2014: 57).

Two configurations follow from this literature on increased communications and marketing efforts. Firstly, numbers in the form of statistics, as well as images, narratives, campaigns and other marketing techniques have become tactics in the successful communication of development problems and impacts. The production of imagery, campaigns, impact narratives and the securing of celebrity endorsements are much harder for resource limited small NGOs who lack dedicated marketing or evaluation teams. Secondly, marketing emphasizes the broad circulation of impact messages rather than the authenticity or usefulness of such messages at specific project or community sites. Impacts conveyed for publics through campaigns and celebrities, or in reports for distant funders, policy makers and administrators make use of the awe of numbers and the attractiveness of imagery at and for distant consumption. These two resulting configurations are both problematic for understanding impacts, and point to an increasing need for critical perspectives in understanding such communications (Richey &

Ponte, 2014). A critical concern for the study that arises from this literature is if and how small NGOs use evaluation results to communicate and market impacts.

Next, the literature suggests that the demands for impact data/knowledge, alongside TIEK prescriptions, produce a capacity gap wherein smaller NGOs remain in confusion and a capacity deficit. NGOs appear to be affected by the confusion that accompanies evaluations, the hopes and fears about daily work, evaluating the right way, or doing a good job. O'Sullivan (2004) acknowledges staff member hopes and fears, political concerns and organisational readiness (or lack of it) when the time comes for evaluations. Evaluation staff must "*weather*" the evaluation storm by going "*outside*", and going "*through the dense fog of evaluation*", before they can see the "*haze lift*" and finally open their eyes to the "*light of day*", arriving at an evaluation place of reflection and understanding (ibid: 76-78). There are presumptions in this language about what evaluations should be like, but it is refreshing to see a commentator acknowledge the internal staff experiences, stress and emotions, issues marginalized in TIEK. For O'Sullivan however, organisational transformation is good, which fits capacity building discourses where NGOs still need to master TIEK skills.

Observers argue that mastering and learning is confusing. Norrish & Sayce (2006: 4-5) acknowledge the "*rich mix*" of impact evaluation, assessment, the "*plethora of issues*", the "*need for consensus*", and a "*literature that abounds with definitions*". The Technical Centre for Agricultural and Rural Cooperation (CTA) in Canada supported Norrish and Sayce's research because of the confusion, disagreements and furrowed brows in their projects.

> "... whenever the subject of evaluation was raised, hot on its heels came the subject
> of impact assessment, and brows became furrowed and opinions heated as
> everyone expressed different views on the subject" (ibid: 5).

Brows are not only furrowed in projects spaces. This is evidenced when researchers declare at the start of a UK government commissioned report on impact evaluation methods, theories and challenges, how difficult it is for themselves to write, and how difficult it is for readers to read about impact (Stern et al, 2012: 5, in "Reading hints" section).

As such, if the literature confirms that experts are confused, how can small NGOs who lack expertise respond to the organisational capacities required to know impact, navigate evaluations, meet managerial goals, master technical specifications, manage digital data, deploy technologies, guarantee participation, and produce scientific rigour? An imbalance is

present in the shift to data/knowledge that underpins NGOs' needs for impact learning. Larger organisations produce, disseminate and legitimate impact methods and knowledge, and have dedicated staff and units to produce media, data, reports, theory, methods, techniques or technologies. For example, The World Bank produces economic development knowledge to which even expert anthropologists must acclimatize in order to sustain their careers (Mosse, 2004a). Universities with Development Studies or Area Studies departments (see Sumner, 2006) such as the Institute for Development Studies (IDS) based at the University of Sussex, hire expert academics to create and publish development and evaluation related knowledge. The UN system publishes a wide array of evaluation "how to" manuals and briefing notes from its diverse tree of sub-agencies. Harvard University uses its network of soft power, research centres and consultancies to promote methods and secure contracts (Thrift, 1997; Brigham & Hayes, 2014). Specialist evaluation organisations, such as 3ie or the OECD-DAC, frame, regulate and bound TIEK discourse. But the capacity question is left open – how do small NGOs respond to publications and prescriptions, the shifting evaluation norms, fads and fashions Leal (2008) identified?

How and where TIEK is produced also affects the capacity gap. Kothari's (2005a: 94) interviews with ex-colonial administrators threw light on the post-colonial development sector, and the academic rise of Development Studies in the 1960s to 1970s, as jobs and expertise shifted towards professionally trained graduates. One ex-administrator remarked:

> *"By 1983, the young, very intelligent, tremendously articulate had taken over but with no overseas experience" (ibid: 92)*

In another paper on how the sector has become professionalized, Kothari (2005b: 440) claims aid orthodoxy is more concerned with "*generating excessive professional knowledge and skill, and 'experts' ... than with alleviating poverty ...*". Schuurman (2009: 834) further argues that academic development institutions operate "*according to a market logic*" having to output students, ideas and concepts, flooding aid discourse with experts, approaches, methods and knowledge. Impact evaluation is just one of many strands in professionalizing drives, but the effects of accelerating TIEK is a real issue, with for example jobs in ministries and NGOs requiring graduates who know "*how to prepare, manage and evaluate development projects, who know how to measure efficiency and increase the impact of projects*" (ibid, 838). These jobs are filled by technically trained graduates, not the poor, grass roots activists or local NGO staff, partly because of the evaluation capacity gap. Small NGOs and NGO staff may not have

access to expert TIEK. Sumner (2006: 647) sees universities as important in terms of being "*training grounds*" that shape future policy and practice. This means that where TIEK is produced and how it circulates is an important dynamic in understanding the transformation of small NGOs towards data/knowledge intensive evaluation.

Despite the capacity gap, confusion, related "*rituals*" and "*headless chicken stuff going on in most development agencies*" (Norrish & Sayce, 2006: 166), NGOs seeking funds, legitimacy and a voice at the table in defining policies or making program decisions need to be able to speak in the voice of the latest development fashion. Otherwise they risk financial and political marginalization (Wallace et al, 2006). Therefore, a further configuration involving NGOs in development 2.0 is that impact capacity itself becomes uneven and commoditized. A question to revisit in the discussion chapter concerns if and how small NGOs experience this impact commodity confusion and capacity gap.

The final critical influence upon NGOs evident in the literature concerns how office work and data/knowledge work displace fieldwork and community engagement. Wallace et al (2006: 165) argue that NGO staff are increasingly required to spend more time and effort on paper work, administration, and bureaucracy. In their research, staff felt more like administrators, not development workers. Staff on the frontline, who had extensive field experience, argued that management tools did not work when field implementations began. Thus, there was a disjuncture between paper objectives and indicators, and the day-to-day realities at project sites with target groups. Staff are left grappling with context and culture. Increasingly computers, screens and digital data mediate the grappling between fields and offices (Heeks, 2018; Beer, 2018). A number of publications highlight participant and field worker voices (Anderson et al, 2012; Norrish & Sayce, 2004; Wallace et al, 2006), however critical observers suggest frontline voices are muted:

> "*The systems and reward structures for NGO staff usually prevent the voices of those on the frontline reaching the distant funders.*" (Wallace et al, 2006: 166-167)

This distancing between stakeholders is further aggravated by the influence of audit cultures (Harper, 2000; Strathern, 2000), and systems of measurement. Strathern argues audit culture has "*broken loose from its moorings*" in finance and accounting, and that audit processes are "*shut out from the 'real' tasks of productive work*", being instead applied to new spaces such as "*evaluating, measuring and reckonings*" (ibid: 1-2). Auditing knowledge and techniques

enter evaluations and produce new activities which displace old ones, in terms of not just fieldworkers, but time, energy, resources, expertise, labour, evaluation agendas, goals and priorities. As so much work is expended by an organisation in order to produce the evidence and environment that audit processes demand, NGO staff are left with less time to tend to field work, engagement, trust building, experimentation, innovation, or empathic understanding. Audit processes involve "*locking up time, personnel and resources*" (ibid), and reshaping evaluation practices, which displace other activities.

Donors often commission short-term evaluators armed with blueprint evaluation methods, which are seen by local NGOs and project participants as exercises in compliance, and accountability, of utility only to the donors (Norrish & Sayce, 2004). NGO field workers in Norrish and Sayce's work emphasised unique local contexts and specific local issues, rather than best practice models, blueprints, replicability or scalability.

This suggests auditing skews towards the needs of donors, management and evaluators in accounting, scaling, and comparing projects, NGOs and budgets, away from fieldworkers, beneficiaries, local community networks or beneficiary lifeworlds. Critical literature on impact assessment for non-profit and public sector organisations raises similar concerns related to auditing and managerial control, for example through quantification (Pinch & Bijker, 1987: 30), "*a need for quantifiable impact*" (Duncombe, 2009: 18), surveys and questionnaires (Townley, 1995), and evaluation dominance by donors, consultants and NGOs who manage evaluations (Kirkpatrick & Hulme, 2001; Duncombe, 2009).

Research has further identified how audit processes are attractive to planners and administrators, even when they create a lack of trust or undermine treasured goals (Wallace et al, 2006: 37). A report by UK think-tank Demos (2003) on local government work with communities acknowledged that: "*strict performance criteria, emphasis on quantitative outputs and formal participatory structures*" can hollow out service provider and user relations.

> "*A heavy audit culture often breeds an atmosphere of distrust and risk aversion, which encourages uniformity in programme design and inhibits the distinctive contribution that CBOs make*" (Demos, 2005: 7).

Workers in smaller organisations dealing with large-scale command and control structures can become demoralized and cynical, in a state of "*grappling with inappropriate and irrelevant*

centrally set targets in contexts unrelated to those on paper" (Seddon, 2005). Critical literature suggests a need to look at the influence of auditing processes on evaluation and how these create distance or unequal power relations, displacing NGO field-work and engagement.

In summary, the critical perspectives illustrate how the upward accountability in NGO service contracts (Wallace et al, 2006; Lewis & Madon, 2004: 120; Brigham & Hayes, 2013: 114) leads to accelerated data/knowledge compliance frameworks, managerial controls and sub-specialisms, such as around evaluations, impacts, marketing and communications. Elbers & Arts (2011: 725) argue that NGOs actively respond to donor pressures by avoiding compliance exposure, influencing and negotiating with donors, shielding, and portraying compliance through information work. However, despite such tactics and relatively invisible (to donors) weapons of the weak (Scott, 1985), uneven power/data/knowledge relations in impact construction and the wider development 2.0 landscape, appear to constrain the kind of TIEK data/knowledge constructed and shape the way small NGOs must re-organize or even transform themselves to fit in with it. The question this leaves for the study concerns how small NGOs navigate the evaluation office work versus engaged field work dilemma.

2.3.3 Critical configurations of governing and anti-politics

As impact is data/knowledge intensive, its governing relates directly to knowledge work and how expertise is mobilized by evaluators or at NGOs. It also relates to how preferences for technical evaluation frame impacts and risk silencing political dialogue. Impact governing draws on Foucault's concept of "governmentality" (1991), and what has come to be labelled "developmentality" (Ilcan & Phillips, 2010; Lie, 2015a; 2015b; McGregor et al, 2013; Norgaard & Deb, 2009). In governing processes, both scientific knowledge and local participation are relevant. For example, Procacci (1991) shows how poverty is governed when scientific expertise is deployed to define problems, populations and subjects, and to justify interventions to fix the defined problems. In this manner, TIEK is part of evaluation governing practices. Lie's (2015) account of developmentality adds a newer aid architecture of partnerships, participation and local ownership of development by recipient communities, NGOs and governments onto the scientific governmentality view, making power relations smoother, more "*indirect, subtle and tacit*" (Lie, 2011: 1).

For the study at hand, three governing configurations arise from the literature. Firstly, how do governing processes shape NGOs? Secondly, how is governing expertise mobilized in aid chains? And thirdly, how does anti-politics, the replacement of dialogue by expert authority, feature in impact governing practices. These configurations are described below.

Firstly, how do governing processes shape NGOs? The notion of governmentality offers significant insights into how impact data/knowledge intensity has consequences for NGOs, how it flows through them and changes them. In governmentality, power is viewed as distributed and productive (Foucault, 1977, 1980, 1991), not concentrated in powerful organisations alone or always destructive. Diverse spheres of life have come to be affected by governing processes, and these often take the shape of competitive, market rationalities (Hayes et al, 2017). Such processes transform organisations like NGOs according to Dean (2010: 21), and "… *reform institutional and individual conduct so they both come to embody the values and orientations of the market*".

NGOs performing impact evaluations are under pressure to align with funders, TIEK and managerial control, as part of how impact governing spreads around the sector. Dean describes '*governing through freedom*' (2010: 262), meaning that individuals freely follow knowledge claims, rather than being forced or coerced. He sees governing as presupposing the freedom of those governed (ibid: 24), their capacity for free thinking and choosing, and their "*mentalities of government*" (Miller & Rose, 1990; Dean, 2010: 24). Mentalities are broader than individual experiences and refer to collective knowledge, expertise, and ways of thinking or working, in *regimes of practice* (Foucault, 1991b; Dean, 2010: 27), such as in how economic (or development) knowledge implicates circulating models, theories, and technical vocabularies (Dean, 2010: 25).

In this critical configuration, NGO staff govern impacts by adopting TIEK, in order to align with markets, compete with other NGOs, claim impact successes and gain funding and legitimacy from donors or partners. NGOs also require sophisticated impact knowledge to inoculate themselves against outside expert manipulation, meaning it is in the interests of NGO and evaluation staff to promote TIEK for their NGOs (Roberts, Jones III, & Fröhling, 2005). As a result, it is understandable for NGOs to choose freely to govern impacts using globally prescribed TIEK and methods that privilege managers and donors, because these choices benefit the NGO. The NGO produces legitimate impact data/knowledge, gains funding

opportunities, and gains status as a reputable organisation. In return, the NGO is regulated by choosing to conform and comply.

As such, impact evaluation has become one technical discourse, which functions in development as:

> *"... on the one hand ... a whole series of specific governmental apparatus, and on the other, in the development of a whole complex of savoirs [knowledges]" (Foucault, 1991: 102-103).*

In this view, certain kinds of facts and truths become dominant within TIEK, and although we should be skeptical about viewing truths, facts or information as value free or apolitical (Knights, 2004: Willcocks, 2004), they do become institutionalized through scientific processes and methods (Introna & Whittaker, 2004). In a governmental view, empathic, intuitive and political knowledge does not carry the same objective weight as scientific or technical accounts of impact.

Changes in working culture away from co-working and trust and towards audit, surveillance and judgements also result from the complex knowledges and apparatus of governing (Townley, 2003; Power, 1996; Strathern, 2000; Porter, 1996). In Lie's related developmentality, donor organisations step back from top down direct control of development projects and impacts, and take on a monitoring and evaluating role (Lie, 2011: 9-10), through auditing, best practices, the normalization of conduct, turning participation into technical or managerial processes (Guijt & Shah, 1998: 3), and regulating targets such as through the MDGs (Ilcan & Phillips, 2010), or the Washington Consensus in the 1990s which opened up world markets to globalization (Lemke, 2002). According to Lemke, NGOs became part of these distributed, informal governing processes:

> *"Foucault's discussion of neoliberal governmentality shows that the so-called retreat of the state is in fact a prolongation of government: neoliberalism is not the end but a transformation of politics that restructures the power relations in society. What we observe today is not a diminishment or reduction of state sovereignty and planning capacities but a displacement from formal to informal techniques of government and the appearance of new actors on the scene of government (e.g., nongovernmental organizations) that indicate fundamental transformations in*

statehood and a new relation between state and civil society actors." (Lemke, 2002: 58)

This literature is important in articulating the diffusion of agencies in impact governing practices. It clarifies how "*knowledge strategies*" (Ilcan & Phillips, 2008; Phillips & Ilcan, 2004), data strategies (Beer, 2018), regulating the visibility of target impacts, and the ways we can legitimately know impact, configure our "*truth claims, and other diverse expert-oriented activities*" (Dean, 2009: 33).

The concept of governmentality and more recently developmentality are important for understanding the previous critical configurations, demands, histories and prescriptive functions of TIEK. They explain why impacts knowledge is circulated in professional networks and edited or filtered using TIEK to secure legitimacy. They also explain how NGOs are pressured to construct impact data/knowledge, to improve their evaluation capacity, and to provide attractive services in competitive markets. Governing and survival are intimately linked.

The question this raises for the study concerns how small NGOs respond to such governing pressures. Do they deliver scientific impacts to aid markets, or construct alternative frames for impacts, political dialogue, and relationship building with other stakeholders?

Secondly in terms of impact governing, TIEK in methods, guidebooks, videos and websites is mobile, moving in documents and knowledge products across aid chains, sites and practices. Harper (2000: 15; 1998) watched circulations of aid documents and information pass over meeting tables, through hunched up bodies, as data entered machines at the IMF in Washington (ibid, 1). Similarly, Mosse (2004a) argues that development knowledge products move and are exchanged within markets. In order to survive and have influence in The World Bank, anthropologists produced and sold "*ideas, packages of concepts, models, methods, or tools*" (ibid, 81). He viewed these products as part of "*a knowledge system, which perpetuates separation of the world of policy rationality from the contingencies of practice*" (ibid, 85), in which the anthropologists produce knowledge as a mobile currency which had value in local World Bank environments.

Mobile knowledge lives on when it travels in wider aid chains and professional networks (Ilcan & Phillips, 2008), beyond the producer departments and institutions, entering broader development markets and bureaucracies, contributing to the expert complex of knowledges

Foucault identified. Such adopted knowledges come to define, normalize and frame legitimate truths and impact. From this wider perspective, Mosse (2004a) identified the local pressures on anthropologists to produce knowledge in order to defend their positions, but additionally this knowledge circulates and contributes to governing configurations across aid networks in the practices of donors, NGOs and evaluators for example. TIEK produced and intended to do work in one place and practice, travels to do other work in other practices and other places.

TIEK taken into fields and offices by travelling aid evaluation consultants and experts in the name of public or community interests is largely unexamined according to Sandercock (1998: 88). It privileges expert plans and pre-established categories such as class, age, gender and so forth. Experts and consultants often end up around a table together, planning a project, working with a "*vague mandate*" from local people and a clear set of objectives from funders (Wallace et al, 2006: 36). Li (2007) observed specialists and experts in hygiene and credit finance alongside World Bank officials competing to influence and change communities in Sulawesi, Indonesia, by instilling habits of choice and competition. For Li, this knowledge was deployed according to the deployers' needs, not the consumers' needs, nor for universal good.

This literature pays attention to the micro-practices of data/knowledge work, how it links to global flows such as TIEK, and pressures experienced by experts. Target communities along with southern partner NGOs are considered as sites for the application of managerial models, specialist knowledge and impact capacity building. Shivji (2006: 23) bemoaned the "*laptop consultants*" armed with rhetorics of participation, flying from country to country, re-packaging carbon copy development programs and plans. This characterization of experts and pre-fabricated knowledge products is seen to hinder participation and authentic partnership with southern NGOs (Wallace et al, 2006; Mowles et al, 2008: 809) and pervade the development landscape, embodying asymmetrical power/knowledge relations. That expert knowledges compete, travel, and form part of diverse aid practices is instrumental in how they normalize thinking and acting, and govern what impacts can be. These flows apply directly to NGO impact evaluation processes because evaluators, like auditors, become assigned and trained in the discourse as qualified to make claims to impact truths, in ways that others, in fields or offices, are not legitimated to do. Porter (1996: ix) suggests that a "*highly disciplined discourse helps to produce knowledge independent of the particular people who make it*". Thus, travelling evaluators are not lone judges, but part of a larger legitimated discourse of TIEK adopted by development agencies and institutions. Kothari (2005a: 93) further argued

one's status as an expert or consultant is linked to the status of the institution one works for, its legitimacy, financial resources, and global sphere of influence. Expert status in this view is graded differently if you work for the World Bank in Washington D.C. or for a small NGO working with hill tribe babies in Thailand. This status issue resonates with the provenance of impact knowledge discussed earlier, and how the knowledge is coupled with centres of knowledge production rather than centres of poverty or marginalisation.

Professionalized practice also produces power configurations that travel geographically beyond claims to efficiency or effectiveness alone. Kothari (2005b: 439) argues that development as an industry has erected a governing expertise, authority and managerial distance that guards against serious participatory challenges. Chambers (1993: 5-6) shows how development sector professionalism maintains itself through a *"repertoire of defence against accordance and threat"*, that assimilates methods used to *"modify, describe and often put some sort of number to the discordance, coding it so that it can be fitted on as an extension of the normal paradigm"*. In this sense, criticism in the literature for quantification in impact evaluation, deployed through surveys and questionnaires as a managerial technique to compare and control distant activities (Townley, 1995), must be accompanied by criticism of participatory methods and participatory forms of evaluation if they too become professionalized, driven by experts and made into mobile methods for distal governing. This is precisely what Lie (2015a) warns against in developmentality. TIEK models and laptop consultants contribute to governing configurations of NGOs globally, regulating the shape and profiles of their impact data, information and knowledge (Ilcan & Phillips, 2008).

The critical question arising from knowledge mobility literature for the current study, concerns how small NGOs respond to mobile TIEK. Are they governed by it? Do they react against it? Or something else?

A third set of governing configurations involves the "anti-politics" associated with Ferguson's research in Lesotho (1994), and considered here as the displacing of political tensions, contexts, conflicts, dialogue, negotiations and power dynamics by TIEK. In impact evaluation, anti-political effects concern the power and politics of evaluation, and contrast with the technical, market focus on rigorous methods plus robust results. This section describes six anti-politics configurations evident in the literature: how development goals expand and change; how failure sustains ongoing technical promises; how debates on social change, political transition, and political power are black-boxed; how anti-politics works at large and small scales; and,

how the complex structural contradictions of poverty rarely feature in TIEK. The last sub-section, reviews calls to reignite political voices in real or honest evaluation.

First of all, anti-politics concerns how goals and objectives change and grow. Escobar (2012) discusses two national programs for food, nutrition and poverty reduction in Columbia in the 1970s to 1980s, the PAN and DLI, which cost a combined total of $550 million (ibid: 133-137). The effects of the programs bore little resemblance to their poverty reduction goals. Instead of nutrition and poverty reduction, major impacts according to Escobar were threefold: the projects made rural farmers vulnerable to urban market fluctuations, set a precedent for national, regional and local government collaboration; and made the rural poor much more visible for future government planning (ibid: 139). In this sense, the two programs unleashed socio-economic and cultural processes that went well beyond their intended scope of agricultural production and poverty reduction leaving Escobar to wonder "*What are strategies like PAN and DLI really about?" (*ibid: 134*)*. Escobar concludes that they are about power and modernity (ibid: 142), referencing Ferguson's critique from Lesotho. Ferguson had questioned the manner in which development agencies defined problems and implemented solutions in Lesotho without taking into consideration the local historical, political or economic context, and argued that the agencies had erected an "*anti-politics machine*" in order to govern target populations.

> *It may be that what is most important about a "development" project is not so much what it fails to do but what it does do. … If the 'instrument effects' of a 'development' project end up forming any kind of strategically coherent or intelligible whole, this is it: the anti-politics machine." (Ferguson, 1994: 256)*

A second aspect of anti-politics literature is that project failure does not lead to review and reform, but to more of the same technical solutions that lack context and political engagement (Ferguson, 1994: 20). The failure of prisons to rehabilitate prisoners (Foucault, 1979/2000: 176-7) did not prohibit the unexpected effect of creating a class of marginalised, undesirable "*delinquents*" that comes to act as a form of societal control on the rest of the population, who do not wish to become delinquents or prisoners themselves. Thus, despite a failure to reform for a century and a half there remains a great reluctance to dispense with the institution of the prison. Likewise, Ferguson contends there is no intentional conspiracy behind the continual failure of aid projects, but that failures result in further institutional fixes (Ferguson: 1994: 156). Identifying and evaluating problems becomes a technical search for new fixes, new

knowledge, and new methods. The lack of social, historical or political critique preserves the technical gaze. This means evaluation risks functioning as a way of justifying ongoing technical interventions.

Thirdly, anti-politics effects obscure power dynamics in diverse ways. Harris (2002) showed how conceptual knowledge can be deployed by organisational actors, such as the World Bank, to obscure power relations. Over 50 years ago, Mitchell (1955: 15, 123, 209) questioned the role and place of technical expertise in mechanically dividing up Egypt to govern its population. Bebbington highlighted how a focus on poverty reduction as a technically framed set of goals and impacts depoliticizes development, arguing that up to the 1990s, NGOs in the Andean region used to speak about social change, but now only about poverty alleviation goals (Bebbington, 2005; Banks & Hulme, 2012: 13). Bächtold's (2015: 2) research on results-based programming in Myanmar contrasts the 1988 discourse of confrontation, street protests, revolutionary ideas, and confronting the military, with that of modern aid, where: *"... reports and project documents produced today by organisations in the international aid architecture contrast sharply with this picture" (ibid: 4)*

For Bächtold, the results focus obfuscates political voices within a modern aid apparatus that produces inequalities and north-south "*power knowledge networks*". Anti-politics in this sense is a front that accompanies the "*technocratic turn*" (Ginty, 2012), in which TIEK becomes normalized in NGO practices. One might add that digital data, big data and data analytics represent newer development 2.0 oriented fronts (Heeks, 2010; Beer, 2018; O'Neill, 2014).

Gaventa and McGee (2013) discuss "Transparency and Accountability Initiatives" (TAIs), a form of accountability evaluation, which aims to ensure better services, budgeting and empowerment outcomes. They call for diverse studies of voice, context, the "*black box of political will*", the relegation of best practice, and promotion of contextual sensitivity (ibid: S22). As Fox (2007) points out, excessive emphasis on scientific rigour, best practices, new technologies and technical approaches to TAI can submerge the politics of accountability (Fox, 2007) meaning we must question what TAIs really are and what they really do. The risk is that the technical paradigm detaches from practice, and pays scant attention to contexts, power or politics (Gaventa & McGee, 2013: S22).

Fourthly, these kinds of anti-politics configurations appear to work at large and small scales. Scott's book "Seeing Like A State" (1998) described large-scale nationwide modernization

failures that ignored local context, know-how and participation, or what Scott termed "*metis*". They depended instead on "*techne*", technical planning rationalities that failed to capture the complexity of people's lives. In Scott's country cases, technical knowledge was mobilized leading to social change and new inequalities of power. In another study, Gabay (2012) discusses how the MDGs regulate and engineer what development is, showing how interventions always meet resistances of different kinds, not just political ones, but historical, cultural, or simply contingent ones. Gabay describes how slum dwellers were moved to tower blocks outside of Mumbai in India as part of a program to meet MDG goal seven[14]. Unexpectedly, youth crime increased as extended families became separated following the moves. Some participants did not understand new plumbed toilet systems in the new apartments and threw their defecated waste out of their tower block windows (ibid: 1262). Scott's and Gabay's works exemplify the unexpected large- and small-scale effects of ignoring local context, knowledge, needs and dialogue.

A fifth issue is the way anti-politics elides the complexity and depth of development. Simply measuring planned and technical outputs overlooks the "*systems, processes and institutions that perpetuate poverty*" according to Banks & Hulme, (2012: 13). This leads to interventions that foreground expertise, management and planning rather than community engagement (Power et al, 2002). This relates to the tensions between big-D development and small-d development. The anti-politics effect is that small-d community development is marginalised in favour of big-D managerial plans, discourse and knowledge fashions, wherein the latter fail to address the structural complexity or depth of poverty (Mohan, 2002: 22; Bebbington, 2005: 943; Banks & Hulme, 2012). The claim is big-D development prolongs poverty and increases inequalities.

Understanding such complexity and working with it is difficult and requires authentic engagement, as well as honest reporting on what is not possible in a given historical and political context. In her account of the $1 billion World Bank program to remake society in East Timor, Li described how aid experts deployed "*neoliberal knowledge*" of competition and accountability to remake communities, despite the very same experts admitting that society was actually "*too difficult ... to render technical or redesign*" (Li, 2007: 230). Technical

---

[14] MDG goal seven was environmental sustainability. Target 7d was to improve the lives of slum dwellers and monitor indicators for numbers of slum-dwellers and improved sanitation users.

discourse and anti-politics were evidently not "what works", to use an aid sector catch-phrase. If it is too difficult for embedded NGOs and expert consultants on the ground, one must ask what can be understood from reports and data at a distance? What do data and reports do?

Finally, many commentators have called for moving beyond technical discourse and anti-politics towards ways of re-embedding power, politics, alternative voices, more authentic accounting of empirical reality, and more honest evaluation into development discourse and practice. Sumner calls for alternative voices in understanding policy and practice complexities (2006: 648). Van Kerkhoff & Lebel (2006: 445) warn against imagining scientific evidence can do all the necessary work alone and argue for a mix of "*scientific, economic, social and political knowledge*" to be brought to bear on environmental issues. Keeley and Scoones (1999) highlight how policy makers work in complex environments and are not rational actors applying linear policies. Knowledge does not speak by itself (Latour, 1987) and increasingly science and society are seen as intertwined and messy (Law, 2004). Beeson & Islam note:

> "... *the contest of ideas in economic policy making can evolve independently of their intellectual merit and empirical credibility. Political interests shape and mediate the process within which policy debate unfolds.*" (Beeson & Islam, 2005: 197)

Many would agree with the need for greater discussion on how evaluation knowledge is entwined with power and politics rather than marginalizing the issue in practice. Alternatives proposed include re-legitimizing marginalised civil society politics and voices (Edwards and Hulme, 1996), recognizing how funding and reporting practices re-configure NGOs as technical service organisations returning "*investments*" back to funders (Duval et al, 2015: 49), rendering visible how the aid architecture constructs opportunities for control (Bächtold, 2015: 13), and finding ways to reinvigorate evaluation as a reality check on technocratic or participatory development rhetoric (van den Berg, 2004: 68). Acknowledging markets and politics in aid chains may not be easy for development agencies or professionals (Hout, 2012: 417-418) and yet it is important in "*keeping the data honest*" (Norris, 2015: 139, referring to Donald Campbell), and differentiating aid altruism and aid investment (Duval et al, 2015).

The question left hanging from this review of anti-politics literature, is if and how, in the study to follow, small NGOs navigate anti-politics. Can NGO impact data/knowledge be "honest"?

2.4 Sharpening the focus on evaluation practices

The literature review has surveyed examples of technical discourse (TIEK) and socio-political perspectives on development impact evaluation, i.e. critical configurations. These are two contrasting domains of knowledge and understanding about contemporary data/knowledge intensive impact evaluations in the development 2.0 landscape. The review of both helps us to sharpen the focus of the study, to move from TIEK prescriptions and critical configurations towards how power is entangled with data/knowledge, within impact evaluation practices.

TIEK data/knowledge is predominantly targeted at and circulates around professional and administrative audiences in development bureaucracies, networks and markets (Wallace et al, 2006; Mosse, 2004a; Quarles von Ufford, 1988; Ilcan & Philips, 2008). This flow is reversed for data capture, where data is primarily captured from intervention target communities, clients, households, business owners, and local workers (Hyman & Dearden, 1998). Thus, the evaluation input/output information meta-model separates data suppliers and data consumers, constituting a professionalized data flow (Ilcan & Philips, 2008), raising questions about the value of evaluation broadly (Barr et al, 2016), and its less clear value for beneficiaries (Duncombe, 2009; Wenar, 2006)[15]. Both scientific evaluation discourse focused on effectiveness, attribution, and empirical robustness, as well as participatory discourse foregrounding local empowerment narratives are evident in TIEK prescriptions.

The critical configurations illustrate how the TIEK method plus results frame and information centric input/output meta-model have repercussions for small NGOs, evaluation agents and evaluation relations. The configurations involve small NGOs, frontline data/knowledge work, and the evaluation power/data/knowledge silences, transformations and inequalities that emerge and become sustained in practice. From the review, three implications emerge, the first two as foundational concerns for the following chapters, and the third a caveat about the wider the corpus of literature.

The first implication is that the history, diverse demands and evaluation configurations constitute a broader range of issues than those confronted by TIEK. TIEK's normative prescriptions produce results that are primarily understood as information to be used in rational

---

[15] There are calls to make the evaluation processes more open and provide more feedback to beneficiaries (e.g. Groves, 2015: 9 and 13), but these are not the primary goals found in TIEK discourse and architectures.

decision-making processes within aid organisations. The breadth and depth of demands and configurations mean that TIEK provides an insufficient response to the problems of power/data/knowledge relations in development impact evaluation. Methods don't explain broader configurations and do not account for aid chain practices or power. Results don't acknowledge power or practice in bureaucratic flows, market competition, or how impact representations are shaped outside of professional standards and empirical data. The target community is the focus of the evaluation gaze, and yet the results are not intended for target communities, but for professional communities, aid chain agents, and decision-makers, all of whom are almost invisible in the TIEK audit gaze.

Key issues were raised in the critical literature around how small NGOs deal with critical configurations around impact evaluation work.

Regarding evaluations, how do small NGOs:

- deal with evaluation issues beyond target sites, TIEK prescriptions, or evaluator skills?
- make space for diverse voices beyond expert problem-solution frames?
- account for hidden politics or how methods change over time?
- balance globally sourced methods and local ways of knowing impact?

Regarding NGO work itself, how do small NGOs:

- tend to the needs of distant contexts and stakeholders versus local ones?
- navigate communications, branding and marketing needs?
- manage expert technical knowledge?
- balance evaluation office work versus field work?

Regarding politics and governing, how do small NGOs:

- weigh compliance and market rationalities against relationship building?
- balance the construction of mobile data/knowledge for professional communities versus local stakeholders?
- navigate technical results and anti-politics pressures alongside political negotiations, with local agents and aid chain agents?

The second implication that emerges from the literature, is the need for a perspective on evaluation that privileges organisational, social and political practices. Such a view of evaluation is a) critical in terms of acknowledging broader socio-economic power dynamics embedded within data/knowledge construction processes, and b) engaged in terms of recognizing and responding to, not just interpreting from a macro, geographic, scholarly or sectoral distance, what evaluation data/knowledge does, at different points along aid chains. Interpreting without engaging risks creating more governing, market or colonizing demands and prescriptions for agents such as small NGOs.

This implication raises a question about three established evaluation perspectives and supports further investigation of a fourth. The first perspective concerns evaluation as technical or scientific. The second addresses evaluation as participatory, but still prescriptive, targeted and often technically configured in practice. A third concerns critical views of evaluation which articulate arguments for evaluation reform. However, there is a case for a fourth perspective which positions evaluations as critique plus engagement, simultaneously, partially and in situ. Mohan (2001) identified the problem that underlies this tension between critical and engaged views. He acknowledged the twin dangers of intervening and of passivity, in the face of:

> *"... our common subjugation to increasingly global material forces and the possibility of transformative dialogues [which] - make the need for alternative forms of collaboration urgent and pressing (Mohan, 2001: 167).*

Following Mohan, and keeping contemporary development 2.0 shifts, opportunities and the configurations discussed above in mind, how can an approach be formed for both critiquing and engaging with impact evaluation power/data/knowledge dynamics? Such an approach requires a deeper discussion of conceptual and methodological frames, the focus of the next two chapters.

However, before embarking on the theory and approach chapters, it is important to recognise a third implication from the literature. This is that there exists an extensive corpus of further literature which deals with the study of and analysis of data, information and knowledge in diverse scientific and socially informed ways. This corpus spans areas such as the philosophy of information and knowledge for example, or technical work on data analytics and data science. The philosophy of information for example draws on older empiricist philosophy and theories of mathematics in order to understand what information is (e.g. Adriaans & van

Benthem 2008a,b; Lenski 2010; Floridi 2002, 2011). Data science and data analytics foreground how data is analysed, interpreted, stored and communicated, drawing on disciplines such as statistics and computer science (e.g. Tukey, 1962; 1977; Cleveland, 2001; Press, 2018)

Of closest relevance in this wide body of literature is work that problematises power or inequalities. One example of this is work by Floridi (e.g. 2014; 2015), which analyses how explosive developments in contemporary Information and Communications Technologies (ITCs) have influenced who we are and how we relate to each other, our organisations and our economic sectors, from entertainment, to education, healthcare or even personal conduct. The fact that organisations now face challenges and difficulties "*all essentially driven by the recording, transmitting and processing powers of ICTs*" (Floridi, 2014: 168) means our production, consumption and mediation of data, information and knowledge, in evaluation or other such information intensive processes, requires serious attention. Floridi sees the circulation of information and knowledge as forms of power, wherein "*power is informational, and exercised through the elaboration and dissemination of norms*" (2014: 176). This contrast with power as a more physical, authoritative force.

Such "norms" amplified through the circulation of information and knowledge are a key aspect of the critical configurations discussed earlier in the literature. These shape the legitimacy of processes such as evaluation, its professionalised methods, acceptable results, relationships, and so forth. In evaluation, methods and results become desired by particular professionals, but inequalities or damage accrued due to the adoption of new data, knowledge or information norms, can remain unclear, even silenced. Floridi advocates not normative ethics per se as a way of dealing with these problems, but the design of information processes and systems which embed the need for ethical choices and reflection. He calls this an "infrastructure" for ethics, "*infraethics*", or pro-ethical design, contrasting it with ethics by design, and gives an example about organ donations (2015: 190).

> "*For example, strategies based on ethics by design may let you opt out of the default preference according to which, by obtaining a driving licence, you are also willing to be an organ donor. Strategies based on pro-ethical design may not allow you to obtain a driving licence unless you have indicated whether you wish to be an organ donor: the unbiased choice is still all yours.*" (Floridi, 2015: 190),

Activities such as obtaining a driving license, or donating one's own organs are well understood in terms of what decisions need to be made, when, and where. However, it is less clear how development impact evaluation systems and processes can be mapped and made more infraethics compliant – it may involve many choices at many sites of evaluation design, data collection, analysis and reporting for example. Where should ethical choices be built in and how should they be enforced in development evaluation? A more detailed map of the micro-practices of evaluation and the power relations involved, who makes what decisions when, who collects, extracts, uses or benefits from evaluation information for example, is required first, before an infrastructure for ethics can be designed.

The next chapter focuses on how data, information and knowledge are related with power dynamics in impact evaluation, in order to understand the socio-political aspects of evaluation in practice. It does not cover the breadth of the extensive corpus in information science, information studies, the philosophy of information or the technical prescriptions of data science and data analytics for example. However it does offer an alternative to contemporary models of data, information and knowledge which have, as will be argued, first of all become implicit and widespread in development evaluation practice, and secondly have become agential in evaluation power/data/knowledge relations, having marginalised the critical importance of power and politics in evaluation data, information, and knowledge construction practices.

3 Theory: Evaluation foundations – from DIKW to the 6Ps

3.1   The conceptual foundations of impact evaluation

The research questions in this study ask how impact data/knowledge is constructed, and how power relations and transformations are generated during impact construction. But, what is the conceptual or theoretical basis for capturing and analysing impact data, and then constructing impact evaluation knowledge products as results for organisational decision-makers? What is impact data? Or impact knowledge? What does impact data/knowledge intensity entail in today's development 2.0 landscape? These questions are the core concerns of this chapter. Responses to them help us to understand the foundations, assumptions and omissions underpinning contemporary development impact evaluation, its data/knowledge flows, its representational meanings, and the socio-political challenges of evaluation configurations. This chapter addresses these concerns by uncovering some of the conceptual foundations of data/knowledge intensive work today and offering an alternative conceptual frame that does not marginalise power and practice in development 2.0 impact evaluation.

These considerations all feature in related debates around the politics of data and knowledge (e.g. Beer, 2018; Gardner & Lewis, 2015), the politics of evaluation evidence (Eyben et al, 2015), and the importance of theorizing power in Information Systems (Walsham, 2001;

Avgerou, 2002; Hayes, 2011: 89) and Knowledge Management (Blackler, 1995; 2011; Easterby-Smith et al 2007: 75). However, development impact evaluation literature does not provide a clear theoretical grounding for power/data/knowledge relations, nor its products or processes. As we saw in the literature review, TIEK does not explain or account for the broad configurations. Therefore, one is left searching for a theoretical and conceptual grounding for how impact data, knowledge, and power are interrelated across aid chain practices, and how impact data sets or narratives move, change and mediate (in) development networks.

There are many models of data and knowledge production, but there is one dominant, pragmatic, and widely used model that cannot be easily avoided. This is the "data, information, knowledge, wisdom" model, also known as the "DIKW Hierarchy" or "DIKW Pyramid". The DIKW legacy in data, information and knowledge management is implicit and pervasive, and there are diverse reasons for this. DIKW[16] provides a model for making functional organisational knowledge; it is a linear, rational model, which frames processes as input and outputs; it adopts a data-centric perspective; and it positions data and knowledge as representational products, not social practices. The argument here is that DIKW, as a generic, implicit and pragmatic model, has not only been foundational in information and knowledge management (Jennex, 2009; Rowley, 2007; Lambe, 2011), but also in development impact evaluation too[17]. DIKW processes underpin impact evaluation processes, as in Figure 3.1.

---

[16] The "wisdom" element is not consistently applied or defined according to Rowley (2007: 164) and (Fricke, 2009: 3). It is problematic and controversial, and has therefore not included as a key focus in this study.

[17] The management of other conceptual property (e.g. media, content etc) is also important. However, the concern here is with power and practice in the use of data to construct and circulate knowledge, and as such, an in-depth interrogation of literature on media and content management is beyond the scope of this study.

Development Impact Evaluation

Data Capture ➡ Data Analysis ➡ Results

DIKW Foundation

Data ➡ Information ➡ Knowledge
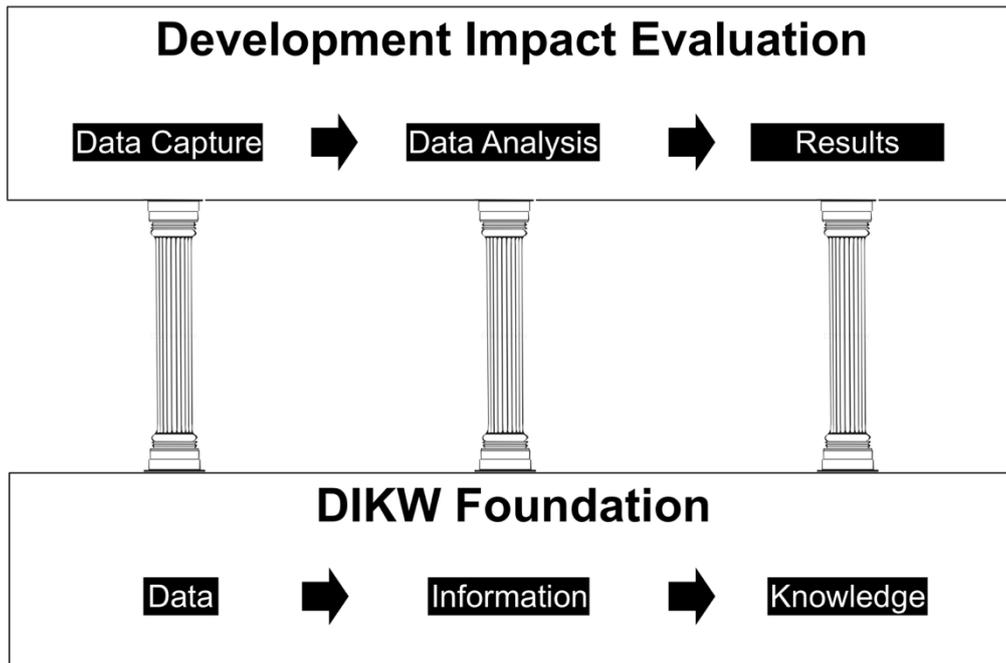
Figure 3.1: Development impact evaluation foundations align with DIKW

This chapter contends that in order to lend volume to issues of power and practice, and the CCs in the previous chapter, the DIKW model must be dislodged or de-centred from its pivotal role in supporting impact evaluation data/knowledge products and processes. An alternative set of sensitivities are elucidated, focusing on data/knowledge products, processes, power and political participation, as they occur within social or organisational practice. This alternative is labelled the 6P sensitivities, or 6Ps for short. The DIKW assumptions and omissions, and much Knowledge Management that draws on the DIKW legacy, have become naturalised within TIEK discourse and practice. This is despite the fact that DIKW and TIEK models follow linear prescriptions that do not sufficiently address complex or messy organisational practices; prize strategic and pragmatic knowledge products over reflective practice; and therefore, marginalise power in impact data/knowledge construction. The chapter therefore follows calls to acknowledge practice and social theory in Knowledge Management (e.g. Williams, 2014; Lambe, 2011: 194; Boell and Cecez-Kecmanovic, 2015). It advocates an alternative set of conceptual sensitivities in order to surface the issues of power, practice, politics and participation submerged by TIEK and DIKW.

Section 3.2 positions the view of power dynamics with related work in Information Systems, Knowledge Management, and Development Studies. Section 3.3 describes the DIKW and criticisms of the model. Section 3.4 recaps the TIEK / DIKW assumptions and omissions, before outlining the value of the 6Ps to elevate products, processes, power, political participation and practice in impact data/knowledge construction. Section 3.5 points to the role of the 6Ps within a comprehensive research approach, both conceptually and methodologically.

3.2 Power as part of data/knowledge intensity

This section positions the problem of how power, data and knowledge are intimately related vis-à-vis germane theoretical work from Information Systems, Knowledge Management and Development Studies. All these literatures feature calls for more research into power relations, but the topic remains marginal in each. The positioning of the theoretical focus does not provide exhaustive accounts of work in each of these fields, but canvases indicative examples to set the scene before the more in-depth discussion of DIKW and the 6P sensitivities in relation to impact evaluation that follows.

3.2.1 Power in IS, KM and Development Studies

Information Systems literature investigates diverse issues around information and power, and also shows how studying power is often avoided. Markus's (1983) paper drew attention to how a management information system (MIS) could transform data flows by changing who has access to what data, and the resulting political conflicts between organisational departments (ibid: 438). Feldman & March (1981) highlighted the importance of information in professional and organisational symbolic behaviours. This contrasted with the established view which suggested information resources are primarily deployed in rational decision-making processes. Walsham's (2001: 57-58) review of power in Information Systems acknowledged diverse concerns: how the language of efficiency is used to gain power (Kling

7Iacano, 1984); the importance of Foucault's work on power/knowledge inseparability, and how techniques and procedures specify legitimate accounts of truth via regimes of truth (Foucault, 1980); how information systems contain data that are representations of knowledge

and truth (Haraway, 1991); and the importance of surveillance, as a form of control (Zuboff, 1988; Foucault, 1977; Knights et al, 1993: 990) and in monitoring (Lyon, 1993; Attewell, 1991). However, Willcocks (2004) and Willcocks & Lioliou (2011) contend that power/knowledge research has often been avoided in Information Systems, even after seminal works advocated critical approaches (e.g. Orlikowsky & Baroudi, 1991). In Wilcock's (2004: 270) view, although Zuboff's "In the Age of the Smart Machine" (1988) became a bestseller in popular technology and management fields, its Foucauldian influences "*fell away almost completely*" as others took up Zuboff's insights. Furthermore, Introna (2003: 239-241) observed how power/knowledge issues in Information Systems journals are often perceived as "*nuisances*". He argued that such nuisances actually constituted the very community and field of Information Systems itself. Willcocks (2004: 241) even suggested that Foucault would study marginalisation within sites of information/knowledge work today in order to understand contemporary power/knowledge dynamics, such as those embedded in "*digital economy rhetoric*" (Willcocks & Lioliou, 2011: 6).

Knowledge management has also featured work on power relations. Easterby-Smith et al (2007) acknowledged the importance of power relations. Hislop et al (2000) looked at knowledge controlled by competing groups in organisations, and Hayes & Walsham (2001) described political enclaves in groupware-mediated knowledge work. Snowden (2002: 4) and Dalkir (2011) both describe different generations of knowledge management, in which the earlier generations either rode roughshod over "*primitive cultures*" (Snowden, 2002: 4), or saw top-down management seek to control knowledge assets, products and containers (Dalkir, 2011). Dalkir's emerging third generation features more top-down categorising and organising of knowledge, whereas Snowden sees a more positive and expansive view of knowledge management in products and complex flows. However, neither explicitly open up questions of power or politics. One key collection in the field is Easterby-Smith and Lyles' "Handbook of Organizational Learning and Knowledge Management" (2011), but power relations are referred to on only 10 of 700 pages, and power is not a central theme in any of the 29 articles. Plaskoff's (2011: 200) contribution does lament the lack of attention to power relations in communities of practice research (see also: Contu & Willmott, 2003), and Hayes' (2011: 89) contribution acknowledges the field's general neglect of power/knowledge. Power hierarchies are mentioned, but only as brief comments in the collected articles, and as related to organisational learning (Brandi & Elkjaer, 2011: 28; Roloff et al, 2011: 260), intercultural relations (Taylor & Osland, 2011: 587-90), or Asian knowledge management (Snell & Hong

(2011: 639). In other works, Spender & Scherer (2007: 13) reference Foucault (1980) in warning about the suppression of power relations inherent in claims of rational decision-making or guarantees of truth. In contrast, one author who has repeatedly foregrounded issues of power and politics is Blackler (e.g. 1993; 1995; 2009; 2011) and his contributions figure prominently in this study.

In Knowledge Management for development (KM4D), research has largely emphasised contextual studies over theoretical debates. Power is considered in terms of hierarchies and status, as well as how the powerful construct knowledge to serve their interests and positions. For example, van der Ham et al (2013) show how in lower- and middle-income countries, healthcare users are perceived as weaker than healthcare professionals due to the latter's knowledge and status. Swaans et al (2013) illustrate how innovation platforms must address stakeholder power differentials. Solnes Miltenburg et al (2013) discuss power hierarchies and knowledge cultures as they interact in Tanzanian healthcare settings. Millington et al (2016) propose a toolkit to empower communities to design their monitoring evidence. Deepek et al (2016) advocate community-based media and personal story production to offset knowledge inequalities, and Jenkin et al (2016) propose the use of films as forms of legitimate local knowledge construction. Such literature draws on bottom-up, community-based knowledge production as an alternative to top-down planning or evaluation controls, thus promoting empowerment for community groups, wherein a knowledge deficit is presumed to exist, and local cultivation is presumed to be a solution. Only Saito-Jenson & Pasgaard (2014) draw on conceptual resources from social practice theories, in their case actor network and research by Mosse (2005), to problematise project success as a result of translations between diverse actors in many sites. Similarly, Brown (2010) offers conceptual insights into multiple knowledge forms and the conflicts between them.

In Development Studies itself, theoretical analysis of power relations is more mature, spanning modernisation as economic catch up with the West during the post-colonial era (see Sumner, 2006), and dependency theory (e.g. Cardoso & Faletto, 1979) or world systems theory (e.g. Wallerstein, 1974) where power is part of macro-level economic and political inequalities. Participatory development has attended to micro-settings, emphasizing empowerment of specific local communities as a response to bureaucratic or market related inequalities (e.g. Chambers, 1997), or small-d development (e.g. Bebbington et al, 2008, Banks & Hulme, 2012: 13). Foucault's analytics of power/knowledge (1977; 2001) has been influential in Critical

Development Studies, and adopted by writers such as Ferguson, Escobar, Li, and Scott. Ferguson (1990/1994) described technocratic knowledge as producing anti-politics effects, because it ignored locally defined needs, aspirations, history, politics, and even economic reality. Escobar deconstructed development discourse, questioning its effects in development encounters (1995/2012). Li (2007) recounted a billion-dollar World Bank program to remake society in East Timor, where development experts deployed neoliberal knowledge (ibid: 230). Scott (1998) described large-scale modernisation failures that ignored local know-how and participation, relying instead on technical planning rationalities and schematic knowledge that failed to capture the complexity of lived lives. Thus, in Development Studies, there are mature conceptualisations of power relations, drawing on economic modernisation, Marxist political economy, and Foucauldian discursive/cultural analysis.

3.2.2 Positioning practice and power in data/knowledge construction

The current study straddles what might be considered a Marxist focus on activities of labour and production, and a Foucauldian concern with discursive formations. It situates power in the micro-practices of (NGOs) constructing impact data/knowledge products for exchange. Studies by Blackler (1995; 2011), Avgerou (2002), Brigham & Hayes (2013), and Hayes & Westrup (2014) encourage attempts to understand the production of data and knowledge and the accompanying power dynamics. Such works, in different ways, problematise power within the social practices of data/knowledge construction.

After critiquing five popular images of knowledge in organisations, Blackler (1995: 1023-1025) had argued for viewing knowledge itself as political processes. The five images focused on were: knowledge as concepts, cognition or abstractions in the brain (which he termed "*embrained*" knowledge); knowledge in our physical interactions in specific settings ("*embodied*"); culturally shared narratives, language or metaphors ("*encultured*"); knowledge embedded in organisational routines ("*embedded*"); and knowledge or information as encoded into books, or electronic media such as emails or databases ("*encoded*"). Blackler's perspective on knowledge as political processes - provisional, pragmatic, situated, contested and mediated - rather than a universal expert asset, problematised the power and politics inherent in analysing, managing, sharing, and storing, in other words, *doing knowledge work*. Blackler suggested that the five popular images of knowledge in fact aligned with "*the needs of*

contemporary capitalism", globalisation and technology-centric knowledge work (ibid: 1040), and thus raised issues of domination and subordination (ibid: 1042). Blackler's shift to knowledge as processes, opened the door to investigation of power and politics.

Referring to Lyotard (1984) and his notion of commodified knowledge, Avgerou (2002: 85-89) discussed how politics and power were implicated in the deployment of ICTs in global development contexts. Here, dominant market-based technological and managerial rationalities and discourses, encounter or subjugate local knowledges, such as shared family knowledge, traditions of conflict avoidance, communal or non-economic-centred knowledge, affectual, traditional or emotional knowledge (Avgerou, 2002: 77). Blackler's concern with situated activities of knowing, and Avgerou's concern with global rationalities complement each other in terms of how one might theorise power dynamics and effects where different knowledges meet in development evaluation encounters.

Brigham & Hayes (2013) looked at such encounters in an e-development initiative, arguing that the mix of technologies and conceptual models for evaluating performance and impact had the consequence of boosting the centrality of donor organisations in setting NGO goals and shaping their services. In a further article Hayes & Westrup (2014) interrogated the role of consultancy organisations in development, questioning their perceived objective, apolitical stance. Consultancies instead, sought to expand and stabilize their positions and influence in development networks, pervasively deploying technologies, models and frameworks in areas such as impact evaluation (ibid: 22). These tactics promote perceptions of professionalism, transparency and accountability, illustrated for example in an Oxfam report on effective consultancies in development work (Rowley & Rubin, 2006: 4), or in efficiencies and effectiveness as required in a DFID call to establish a global evaluation agreement (Lietch, 2012: 2). Consultancy tactics and discourses paint NGOs as unprofessional, ineffective, or lagging in terms of IT capabilities. This suggests consultancy tactics, measures and technical models used to fix social problems are not value-free or neutral, but politically geared towards reconfiguring perceptions of evaluating, perceptions of NGOs, and stabilising a consultancy's own positions and ability to secure future contracts (Brigham & Hayes, 2013: 25). In these ways, such models and frameworks appear to promote efficiency and professional values, but can function as Trojan horses for the promotion of neoliberal development agendas (ibid: 27). These critical perspectives draw on approaches to practices, institutional (Avgerou, 2002), and

socio-organisational, such as activity theory (Blacker, 1995; 2011), or actor networks (Hayes & Westrup, 2014) to understand power and practice in data/knowledge intensive work.

The current study theorises power/data/knowledge relations as key parts of impact evaluation data/knowledge construction, wherein certain impact data/knowledge is legitimised and made visible by prescriptive and widely distributed methods in response to sector demands, and where other data/knowledge, impacts, effects and configurations of the demand-method-result dynamic are discarded, delegitimised, or subjugated. Exposing and making palatable this bifurcation of data and knowledge for action by evaluators, NGOs or donors involves two arguments. It requires decentring the pervasive received wisdom of DIKW and the conceptual assumptions and omissions that underpin impact evaluation. However, exposure alone is not sufficient. It also involves creating alternatives and sensitivities to power and politics that evaluation data/knowledge practitioners can pragmatically reflect or draw upon. The next two sections consider firstly decentring DIKW and secondly, alternative sensitivities.

3.3   Dislodging the DIKW foundation

As detailed above, DIKW came to predominate in information and knowledge management from the 1990s onwards. In its various derivative forms (e.g. Ackoff, 1989; Zeleny, 1987; Davenport & Prusak, 1998), it has come to underpin impact evaluation discourse and practice, despite a problematic legacy (e.g. Williams, 2014; Lambe, 2010; Fricke, 2009). DIKW's historical roots are broad based, however, in the 1990s when it became popular, its literature reoriented towards technology, consulting and managerial functions (Lambe, 2011). DIKW's problematic legacy involves its data-centric foundational data, its pragmatic privileging of knowledge as products, and its focus on making such products for decision-making.

3.3.1 DIKW – still a widespread and implicit model

DIKW's credentials as a conceptual foundation for contemporary data/knowledge work, processes and flows are quite possibly unsurpassed. Walsham (2001: 36) called data "*the most fundamental elements of the digital age*". Tuomi (1999) describes the hierarchy (see Figure 3.2) as conventional, intuitive and generally accepted in Information Systems and Knowledge

Management. Fricke (2009: 1) referred to DIKW as part of the "*the canon of information science and management*", and Sharma (2008) claimed that DIKW is everywhere, like "*an urban legend", an information pyramid in information science, or a "knowledge pyramid"* in knowledge management. Lambe (2011: 187-189) views DIKW as a convenient and eagerly adopted "*iconic framework*" that congealed with Knowledge Management in the 1990s, becoming entrenched in Information Systems. [18]

DIKW Pyramid

Each level in the hierarchical model contains assets that can be processed to make the next level higher up in the pyramid.

WISDOM
Evaluated knowledge

KNOWLEDGE
Information applied to contexts

INFORMATION
Structured, organised data; meaningful messages

DATA
Objective, value free, discrete, facts about events; symbols stored in technologies
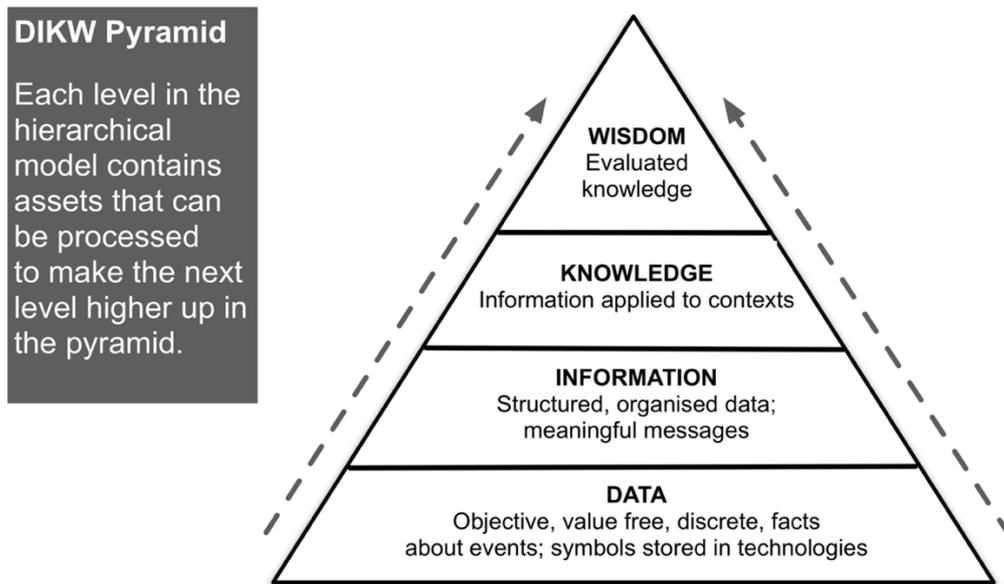
Figure 3.2: Typical DIKW model, with data as the foundation for knowledge

DIKW has been drawn on, used and altered in different ways, but remains an entrenched and taken-for-granted building block in information and knowledge management. Two of its main features are its view of data, information and knowledge as products, and of data as the foundation for knowledge.

---

[18] Boell and Cecez-Kecmanovic (2015: 2) cite Rowley (2007: 168) on DIKW as the "most prevalent approach to information in IS textbooks" and a "core of the knowledge revolution, information systems and knowledge management".

3.3.2 Data, information and knowledge as assets and products

The consensus in Information and Knowledge Management fields is that data, information and knowledge can be considered as assets or products (Walsham, 2001: 36) that can be stored or transmitted (Boland, 1987), for example through documents, databases, routines or brains. Davenport and Prusak's (1998) seminal text on knowledge working is a useful outline of this asset-based perspective.

What then is a data product? As fundamental to the digital age, data is considered as atomised, empirically derived properties, fixed in cells, and characterised as discrete, objective facts about events. Data do not explain why processes happen, fail or succeed, and are generally viewed as devoid of judgements, values or opinions (Davenport & Prusak, 1998: 2-3). According to Davenport and Prusak, data are stored in "*some sort of technology system*", inputted into systems by "*departments such as finance, accounting and marketing*", managed by Information Systems units, and evaluated by cost, speed, capacity, relevance or clarity. Banks and governments need data, and data needs to be managed as part of an organisation's culture. From this empirical perspective, data contributes to scientific legitimacy, but has no meaning by itself. It is a factual raw material used to make information (ibid).

Raw data assets (Earl, 1994: 59) are needed to construct the next level in the DIKW - information. In the model, information is a structured or organised asset derived from the raw atomised particles of data. Information is found in records or transactions, and can be described as meanings or messages for senders and receivers. In this sense, information changes the receiver's attitudes or actions (Davenport & Prusak, 1998: 3). Information products move around organisations in visible, definite networks of wires, post offices, and e-mail boxes, or in softer, less formal, less visible, ad hoc networks, or example in hand delivered notes or personal messages. Unlike data, information has meaning, relevance, purpose and context; it can be calculated or analysed before sending. Technologies perform such calculations while humans add context and purpose.

In constructing knowledge, the next level in DIKW, Davenport & Prusak argue that information must be contextualized, as the message enters the receiver's actions and decision-making context. Knowledge in this sense is deeper, broader and richer than data or information. It is complex, not neat or simple. Knowledge is found in individuals, minds, brains, or in books, routines and conversations (ibid: 6). Knowledge is:

"A fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of the knowers. In organizations, it often becomes embedded not only in documents or repositories but also in organizational routines, processes, practices, and norms."
(Davenport & Prusak, 1998: 5)

In describing data, information and knowledge, Davenport and Prusak's stated aim is to find out what knowledge is: what it looks like or sounds like, who has it, where it is, how to use it every day; and to help managers use knowledge to improve organizational effectiveness, efficiency, productivity and innovation (ibid: xi). Their book aims to answer the question: *"What do I do on a Monday morning to help make our organization's use of knowledge more effective, efficient, productive, and innovative?"* (Davenport & Prusak, 1998: xi).

The view of data, information and knowledge as products constructed in organisational processes fits well with TIEK, and the demands in the development sector for impact to be rendered as information-based results and delivered to organisational units responsible for decision-making on issues such as fund allocations, project efficiency, or policy development etc. Groves' (2015) model for analysing impacts and disseminating results is implicitly based on this product view, as are the results-based management foundations of the UN Women's Gender Evaluation Handbook (2015). The OECD-DAC principles (1991), 3ie definitions (2012), and the gold-standard experimental RCTs (Duflo & Kremer, 2005) are all founded on capturing empirical data for analysis and producing robust impact narratives from that data. Similarly, Davies & Dart's (2015) "Most Significant Change" qualitative approach also views impact narratives as collected, documented and passed up the organisational hierarchy for review and selection. Wallace et al's (2006) depiction of the aid chain in which NGOs work also sees an exchange of assets – reports and bids exchanged for financial investments.

Development sector knowledge management practitioners recognize the tendency for DIKW perspectives to reify assets and strip context, but also see the curation of assets as part of their work, for example, in how Web 2.0 can be seen as a "*codification adventure*" to make knowledge products (Acuna, 2012). Wilson & Best's (2005: iii) editorial for the Information Technology and Development journal signifies the importance of the consensus on data,

information and knowledge as products in ICT4D. They signal DIKW principles in "*the way data becomes information, and information may be turned into knowledge by individuals to aid in their various development aims*", and also in how professionals seek the right information "*out there*" for capture and repurpose:

> *"We know the information we seek, which we could turn into useful knowledge for ourselves, is out there — "It's on the web" — or at least we believe it is. Too many times we discover that it isn't actually there. Hence, the purpose of this volume is to help create more of this knowledge that we seek—necessary knowledge." (Wilson & Best, 2005: v)*

In summary, the asset/product perspective in the DIKW, involves three conceptual stances:

- a reification of data/knowledge as things to be shaped for particular goals, but goals that are not questioned because the raw data foundation is allegedly judgement-free;
- a pragmatic view that is easy to understand for students, clients or non-specialists;
- an understanding that the product view cannot stand alone without acknowledging organisational or social processes used to construct or move data or knowledge assets.

What is most important for the current study, is that the reified and simplified product view lacks concepts, vocabulary or analytical purchase on power and practice, which are so important in understanding what impact data/knowledge can legitimately be, or what it actually does. Could a view of data, information and knowledge processes address power relations?

3.3.3 From products to processes

Rowley (2007) suggests that although DIKW definitions of products have remained consistent over time in Information Systems and Knowledge Management literatures, there are still many differences in how researchers see processes. The basic DIKW processes implicated in the product view relate to how data is made into information, and how information is made into knowledge (Figure 3.3). The data foundation remains empirical, judgement-free.

DIKW illustration of how empirical data enables construction of knowledge
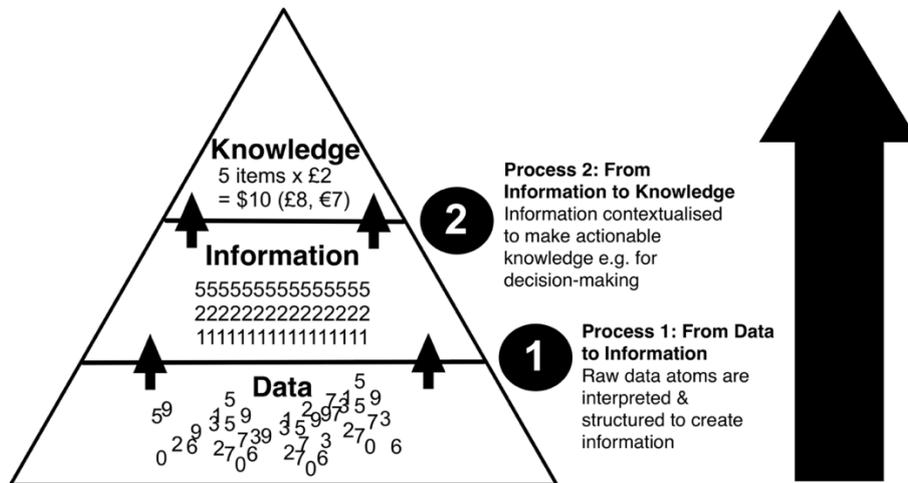
Figure 3.3: The DIKW view of data as processed to make information and knowledge

Research into the management of information and knowledge expanded through the 1990s and 2000s. It identified many diverse organisational processes, and advocated numerous prescriptive models in support of commercial innovation and performance improvements, but not so as to problematise data/knowledge power relations. For example, Leonard-Barton's "Wellsprings of knowledge" (1995) situates information and knowledge management within organisational learning, as does Von Krogh & Roos (1995) "Organizational Epistemology". Nonaka & Takeuchi's (1995) influential "SECI" model describes knowledge creation through processes of socialisation, internalisation, combination and externalisation, in support of product innovation. Choo (1998) describes a sense-making approach to knowledge creation, Wiig (1993) foregrounds knowledge as utilisation, and Bhatt (2000; 2002) emphasises knowledge work in quality management. Alavi & Leader (2001) advocate a framework featuring processes of knowledge creation, storage and retrieval, transfer, and application, wherein technologies facilitate access and storage of conceptual assets[19].

---

[19] Boisot's (1998; 2004) information space model (I-Space) stresses how abstract knowledge travels faster than situated knowledge. Technologies sacrifice the "richness and source of context" in favour of speed and efficiency (2004: 10). Such sacrifices are design choices for data/knowledge creation and movement (Boisot, 1998: 57, 157).

One paper that does draw attention to the social processes underpinning DIKW and knowledge management, if not power explicitly, is Tuomi's (1999) call to invert the DIKW pyramid. For Tuomi, knowledge precedes data. People's tacit knowledge needs to be shared or verbalised in order to structure information as textual representations and put them "*in a file, for example*" (ibid: 8). And information then needs to be broken down or atomised to provide data for manipulation in computer memory systems (Figure 3.4). To achieve this, databases require pre-designed grammars or structural syntax, which produce "*isolated*" and "*independent database entries*" that are "*detached from meaning*" (ibid: 107), i.e. cells. In Tuomi's inverse pyramid, data make knowledge - knowledge is required to make raw data or measurements.
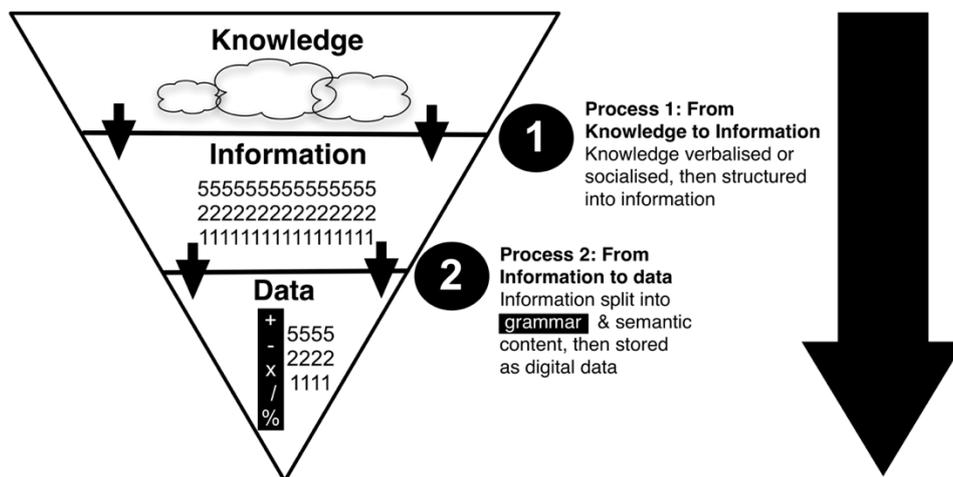
Figure 3.4: Inverse DIKW, where knowledge is required before making data

Work on incorporating social relations into knowledge construction had been evident in numerous places during this period. For example, in Nonaka's & Takeuchi's work (1995), which drew on Polanyi's (1986) tacit/explicit knowledge distinction; in Suchman's (1987) account of situated activity; in Lam's (1997) research on cross cultural engineering collaborations; and in Lave and Wenger's (1991) studies of communities of practice. Lambe adds that:

> "… data is the product of a knowledge-driven, purposeful piece of design work. The DIKW model implies the opposite, that knowledge is the product of a series of operations upon data." (Lambe, 2010)

Weinberger (2010) sees DIKW perspectives, such as those of Zeleny (1987) and Ackoff (1989), along with much of the DIKW literature, as a managerial response to the information overflow in the information age, rather than a serious account of "*social, goal-driven, contextual and culturally-bound*" processes. However, neither Weinberger's reflections nor Tuomi's shift to the social nature of knowledge, raised issues of power or control. The focus of research remained limited, on leveraging social complexity to boost organisational performance.

Dalkir (2011) and Snowden (2002) both outline three generations of knowledge management practice. Dalkir see a first top-down phase, involving containing and storing knowledge in databases and documents; a second, bottom-up, people- and community-centred phase drawing on communities of practice; and an emerging third phase that emphasises once again content and its organisation for access and use. Dalkir advocates integration of a series of knowledge management processes for capturing, selecting, codifying, refining, sharing, accessing, applying, evaluating or reusing knowledge (Dalkir, 2011: 53-56). Snowden (2002: 4) by contrast, saw a first, pre-1995 technology-centric generation. It featured managers and consultants riding "*roughshod across pre-existing 'primitive' cultures with the intent of enrichment and enlightenment that too frequently degenerated into rape and pillage*". A second phase was ushered in by Nonaka & Takeuchi's "SECI" model (1995), although Snowden suggests that Nonaka recanted his view of tacit/explicit as separate phenomena in favour later of a dialectic (Nonaka & Konno, 1998; in Snowden, 2002: 5). Snowden asserts a third phase for managing conditions, flows and complexities, not just moving assets. Snowden labels his model with the Welsh term "Cynefin", which translates as "habitat". The model embraces complexity beyond scientific or Taylorist rational thinking.

Finally, Williams (2014) reviews various DIKW and knowledge models in his recent paper and argues that a distinct new model is required. His "Action Knowledge Information" model (AKI) foregrounds actions and removes hierarchical dependency on data. Williams makes action a key pillar, but along with Dalkir, Snowden, and other process-based perspectives, he does not explicitly address the DIKW power lacuna. Snowden (2002) says we must understand knowledge as products and processes, things and flows, but power relations, despite the roughshod riding of bygone days, are not considered in today's chaos or complexities (ibid: 2).

3.3.4 Criticisms of the DIKW legacy in Knowledge Management

DIKW has attracted inadequate critical attention according to Rowley (2007: 163), although there have been direct philosophical and historical appraisals. Fricke (2009: 1), for example, interrogates the logical and methodological foundations of DIKW as the "*canon of information systems and management*". He argues, following Popper and Kuhn, that the raw data foundation is not a secure bedrock; that DIKW encourages "*mindless and meaningless*" (ibid: 5) volumes of data capture in the hope of refining knowledge later; that data may be true, false, fallible, weak, uncertain, non-empirical, future-oriented, or fictional; that information and knowledge are closely related, but contextually diverse; and that DIKW does not deal well with knowledge as recordable or explicit. Fricke advocates abandoning DIKW, although he warns of the "*intellectual and theoretical vacuum over the nature of data, information, knowledge, wisdom and their interrelationships*" (ibid: 2) that will remain in its host disciplines. Another view draws on Wittgenstein's later philosophy (1953), where Boell & Cecez-Kemanovic (2015) argue that information should not be seen as representational alone, but as contextual in use, for example by IT staff, bankers or customers.

In tracing the genealogy of knowledge management research through the economists Machlup (1962) and Arrow (1962), Lambe (2011) argues that 1990s approaches to knowledge management were born out of technology and consultancy agendas, and were clothed for adoption in the management of organisations (Koenig & Neveroski, 2008: 244; Lambe, 2011: 176). In this formulation, 1990s knowledge management forgot it's deeper, wider economic, societal, educational, and public policy roots. Wallace (2007: 2) goes so far as to say that knowledge management was truncated by Nonaka & Takeuchi's work in 1995, its roots in philosophy, economics, education, and information and library studies severed. The role of DIKW in Knowledge Management grew out from Computer Science and Information Science developments in the 1970s-80s, as those disciplines sought to establish their legitimacy and utility for management and decision-making. DIKW lent credence to these fields, linked data to good decisions, and supported the establishment of technology units in organisations. Williams (2014: 9-10) argues that DIKW was a functional, intuitive, attractive and "*slick*" kind of model. For Lambe (2011: 189). it was convenient and explanatory too, appealing to technologists and executives who wanted stable, logic-driven rationales. As such, DIKW spread throughout knowledge management in the 1990s, privileging commerce, consulting and

computing concerns (ibid: 189). Lambe concludes that DIKW has now become an obstacle. Knowledge Management has become reliant on a DIKW legacy of technical, data-centric assumptions, thus omitting insights from social theory (ibid: 190). It is now in a "*malaise*", and unable to evolve with reference to its deeper, richer historical roots (ibid, 2011: 191).

Practitioners are also aware of the limits of the DIKW model, but depend on its enduring utility. This is illustrated in a 2010 Knowledge Management for Development "D-Group" online forum discussion (see KM4D, 2012). In 2016, there were over 800 members in this group, and many contributed to the thread, citing limits to DIKW such as: a lack of cultural sensitivity; an inability to deal with diverse forms of knowledge (e.g. Zen Buddhism versus London cab driver knowledge of city routes); that DIKW was a managerial model that privileged computer/information science perspectives; and that it rendered data and information as objective and tangible in technologies, but knowledge and wisdom as inner, subjective mental phenomena. At the same time, commentators pointed out DIKW's continued usefulness and relevance. It provided clear definitions, furnished explanations, it acted as a clear model for non-specialists, non-KM people, departmental managers, and students, and it helped in dealing with clients and remained a useful resource for experts to return to and refresh their understanding. One commenter argued it was still better to base social development decisions on DIKW evidence, rather than faith alone. DIKW's continued implicit diffusion is also demonstrated in a recent KM4D journal call for papers (KM4D, 2016, August 24), which elicits research into organising data, curation, data quality, usability, access, sharing and ways of facilitating these processes. The community discussion cited above and call for papers both suggest that power and politics in knowledge management remains slippery and difficult to grasp.

As argued earlier, impact evaluation is founded in practice on a method-plus-results frame, a meta-model of data capture and analysis as inputs and outputs, and the supply of impact products as results delivered to organisational decision-makers. Thus, impact results, whether constructed through technical scientific methods or participatory ones, both of which target sites of poverty, are data/knowledge assets that flow across agents and organisations in aid chains and networks. Such data/knowledge flows implicate relations of power and politics according to authors in Development Studies, Evaluation, Knowledge Management, and Information Systems research (e.g. Escobar, 2011; Eybren et al, 2015; Picciotto, 2013; Norris, 2015; Gardner & Lewis, 2015; Mosse, 2011; 2013; Blackler, 1995; 2011; Hayes, 2011; Hayes

& Westrup, 2014; Walsham, 2001; Avgerou, 2002). As such, this detour into DIKW has been critically important, for understanding how impact data/knowledge is conceptualised as products, assets, processes and flows for managerial decision-making and functional utility. Issues of social practice, power and politics are largely elided in the DIKW legacy, but can be traced back historically to its 1990s embedding in and adoption by computing, consulting, commerce and management practice (Lambe, 2011).

DIKW therefore carries much weight in the erasure of power and politics (Figure 3.5), in favour of a technical, data-centric foundation for knowledge production (c.f. Ackoff, 1989; Davenport & Prusak, 1998). Knowledge management processes omit power and politics, but do leverage social practice and recognise how knowledge and complex design processes precede data construction (e.g. Tuomi, 1999; Jennex, 2007; Lambe, 2011). Critiques of DIKW have identified philosophical, methodological, and historical problems with the model, as well as a need to bring more social theory into our understanding of DIKW and Knowledge Management. But again, these enquiries failed to foreground power or politics. Even Jennex's review of DIKW (2007), which argues for a two-way flow, a synergy of Ackoff's perspective from data-to-knowledge and Tuomi's knowledge-to-data depending on contexts and goals, fails to question goals or address power dynamics. Snowden (2002: 4) admitted the older generation of knowledge work did *"rape and pillage … primitive"* cultures, but did not problematise what it does for today's generation. Fricke (2009: 9) accepted that representations did *"carve up reality"*, but did not discuss the implications of this. Tuomi (1999: 111-113) referenced Polanyi (1986) and Vygotsky (1962) regarding the mediation of knowledge or data as peripheral or non-focal, acknowledging the importance of social institutions, and divisions of labour. However, Tuomi did not deal with how institutions might mediate power/data/knowledge relations.
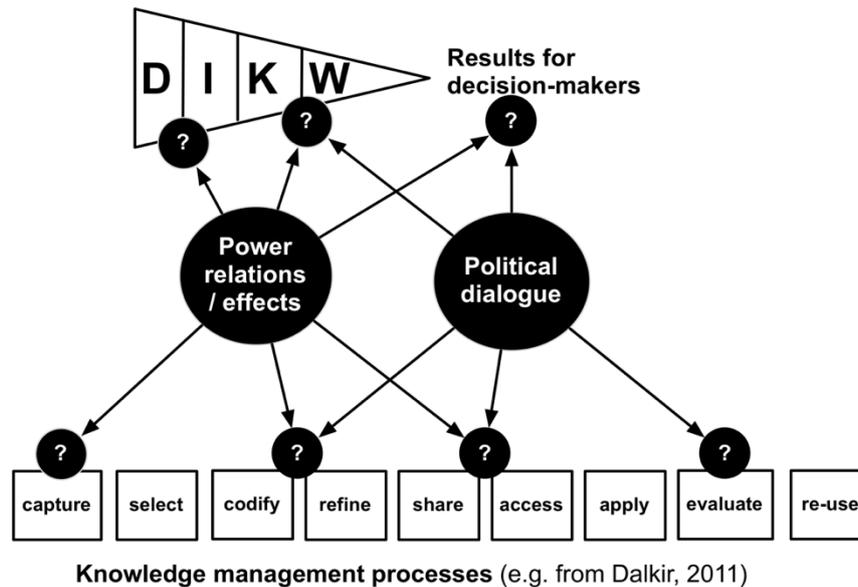
Figure 3.5: Where are power or politics in DIKW and knowledge management?

In conclusion, if an implicit model does not address goals, materials, institutions, power, politics, practice or participation in processes of data/knowledge construction, then it contributes to eliding them. What is required is a reinsertion of power, practice, politics and participation into impact evaluation, in order to decentre the DIKW legacy of omission.

3.4   The 6P Sensitivities: an alternative perspective

3.4.1 Reviewing the assumptions, omissions and configurations

As stated at the start of this chapter, to address the theoretical problem of power/data/knowledge relations it is necessary to decentre the inadequate DIKW model of knowledge production that underpins impact products and processes in the data/knowledge intensive development sector today. The focus on TIEK prescriptions, managerial decision-

making, and empirical data as the foundation of robust evaluation, means that social practice, power relations, political dialogue, and critical configurations beyond functional efficiency, effectiveness and performance have often been marginalised or elided in evaluation. One is left wondering where power, practice and politics are located in the DIKW legacy, in managing knowledge, or evaluating impacts, because they are not explicitly dealt with in TIEK or DIKW (Figure 3.6). Such omissions relate directly to the configurations, concerning how evaluation is understood, how NGOs change, and how impact knowledge is governed. This is because impact evaluation data/knowledge flows rely on the implicit conceptual foundation of DIKW and knowledge management processes that emphasise data-centric views, technologies, and managerial controls, not power or politics, in the construction of impact representations.
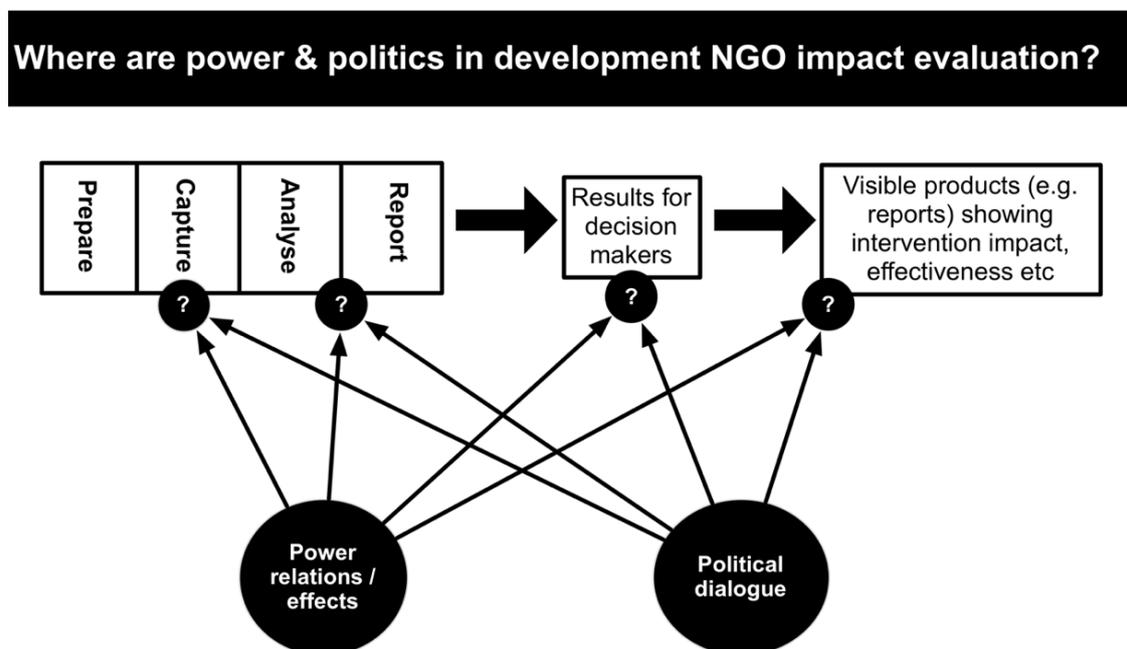
Figure 3.6: Where are power or politics in impact evaluation practices?

In response to such omissions, this section raises a set of sensitivities, which help in reasserting and reconceptualising power and practice in impact data/knowledge construction. These sensitivities acknowledge the DIKW legacy, knowledge management assumptions and omissions, and also the configurations brought about by development sector impact demands and TIEK prescriptions. Table 3.1 recaps the assumptions and omissions evident from the discussion of DIKW and knowledge management models, along with the configurations from Chapter 2.

DIKW legacy and KM assumptions

- Early DIKW data as empirical, factual foundation
- DIKW/KM as pragmatic; goals not questioned
- DIKW products represent real world states
- DIKW/KM technology leads to better decisions
- KM leverages social insight to serve agent goals
- Later DIKW / KM accept design precedes data
- Simplicity needed for clients, students, non-experts
- DIKW/KM focus on management / IT since 1990s
- DIKW/KM is evidence-based, not faith-based

DIKW legacy and KM omissions

- Methodological & logical flaws in DIKW relations
- Data is diverse; empirical foundation not secure
- Not all knowledge is constructed from empirical data
- DIKW/KM social sciences roots truncated by 1990s
- DIKW/KM elide social/organisational practice
- DIKW/KM elide power relations / configurations
- DIKW/KM elide politics & participation
- DIKW/KM elide insights from critical / social theory

Critical configurations

- How evaluations are understood and change
  - Seeing beyond individuals or methods
  - Seeing expert problems and diverse voices
  - Seeing hidden politics and method effects
  - Seeing the provenance of technical discourse
- NGO transformations
  - Outsourced service providers
  - Data/knowledge displaces field work
  - Confusion, capacity & commodities
  - Communications, marketing & branding
- Governing & anti-politics
  - NGO impact governing
  - Impact data/knowledge mobility
  - Impact anti-politics

Table 3.1: DIKW / KM assumptions and omissions, and critical configurations

3.4.2 An alternative – the 6P sensitivities

Impact evaluation power relations and configurations require critical interrogation and renewed engagement because these configurations result from the combination of sector demands and TIEK discussed earlier. The DIKW model and its legacy in knowledge management carry into evaluation discourse and practice assumptions and omissions that foreground technology-centric data foundations for managerial strategy and decision-making. AT the same time, they background approaches to social and organisational practice, power relations, and political dialogue. To avoid eliding the power relations generated through impact data/knowledge construction, an alternative theoretical perspective must take account of data/knowledge as products, processes, power relations, and political participation within accounts of specific social or organisational practice. Why we need to pay theoretical attention to each of these 6P sensitivities is outlined below.

Why products?

A critical view of power/data/knowledge requires us to focus on products for a number of reasons, Firstly, because the sector impact configurations are tangible, and will not simply go away. Practitioners are under pressure to produce impact data/knowledge products or assets for decision-making, and this requires management of products and assets through capture, selection, codification, refinement, sharing, access, application, evaluation or reuse etc (Dalkir, 2011: 53-56). Secondly, the DIKW discourse is still widespread if implicit in sector work processes, in evaluation, policy analysis, project planning, funding decisions, and performance management. This sustains a product-centred focus, and the need for data/knowledge products to serve organisational goals that are negotiated and acted upon through reports, emails, meetings, arguments, claims etc. These draw on empirical data as assets or narrative arguments to satisfy specific demands such as project goal success, comparing project costs and benefits, analysing policies, or mitigating risks. Thirdly, it is wise to pay respect to products in order to facilitate engagement with practitioners, evaluators and NGOs who may not be philosophers, critical analysts, or have time and resources to develop responses to neoliberal development.

Practitioners must face Davenport & Prusak's (1998) pragmatic everyday data/knowledge work demands. Starting an analysis with practitioner products and demands for those products legitimates the practices of NGOs or evaluators within the analysis of impact evaluation. Spender & Scherer (2007: 7) sum up the need to take seriously knowledge as products or assets:

> *"Whatever doubts we academics harbour about the much-trumpeted paradigm shift into the Information Age, practitioners generally see KM as driven by competitive pressures and the need to manage the organization's intangible assets more efficiently."*

Managing impact data/knowledge remains crucially important in the development sector. NGOs are under intense pressure to create products, and to market impact and effectiveness as sharable assets. Whatever doubts researchers harbour about knowledge conceived as assets, practitioners see data/knowledge products, driven by competitive and bureaucratic demands, as requiring serious management and attention (ibid). The product bias is not going away, but at the same time, interrogating the politics of knowledge generation (Gardner & Lewis, 2015: 179) is an important progressive aim. Viewing data/knowledge as products constructed for circulation in aid markets and bureaucracies (Mosse, 2004a) enables researchers or stakeholders to problematise power and politics as part of normalised, mundane knowledge work processes, where power may be elided or silenced in order to stabilise assets for exchange. These products are part of what Wallace et al (2006) see as rational management, or Miller & Rose (2008) consider as part of expert governing practices.

This applies equally to commercial business and the new data/knowledge intensive aid environment discussed in the literature review. Products are entry points for examining the construction or deconstruction of those products. Mosse's (2004a: 77) discussion of "*social knowledge products*" and anthropologists at the World Bank exemplifies this kind of analysis. Mosse looks at how and why these products are constructed and adapted. World Bank anthropologist knowledge products should not be too radical or drift too far from the economic discourse prevalent in the bank, but nor too conservative because they must add new value to economic-centred views. From Quarles van Ufford (1988: 77) we learn that ideas, goals and knowledge are produced in aid bureaucracies that seek their own maintenance and survival (Mosse, 2004a: 77), and in combination with Avgerou's discussion of global rationalities (2002) and Craig & Porter's (2006) travelling rationalities Identifying such rules, incentives, and forms of accountability and transparency requires an approach to data/knowledge products,

that begins with initial engagements with practitioners around what they consider to be their data/knowledge products.

In attempting to transcend DIKW, Williams (2014: 17-18) argues for more attention to the specific elements of information and knowledge, many of which can be considered assets or products. Information may include facts, fiction, data, records, designs, and laws. Knowledge incorporates beliefs, values, concepts, ideas, insights, skills, and truths. Identifying such conceptually rich products enables attention to them as situated and contestable in Blackler's terms (1995), and thus exploration of their mobility, governing roles, and relations of power in practice. Data/knowledge products then are a point from which to start an analysis of power relations and configurations, and initiate engagement with practitioners. Practitioners construct products, researchers deconstruct them. This interplay means the products are crucial for studies of power/data/knowledge.

Why processes?

Products showing the results of development regulation and policy are privileged over the processes that produce them (Escobar, 1995; Mosse, 2011: 2). Yet processes are important in studying power/data/knowledge relations for a number of reasons. Products need constructing, and processes are implicit in the making of data/knowledge products. As with products, practitioners will be aware of many of the processes involved in producing impact data and knowledge. Evaluators will be able to describe evaluations processes. NGO staff will be able to talk about aspects of their projects, their intended impacts, how they go about making reports, recording data etc. TIEK prescribes a variety of process options, for example, in conducting qualitative evaluation, or in applying a statistical model to data collected from the field. Furthermore, researching processes also means engaging with what happens to data/knowledge as it moves in the organisation: how it is captured, selected, shared, applied, in support of specific organisational functions and decision-making. And finally, DIKW, knowledge and information management speak to diverse processes, whether producing information from data (Davenport & Prusak, 1998), externalising or internalising tacit and explicit knowledge (Nonaka & Takeuchi, 1995), using knowledge to construct data (Tuomi, 1999), or the processes involved in managing knowledge to support other organisational functions. However, Tuomi's move towards social processes for data/knowledge construction,

and his referencing of institutions, rules and divisions of labour (ibid: 113) only opens the door so far on social processes in managing knowledge, albeit beyond the "*What shall I do on Monday*" pragmatics of Davenport & Prusak (1998: xi).

Blackler's (1995) expands the knowledge management concept of assets and products in brains (embraced), documents (embedded), culture (encultured) etc to knowledge as itself a process of knowing, enacted in different ways, contested, situated, partial and so forth. This allows in more overt analysis of power and politics, and shows that organisational activities are key to processes of knowing. More recently, Williams' (2014) AKI model foregrounds action as one of the three core concerns of knowledge management, but Williams does not focus on power. This is the case with virtually all formal DIKW and knowledge management models: processes are important, but power relations remain unclear.

The key concern regarding processes is how they can help us understand power dynamics in development impact data/knowledge construction. In KM4D Beaulieu's (2013: 37) research with civil society organisations doing knowledge work in Ghana concluded that looking at knowledge activities and processes enables an understanding of how stakeholders determine what counts as evidence or knowledge in policy debates, and argued that this process was rarely exempt from power issues. Such work, alongside Blackler's (1995) insights into situated knowing activities and Avgerou's (2002) discussion of global rationalities and subjugated knowledges, shows how a process-centred view of the construction and movement of data/knowledge products across practices facilitates greater understanding of power/data/knowledge flows and configurations. Products alone don't show these relations and configurations. This is why processes must remain a part of studying power dynamics in impact data/knowledge construction and management.

Why power?

Sensitivity to power relations in data/knowledge construction links back to development 2.0 opportunities and challenges, the silent configurations, and the assumptions and omissions of DIKW and TIEK discussed in the two previous chapters. Therefore, if one wishes to understand the effects of knowledge construction beyond prescribed methods plus results, then an approach to understanding power relations and inequalities becomes essential. The aim is not to define all power relations or preclude how to analyse power in these sensitivities, but to

advocate for their incorporation and conceptual inclusion when analysing impact evaluation processes and products.

Of course, as we have seen, there is little consensus over what power relations are or what they concern, and there are a number of related issues to be taken into account. In this chapter for example, the objective legitimacy of data was problematised (e.g. Fricke, 2007; Lambe, 2011), the knowledge and values used to make data were questioned (Tuomi, 1999; Lambe, 2011), the representational content was questioned in terms of its use in particular practices (Boell & Cecez-Kemanovic, 2015), top-down control of data/knowledge was queried (Dalkir, 2011; Snowden, 2002), and the omission of social accounts and contexts in which data and knowledge are produced, by models drawing on DIKW, was criticized (Weinberger, 2010; Williams, 2014; Lambe, 2011). From an Information Systems and ICT4D perspective, Walsham acknowledges power dynamics in relation to rhetorics of efficiency (Kling and Icano, 1984), the inseparability of knowledge and power (Foucault, 1980), surveillance (Zuboff, 1988; Foucault, 1979), and how power is avoided in Information Systems (e.g. Willcocks, 2004; Introna, 2003). Numerous authors have called for more attention to power relations in Knowledge Management too (e.g. Easterly-Smith et al, 2007; Hayes, 2011; Plaskoff, 2011; Spender & Scherer, 2007; and Swaans (2013)). In general, this chapter has argued for data/knowledge models to be more sensitive to power relations and effects, and to correct its elision in DIKW-inspired knowledge management research and TIEK prescriptions for evaluation practitioners.

Despite the difficulty of dealing with all the diverse dimensions of power, an approach to power in evaluation data/knowledge construction must select and describe one form or another. DIKW has not proved adequate for this task. If Walsham (2001: 56) is right and power is of "*fundamental importance*", underlying "*all individual, organizational and societal practices*", because "*power is endemic to all human activities*". then impact evaluation and KM concepts, models and prescriptions must be broadened and deepened. Neglect and omission of such power dynamics becomes a site for research itself. The DIKW focus on products, and the knowledge management focus on processes for making products both require an understanding of power, of what is included or excluded from existing data/knowledge work, impact statistics or narratives. In short, power dynamics threaten to undermine both big development rationalities and small development engagements, and should be seen as endemic to both target sites of poverty, risk or marginalisation, and less visible sites within aid chains and networks.

This is a political challenge to techno-managerial views of impact that seek to govern other voices.

Why political participation?

Sensitivity to politics, participation and dialogue is important in trying to understand the widely diffused power relations inherent in the construction of impact data/knowledge. TIEK curtails opportunities for political processes beyond target intervention sites. This limitation is a key justification for expanding the scope of evaluation to include participation and politics beyond intervention sites alone, and into wider aid chains and networks.

Furthermore, DIKW concepts that underpin data-centric knowledge production in impact evaluation serve decision-making and organisational goals which are not necessarily universal, nor open to political engagement, nor explicitly implicated in any need to facilitate political understandings between aid funders, NGOs and participants. Related skepticism about participation as tyranny (e.g. Cook & Kothari, 2001), partnerships as means of control (Contu and Girei, 2014), or the decapitation of authentic participation (Leal, 2007) also need to be taken into account to find out if/how impact data/knowledge construction can support authentic possibilities for political participation that stretch beyond a project's targeted people and sites. DIKW risks supporting the smooth decapitation of participation because it offers no resources for the analysis of politics or participation. It thus elides them. Other knowledge management models that ignores political participation also risk reproducing a similar anti-politics.

Walsham (2001) considered politics to be "*attempts to use resources, including power relationships, to achieve particular ends*". For Fricke (2009), the carving up of data/knowledge representations to make such resources presumably involved values or goals, and similarly for Tuomi (1999), the knowledge used to make data draws on existing goals and values. Spender & Scherer (2007: 13) advocated embracing diverse values, discourses, epistemologies and interactions in Knowledge Management as a more appealing way forward than calls for a "*one true*" foundation. If diversity is to be achieved, then embracing differences in data/knowledge construction, finding alternative ways of doing it and of valuing knowledge, must involve a sensitivity to political differences and thus political participation and dialogue. Political participation must be acknowledged in data/knowledge construction if knowledge management

is to transcend its malaise (Lambe, 2011) and become a legitimate, inclusive and effective resource base.

The relevance of the politics of data/knowledge creation to development and evaluation has been recognised by numerous authors (e.g. Gardner & Lewis, 2015; Norris, 2015; Picciotto, 2015: 153; Eyben et al, 2015; Norrish & Sayce, 2004; Wallace et al, 2006). Escobar has called for a "*pluriverse*", of different views and exchanges (2011: xxxii-xxxiii; 2018). Ferguson (1994: 282-288) argued that engagement should be subject to it being specific, limited, conditional and marginal. Gaventa & McGee (2013) argue against the black-boxing of political differences in accountability evaluations, and Fox (2007) warns how overly technical approaches obscure politics. Such warnings are well-heeded in "small-d" development, which draws on community-based negotiations. But if development is constructed across larger aid chains and networks, then political representation and dialogue needs to travel these channels too, beyond the sites of the poor. Perhaps we need wider views of politics and participation, is required, one that promotes tools, devices and techniques to do political participation as part of standard evaluation processes. The alternative, as we have discussed, is anti-politics. Knutsson's (2014: 805) work with Rwandan civil society organisations (CSOs) shows how post-political governmental arrangements and resistance to them are dispersed, often co-occurring, in reporting practices, interpreting project goals or challenging technocratic development mechanisms for example. Politics, resistance and participation are not only at the site of poverty or marginalisation. Knutsson sees the smooth developmentality and anti-politics, following Lie (2015), in today's technocratic agents who assume anti-politics is the norm, not a side effect of development projects in the way Ferguson originally depicted it. Power et al's (2002: 77) observation that development as 'practice' is envisaged as and compartmentalised in field-practice becomes important here because field problems are "*not allowed to permeate the organisation as a whole*". Participation and politics must therefore be viewed across evaluation chains and networks, across reporting, regulations, data capture, and decision-making etc.

The diffusion of participation and politics outside the target site of a development project is described in Hayes & Westrup's (2012: 27) account of participatory development involving processes and strategies that render target locales actionable. Participatory strategies produce representations, which perform upon local settings, objectifying them and bringing them under the influence of development projects (Green, 2009). These strategies are deployed by "macro

actors", such as large NGOs, working across many locations, between boardrooms in New York and fields in rural India. These actors are more interested in and more capable of deploying such strategies, technologies, and reporting or measuring techniques than local participants. Thus, even participatory development, bottom-up knowledge management, and participatory evaluation involve diffused relations of power across networked practices, mediated by reports, technologies, and methods of measuring.

Two ways of approaching politics in development impact evaluation can be drawn from observations by Li and Dean. Li (2007: 10-14) discusses practices of government and practices of politics, in which political acts of resistance can puncture the expert and smooth technical development matrix. These punctures lead to practices of government that seek to close the punctures so as to maintain technical power/legitimacy. Li views political interests and agendas as part of the mechanisms of development, rather than masters over them. Acknowledging such practices and interests as part and parcel of inevitable political jostling, not as untouchable governing macro processes or contingent micro nuisances, opens spaces for research or dialogue about changes, impacts, and the unexpected results of development projects such as those that Li describes in detail. In a similar vein, Dean's (2010) descriptions of increasing technologies of government (such as TIEK) and increasing forms of participation, such as community action, together constitute new spaces, both for research into politics and for stakeholder political participation. Acknowledging both increasing controls and community responses is part and parcel of understanding the back and forth politics of development evaluation. In recognising political exchanges rather than eliding them, informed approaches to politics and participation are required that open up opportunities for dialogue in different spaces, experimentation, and often legitimate conflict. Social theories of practice and of networks have a role to play here. They can see impact data/knowledge as constructed across many aid-chain sites and practices, as representational commodities, as political strategies, and as opportunities for political participation. DIKW pyramids, expert knowledge management, and TIEK curtail such openings and shrink such spaces in the name of smooth effectiveness, efficiency and claims to govern.

Why in practice?

Understanding data and knowledge as part of practices expands our views of what they are and what they do, beyond representations in rational decision-making. In this vein, Feldman & March (1981) problematised the representational limits of information work, criticising dated engineering models and decision theories. They argued against organisations collecting gluts of information; data gluts would be of limited value in decision-making or planning anyway, and could be used for surveillance or post-hoc rationalisations after decision events. Thus, information enables information behaviours that are symbolic, not just instrumental, and which may signal professional virtue, competence, expertise or authority. According to Feldman and March (ibid), rational decision-making theories elide such social analysis. Boell & Cecez-Kecmanovic (2015) would agree with moves towards understanding information practices and usage. Nonaka & Takeuchi (1995) also describe socialisation processes in the construction of information and knowledge, and their work has fed into many accounts of knowledge management today, from Dalkir's (2011) multiple knowledge processes to Williams (2014) inclusion of activity. Spender & Scherer (2007) add that Polanyi's tacit knowledge, so influential in the discipline of Knowledge Management, similarly takes us away from information as representations and towards practices:

> *"The notion of tacit knowledge takes us beyond information and into a domain of practice, so setting up the possibility of a more complex three-way interaction between data, meaning and practice … [P]hilosophical support for this move can be found in the writings of Wittgenstein, Heidegger, Nishida and others. It helps us see how an epistemology of KM spins out of the organization's characterizing practices" (ibid: 14).*

Spender and Scherer cite Schreyögg & Geiger (2007) who identify an anxiety in our understanding of knowledge work and the limits of human rationality. This anxiety moves us away from data/knowledge as purely part of rational man's calculations and interests, or what Reckwitz (2002: 245) calls "*homo-economicus*", and away too from sociological rules, norms and consensus, or "*homo-sociologicus*". In practice-based views, data, information and knowledge become active elements within practice, not determinants of practice.

If practices are an effective lens for sensitising impact evaluation data/knowledge construction, and data/knowledge is part of our social practices, then what resources beyond DIKW do we have to analyse practices and to take account of data/knowledge products, processes, power and political participation in these practices? Before answering this question, a caveat is

required. We must be clear that a practice does not exist in a vacuum and should not be reified or bounded. Local knowledge, local practice, and local participation are not the possessions of an evaluation target group or an evaluator. One must be clear that practice, and power, are not simply localised at the development project target site. So-called "local" knowledge or sites are "*constituted by economic, social, cultural and political relations and flows of commodities, information and people that extend far beyond a given locality*" (Mohan & Stokes, 2000: 264). Participatory evaluation and bottom up knowledge management can guard against accusations of lacking authenticity by addressing data/knowledge circulation beyond, for example, each local site of data capture, knowledge sharing or impact reporting. This observation follows Hayes & Westrup (2012: 27) and Green (2009) in questioning performative representational techniques that are deployed to objectify the local and bring it under the influence of the unquestioned project. These representations perform work in practice, carving up Fricke's realities. In this view, impact evaluation is one such set of techniques. As such, critical examinations of knowledge management models and impact evaluation methods, involve power and knowledge in defining, arranging, and carving up many practices or relations. This means that concepts and models can and do shape practices, although not always in wholly predictable ways.

3.5   Towards the 6Ps for critical engagement

Set against the study backdrop of an increasingly data/knowledge intensive development 2.0, and the need to understand power/data/knowledge relations in development NGO impact evaluation, this chapter has outlined two conceptual moves. Firstly, it confirms the influence of the DIKW pyramid's legacy and its pivotal role as a foundation for development impact evaluation. DIKW and derivative knowledge management models articulate a conceptual basis for understanding evaluation products and processes in organisational exchanges in ways that scientific and participatory methods of impact evaluation have not explicitly theorised. This influence is arguably hard to underestimate as it is widespread, normalised and implicit in TIEK prescriptions. Scientific and participatory forms of evaluation prescribe techniques or commitments to equal participation at intervention sites, but they do not explain power inequalities across aid-chain practices.

DIKW is found to be, at best, inadequate as a conceptual basis for dealing with power, politics and practice in impact evaluation; at worst DIKW and related knowledge management views and tools risk actively contributing to the marginalisation of power and the critical configurations. DIKW's roots in economics, policy, culture, and society have been marginalised, superseded by narrow data-centric considerations and agendas from computing, technology, management and consultancy (Lambe, 2011 Weinberger, 2010). TIEK products and processes have been exposed to a similar trajectory. Consequently, disrupting the DIKW/TIEK influence by shedding light upon their assumptions and omissions - principally the elision of power and practice - has been the first theoretical move in this chapter.

The second move has been the exploration of an alternative conceptual basis for impact evaluation which does address power and practice. Thus, the second half of the chapter outlined a series of sensitivities, termed the 6Ps, which offer an alternative framework to DIKW/TIEK. The 6Ps support a re-embedding and renewed theorisation of practice, power and political participation in impact data/knowledge construction. This a) does not reject the development practitioner need to work with data/knowledge products and processes; and, b) does not ignore practitioners' diverse and situated responses to sector demands for knowing impact. This conceptual move involves experimenting with critical views and studying organisational practices within practitioner spaces. Approaching the practices of impact evaluation is essential to draw attention to the configurations, omissions and assumptions in evaluation work.

Operationalising the 6Ps as part of a critical approach to impact power/data/knowledge to be used in an empirical setting has not been the aim of this chapter. Nor has articulating an engaged approach to working with NGOs in order to understand their situated practices and responses to evaluation demands and configurations. A more comprehensive conceptual and methodological treatment is required to describe such a critical and engaged approach[20]. The following chapters develop this approach and incorporate the 6P sensitivities as an alternative way of seeing and hearing impact power/data/knowledge relations.

---

[20] There are a variety of social practice theories and perspectives, which support critiques of power/data/knowledge or engagements with participants and their lifeworlds. However, there are few that promote both conceptual and methodological resources to attempt to both critique power and engage with partners.

4 Approach and Design: critical engagement

4.1 Introduction

There are six parts to the chapter. Following this introduction, section 4.2 discusses how scholarly critiques are often in conflict with attempts at research engagement, and the need to explore a middle ground between critique and engagement. Section 4.3 describes how cultural historical activity theory (CHAT) provides resources for an engagement component in the study's approach. Section 4.4 describes CHAT resources for conceptually articulating the 6P sensitivities advocated in the previous theory chapter. In section 4.5, the empirical research design for two NGO case explorations is described. Finally, section 4.6 briefly summarises and reconfirms the salience of the critically engaged approach spelled out here.

The methodological frame is designed to explore critical insights in engaged field work interactions with research partners, in the wilds of evaluation practices so to speak, not in scholarly arguments (Law, 2008: 150). The aim is first to understand how NGOs practice evaluations, and then how they respond to, align with, adopt, adapt, resist or reject different critical insights. Critique and engagement are thus combined in this study as a response to their separation in many areas of discourse and practice in which practical engagement, participation, collaboration, or consultancies are deemed of limited critical or theoretical value (e.g. Cooke & Kothari, 2001: 15; Burrell, 1993), or where critical perspectives are dismissed by practitioners as too abstract or unrealistic (Mosse, 2006: 941; 2007: 14; Wallace et al, 2006;

Green, 2018), This allows us to bring critical insights into more engaged, practical spaces for understanding and responding to the impact configurations, omissions, assumptions and inequalities found in earlier chapters. Critical engagement heeds Mohan's (2001: 167) advice on the dangers of passivity and the potential of transformative dialogue in the face of subjugating forces in development. It also heeds the warning that purely interpretive scholarly work can fail to trickle-down into development practice (Harris, 2016).

CHAT is outlined here as a "*alternative research approach*" (Engeström, 2000a: 141) from which to draw resources for collaborative research in development (Kontinen, 2007; Karanasios, 2014). Methodological resources for engagement in the study include CHAT's focus on practitioner's own activities, multi-voiced views, dialogue, offering advice, tools for expansive learning, encouragement of critical reflection, and probing possible changes.

The conceptual framework presented here is based on the 6P sensitivities discussed in the previous chapter. The 6Ps provide an alternative, critical frame for understanding the products, processes, power, political participation, and practices of development NGO impact evaluation. This alternative avoids many of the assumptions, omissions and configurations produced by over-reliance on TIEK and the legacy of DIKW. This includes their linear logics and "methods plus results" obsessions. CHAT offers concepts and vocabularies for articulating the 6P sensitivities. It is one approach to social practice (Engeström, 1987; Blackler, 1995; Miettinen et al, 2013; Nicolini, 2014) and its concepts and tools support an understanding of practice and power (Blackler, 2011; Karanasios, 2014; Simeonova et al, 2018b; Kelly, 2018). CHAT's resources for articulating the 6Ps include activity as a unit of analysis, activity systems, activity networks, use/exchange value contradictions, and the novel notion presented below of temporal activity chains. CHAT is used to operationalise the 6Ps and support critical engagement.

The final section of this chapter describes the research design, research aims, background philosophy, core questions, NGO case design, and case selection criteria for this research. Two small NGO empirical cases are introduced: Rural India, an expert evaluator funded by a well-resourced philanthropic sponsor; and HTSG Thailand, novice evaluators, who compete contract to contract, bid to bid, ad hoc training to ad hoc training, to secure finance and build impact evaluation expertise. Access to the cases, the methods used to generate data, and the analysis of data are discussed in detail. This section ends with an outline of further concerns regarding research limitations, generalisability, ethics, anonymity, and important reflections on the approach and use of CHAT. The research design is critical, engaged, and collaborative,

involving both a consulting perspective to answer requests for evaluation help by the NGOs, but also an exploration perspective, which probes impact power/data/knowledge contexts, configurations and reactions.

4.2   Critique, engagement and a middle ground

How to critique and how to engage with development research subjects are well-established research debates, with many observers suggesting that scholarly critique and practitioner engagement are in conflict. How to critique whilst engaging is less clear. This section outlines some of the tensions between critique and engagement, and makes a case for an explicit middle ground.

There is a history of reflexive work on the dangers of engaging with development-sector actors, characterised clearly in Lewis' (2005) depiction of three kinds of development anthropologists. *Antagonistic observers* critique development ideas, processes and institutions, but make no claims to offer solutions. *Reluctant participants* study cultures and generally avoid engagement with development agencies, although this group took up consultancies in the 1980s when anthropology funding dried up. Finally, *engaged activists* draw on applied anthropology and collaborate with development organisations. Lewis outlines criticisms of the latter group, their involvement in controlling communities, their dismissal as second-rate scholars, and their perceived betrayal of the principle of cultural relativism, when people external to a culture seek to change it. Schonhuth (2002) further argues against external influence from researchers on pure or untouched cultures for ethical reasons, and Ferguson (1994) sets conditions on collaborations, such as engaging only in "*one's own society*" (ibid: 286), or political engagements with social movements, churches, unions or civil society organisations. In addition, Li (2007: 4) warns about the unexpected backlash to well-meaning interventions and the mistaken diagnoses of power deficits as existing in target communities exclusively (ibid: 275). Cook & Kothari (2001) show how proponents of participation have been naive about power (ibid: 14), arguing that critiques of participation are "*a starting point for those who might try to redeem it*" or *"abandon"* it (ibid). They do not specify what follows such critical first steps.

Others view the situation differently. Nauta (2004: 96) argues that although Ferguson is sceptical of engagement, he is not cynical. Henkel & Stirrup (2001: 183) further maintain that increasing knowledge of "out there" poverty is not the priority now, but rather finding critical tools to analyse development's own apparatus. Escobar (2008: 202-3) calls for more attention to agency, connections, contexts, historicity and praxis, the politics of expertise, and what Gibson-Graham (2006: 192) terms the dream of development projects as ethical and political engagements and experiments. One conceptualisation of this dream is Tsing's (2004) "engaged universals", for example human rights, which grew in specific locales, gained attention through advocacy chains, and became globalised knowledge with progressive potentials (Mosse, 2013: 232). Although such progressive knowledge holds value, criticising NGOs in the field as a basis for scholarly debates also frustrates researchers such as Nauta (2004), who argue instead for more care (e.g. Quarles van Ufford & Giri, 2003), and more authentic field contributions.

These views support the idea that critiques must be interrogated in the field, outside scholarly spaces, in order to challenge institutional frontiers (Marcus, 1998: 204, 228). This is because the consumption and effect of critiques are not certain; they can disempower, threaten professionals and their representations, cause upset, or even rage (Mosse, 2006: 941). Mosse advocates citizen engagement and diverse voices in debates on poverty (Mosse, 2011: 8; see St Clair, 2006: 82) and insists that engagement also means that reflection flows two ways, towards participants and researchers, as well as towards their institutional or disciplinary contexts (Mosse, 2013: 227). This point is well illustrated in Harris' (2016: 177-178) article on how ICT4D research fails the poor, how the "*trickle down*" view of research impacting downstream from academic journals to aid policy and practice is "*hard to justify*", and how researchers and practitioners inhabit "*parallel worlds*", in one of which impact is either for academic careers and departments, and in the other, for development policy and practice. Harris acknowledges that researchers aim to influence policy and practice, but their institutional environments discourage them from doing so [21]. If these problems are valid, it means "*a good deal more work remains to find varied constructive modes of engagement with international development, encounters with development policy and its parallel ways of doing knowledge*" (Mosse, 2013: 240). The implication is that critical research must somehow consolidate its

---

[21] Harris also points to former Princeton dean Ann-Marie Slaughter's claim that all disciplines have become "*more specialised, more quantitative, less accessible*"; and that many PhD programs glorify "arcane unintelligibility while disdaining impact and audience" (Harris, 2016; Kristof, 2014).

place in development practice, "*not merely as frustrated post-project critics, but as implementing partners*" (Silltoe, 2002: 1).

The arguments for and against critique or engagement are well examined in Fournier and Grey's (2000) article on critical management. They ask how critique can avoid dilution during engagement with technical managerial discourses, resonating with the question of how development knowledge products can be de-radicalised (Mosse, 2004a) or participation decapitated (Leal, 2007). Fournier & Grey identify two positions for critique: a reformist one advocating critique and engagement (e.g. Alvesson & Willmott, 1996: 18) and two-way exchanges (e.g. Anthony, 1998: 279); and an anti-management one which views the critical role as antagonistic, to "*piss in public*" (Burrell, 1993), with dialogue considered the weapon of the powerful and managers considered a homogenous group. Polemics have a place in scholarly practices, but arguably less so in development policy:

> *"Critical discourses on ICT and development run the risk of having a polemic or moralizing character, of little scholarly value and unconvincing in policy circles."*
> *(Avgerou, 2010: 12)*

Fournier and Grey conclude that critical researchers should explore engagements with diverse groups, for example unions, women's groups or civil society, and not just with corporate managers.

Given these tensions and conflicts between critique and engagement, researching NGOs can still be viewed as a legitimate site of both engagement and critique, in order to learn how NGOs respond to demands for impact data/knowledge, how they are changing, and how they navigate governing processes. Accepting the risks of dilution, this brings into view a fertile middle ground for exploring how critiques might become more pragmatic, and yet retain vitality. Such an exploration suggests there are spaces between and lessons to be learnt from unreflective engagement and esoteric critique (Figure 4.1).
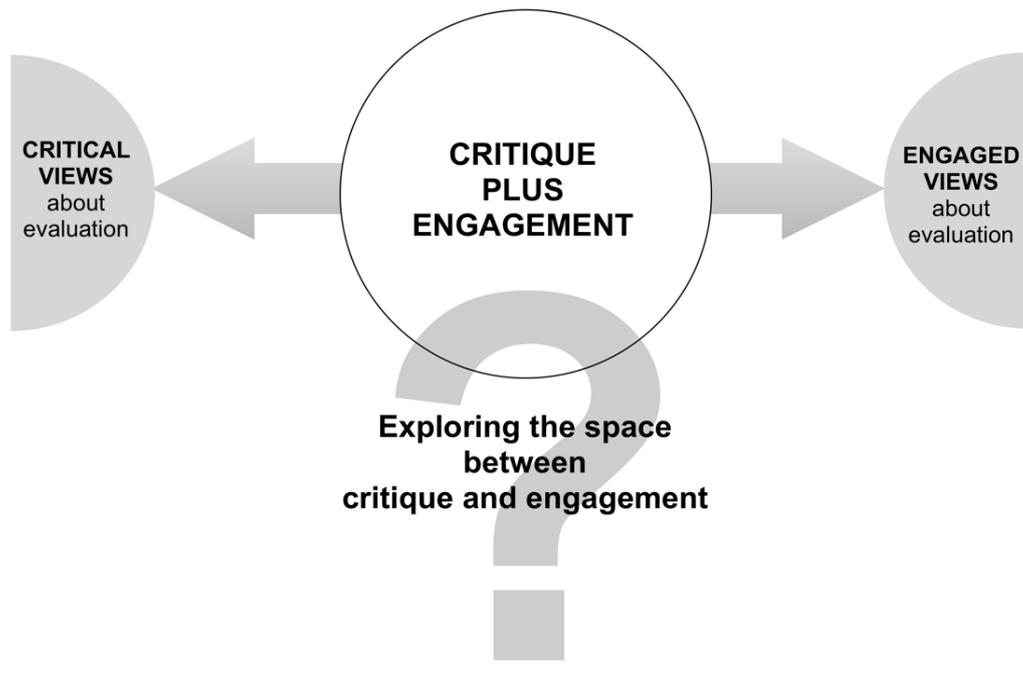
Figure 4.1: Exploring the space between critique and engagement

Amongst many new entrants in the aid sector, from philanthropies and celebrities (Richey & Ponte, 2014) to for-profit consultancies (Hayes & Westrup, 2014), what then becomes of the critic's role? How can critical insights live in the wilds of evaluation practice, prompt reflection, or elevate opportunities for political dialogue, or collaborative evaluation? These questions foreshadow a critical shift from *whether one should engage and critique, to how one does.*

As a first step in reducing the elision of power and practice evident in TIEK, critical engagement needs to emphasise the 6P sensitivities: data/knowledge products, processes, power dynamics, political participation, and practice. Critical perspectives on social practice help us understand these sensitivities in the context of development, because assessments of political-economy at the macro level lead to mixed results regarding specific processes (Hout, 2012), and evidence-based programming does not prioritise context, practice or political engagement (Hickey, 2012).

Critical Development Studies have drawn on the analysis of power and knowledge provided by Foucault (e.g. 1977; 1979; 1980) to interrogate the sector (e.g. Ferguson, 1990; Escobar, 2012; Li, 2007; Lie 2015). Foucault made occasional pleas for practitioners to use and adapt

his concepts or what he once called "*gadgets*" (Foucault, 1980: 65), and for researchers to be "*hyper-active*" (Foucault, 1983: 256). Yet, practical methods of engagement are not easily extracted from his oeuvre.

Many theoretical perspectives do offer resources for understanding social practice (Reckwitz, 2012: Miettinen et al, 2009; Nicolini, 2014). These include for example, structuration theory (Giddens, 1984) and actor network theory or ANT (e.g. Latour, 2005), which have both been used extensively in interpretive studies of information systems (Walsham & Sahay, 2006). ANT has received more attention in contemporary development (Escobar, 2012: xv). However, Escobar also questioned the "*muddled*" agencies and the lack of attention to uneven power relations in network approaches. Mosse (2013: 232) reiterates Escobar's three arguments (2012: xv) that network approaches a) fail to see power as granulated or related to political economy; b) fail to challenge the status quo, risking a "*hermeneutics of cynicism*" about agency; and c) deny universal "*western*" notions of development or modernity, thus constraining counter-arguments. This means some network approaches to practice may lack resources for collaborative agency.

Action research has been suggested as one way of promoting research engagement (Walsham & Sahay, 2006: 19), with participatory design as another. Participatory approaches promote engagement with communities in order to understand local factors such as local needs, problem-solution frames, and knowledge. Local action is a priority in participatory development (e.g. Chambers, 1994), participatory evaluation (e.g. Cameron 2006; Holland, 2013), participatory action research (e.g. Fals-Border & Rahman, 1991), and information or knowledge management using "communities of knowing" (Boland & Tenkasi, 1995), "communities of practice" (Lave & Wenger 1991) or "community informatics" (Gurstein, 2000). Participatory design approaches (e.g. design thinking or service design) also stress local participation, experimentation and even empathy as part of innovation toolkits to support economic and social development (e.g. IDEO, 2014; Edmunds & Cook, 2014). These approaches provide pragmatic concepts and techniques for engagement, knowledge sharing, tool use and reflection on local activities.

However, action research and participatory design have attracted criticism for not dealing with development in practice, power or politics (e.g. Nichols, 1999; Mama, 2000). Communities of practice are limited in terms of dealing with power (Plaskoff, 2011: 199-203; Contu & Willmott, 2003) and there is concern regarding their adoption within managerialist discourses

(Hughes et al, 2007; Plaskoff, 2011: 200), an issue which resonates with concerns about faux-participation (Leal, 2007; Cook & Kothari, 2001; Contu & Girei, 2014). Cooke (2005) has identified the roots of action research in US colonial encounters with native Americans, where resistance was pacified using superficial participation (Collier, 1945), and Hayes & Westrup (2014) have further shown how participation can be professionally and strategically deployed to render local sites controllable in a form of technical anti-politics opposed to multi-voiced participation. As such, many participatory claims are weakened by seeing power and participation as encapsulated in the field (i.e. sites, communities or workshops), rather than diffused across the aid chain, in networked sites, products, processes and practices.

In selecting a framework for critical engagement and sensitivity to the 6Ps, this raises two issues: firstly, participatory design and action research do not privilege networked power relations; and secondly, critical theories do not foreground methods for collaborative engagement. What is required is a framework that addresses the space between critique and engagement, and tactics to combine them as critique plus engagement.

4.3   CHAT as a framework for critical engagement

4.3.1 Introduction to CHAT

To understand and respond to the problem of critique versus engagement, the study approach must therefore address the 6P sensitivities, test critiques of evaluation power and practice in the field, and engage with NGOs performing evaluation practices.

The framework adopted here to operationalise both the 6Ps and critical engagement is cultural historical activity theory, or "CHAT" (Engeström, 1987; 2009; Blackler, 1995; 2011). CHAT draws on Vygotsky's (1978) social constructivist view of developing practice. It rejects positivist or objectivist ontologies that separate the "knower" from the world known, and views organisations as historically and culturally formed relations (Engeström's, 1987). Knowing and acting in CHAT are not considered as individual cognition or behaviour (Engeström, 2000b; Packer & Goicoechea, 2000), but are part of wider cultural and historical processes. This commitment to practice is rooted in CHAT's history and the work of Vygotsky (1978) and

Leontyev (1978). Leontyev sought to build an approach to psychology which utilised Marx's rejection of Hegelian idealism, and favoured a focus on material practice (e.g. Marx, 1845/1969), emphasising the importance of language, signs and group collaboration. CHAT encompasses a broad school of thought applied to domains such as Information Systems, Organisation Studies, and Education. A small literature is now emerging in Development Studies and ICT4D (see Karanasios, 2014).

This section describes CHAT's conceptual perspectives on engagement, on practice as networked activities and contradictions, and on power as the incremental mediation of data/knowledge in temporal activity sequences. Using CHAT to investigate power, knowledge and data relations follows calls from numerous authors, in Knowledge Management (Blackler & Regan, 2009; Blackler, 1995; 2011); in organisation studies (Engeström, 2006; Jarzabkowski, 2003; Canary & McPhee, 2009; Kelly & Cowen, 2014; Kelly, 2016); and in IS/ICT4D (Korpela et al, 2004: 453; Karanasios & Allen, 2013; Karanasios, 2014). Blackler (2009) and Engeström (2011) themselves use CHAT to critique power structures in collaborations with research stakeholders. As indicated by this work, CHAT has resources for both engagement with research participants and critiques of networked power and practice. It should be noted that CHAT offers a diverse range of concepts and tools for analysis and engagement, although the study does not deploy all of these. The reasons for using specific concepts, but not others are clarified through the chapter and reflected upon in section 4.4.7.

4.3.2 CHAT and engagement

There are a range of concepts and principles in CHAT that support engagement and collaboration with and between research participants, through dialogue, discussion, experimentation and learning (Engeström, 1987). Although less suited to large-scale studies or national-level enquiries (Karanasios, 2014: 11), CHAT is rather well-suited to analysis of collaborations with research partners that focus on their practices (e.g. Virkkunen & Shelley Newnham, 2013), for example in hospitals (Engeström, 2001) or emergency services (Allen et al, 2014). CHAT studies support dialogue, allowing participants to discuss problems or conflicts within work processes, to surface systemic contradictions in their work, to expand their views of activities shared with other participants, and to co-design experimental changes intended to resolve such conflicts and contradictions. As such, an affinity exists between CHAT

and other research approaches that foreground multi-voiced participation, such as participatory design or action research [22]. Engeström elaborates, an activity system:

> *"… is by definition a multi-voiced formation. An expansive cycle is a re-orchestration of those voices, of the different viewpoints and approaches of the various participants." (Engeström, 1991: 14-15; 1999: 8)*

Different CHAT facilitation methods promote collaborative enquiry, such as the "Change Laboratory" method (Engeström et al, 1996; Engeström et al, 2001; Virkkunen & Newnham, 2013) and the "Clinic of Activity" (Clot, 2009; Clot & Kostulski, 2011). CHAT has also been used in other research settings each with diverse methods for generating data, from human computer interaction studies (Nardi, 1996) and participatory ethnography (e.g. Kontinen, 2007), to collaborative design (Zahedi & Tessier, 2018), emancipatory initiatives (Karanasios, 2014), and case studies with farmers (Mukute & Lotz-Sisitka, 2012). In this sense, CHAT offers a flexible framework for engagement and analysis. It has been described as "*more operational*" than actor network theory or Giddens' (1984) structuration sociology, and more "*theoretically founded*" than participatory design (Korpela et al, 2004: 455).

In both NGO cases, research participants were separated across cities, countries and continents, and roles, from the CEO of a philanthropy in London, to NGO directors and rural beneficiaries in India and Thailand respectively. Thus, a typical change lab with a series of face to face meetings for diverse participants, and review of video evidence in group discussions for example, was not feasible. This separation, characterised by the geographic separation and socio-economic diversity of aid funders, NGO managers, specialists, evaluators, and aid beneficiaries was crucially influential upon the engagement process and research design, with CHAT's view of overlapping activities and the value of dialogue between participants requiring support through activities on the part of the author, including: testing critiques with research subjects, suggesting wider evaluation frames and voices, offering consulting advice, and conducting workshops. The engagements therefore involved an active researcher, not a passive observer. As such, in the research design below, further tools are specified which helped the author test critical insights and alternative evaluation options to better understand

---

[22] See Somekh and Nissen's (2011) introduction to Mind, Culture and Activity's special edition on CHAT and action research for more on this frequently debated relationship.

power/data/knowledge flows and relations with the research participants. In effect, the author tried to probe NGO evaluation practices (Gaver et al, 2006), or stimulate alternative ways of doing evaluation (Salancik, 1978). The aim was to encourage reflection and dialogue, and shed light on evaluation preferences and priorities; thus illuminating too issues of silent or unobtrusive power and practice (Blackler, 2011) in evaluation processes.

In typical CHAT studies, engagement and organisational changes are viewed as transformative processes, which are analysed using CHAT's expansive learning cycle. The cycle consists of stages including questioning and analysing current work activities, modelling and examining new tools and new activities, implementing new solutions, and finally evaluating and reflecting on changes made (Engeström, 2001: 152; Virkkunen & Newnham, 2014; Mukute & Lotz-Sisitka, 2012: 345). The expansive learning cycle (Figure 4.2) is typically used to scaffold interventions, deal with contradictions, formulate new ways of working and promote transformative work practices.
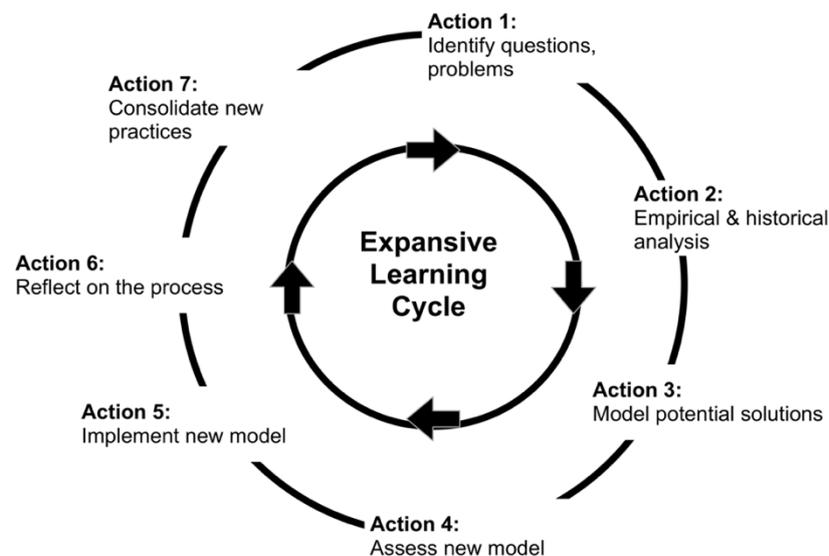
Figure 4.2: CHAT expansive learning cycle [23]

The expansive learning cycle is a significant aspect of CHAT in many studies and bears some affinities with action research cycles which focus on cycles of problem identification, planning, action, and renewed reflection. (e.g. Virkkunen & Shelly Newnham, 2013: 3-10; Lewin, 1943,

---

[23] Figure 4.2 is an adaptation, drawing from Engeström's original (2001: 152) and Kontinen's version (2013: 113).

1946, 1947a; Argyris, 1985; Whyte, 1991; Reason & Bradbury, 2001). Cycles of expansive learning in CHAT focus on empirical engagement with research participants as the sources of knowledge and activity. In this sense, the expansive learning processes used in formal CHAT interventions is not like a standard learning process which focuses on acquisition of skills or knowledge that can be identified in advance, for example in a curriculum, or on individual mastery and behaviour change (Engeström, 2001). Instead, expansive learning involves situated learning, change and negotiation through collaborative work and dialogue.

> *"People and organizations are all the time learning something that is not stable,*
> *not even defined or understood ahead of time." (Engeström, 2001: 137-138).*

In formal CHAT interventions, researchers conduct a series of group sessions with work units and diverse stakeholders, exemplified in Engeström's (2001) study of physicians, nurses, staff, managers, primary healthcare centres, and hospitals. These groups met 10 times for three-hour meetings over sessions ending in February 1998 (ibid: 139). One of the key elements in such CHAT studies is to surface multiple voices and perspectives on work activities, and generate dialogue to explore problems, conflicts, dilemmas and opportunities for new ways of work.

Mukute & Lotz-Sisitka (2012) show how CHAT can be used to generate diverse kinds of data in engaged development-sector research. Different data types emerge from interviews, insider experiences, gatekeeper interactions, document analysis, visits to research sites, trainings, workshops, ongoing collaborations with participants, and the use of audio or video transcripts. Kontinen's (2007) ethnographic field work also featured both observation and active participation, in meetings and workshops, specifically with development NGOs. These studies reconfirm CHAT's resources for engaging with research participants, stimulating dialogue, and probing existing practices and opportunities for change.

As a result of such resources and concepts, CHAT not only offers critical views of power and knowledge flows, but also supports participatory engagement with research partners beyond interpretive studies alone. However, CHAT research is not always successful in promoting dialogue, learning or change amongst participants. In Engeström's work with a TV production crew (2008: 22-47; also Engeström & Mazzocco, 1995), problems, conflicts, dilemmas and contradictions were "*masked*" or silent, unclear and not acted upon by participants. There were many reasons for this, but Engeström concludes that a lack of communication across activity systems, i.e. TV show production, management, and marketing activities contributed to the

masking of problems. Effectively, problems were silenced and normalised. Avis (2004) also critiques some CHAT work for promotion of corporate transformative narratives.

In terms of the engagement process in this study, the researcher used technical impact evaluation knowledge (TIEK), CHAT concepts, and critical development insights to both respond to participant NGO requests for technical advice, and to probe their activities in order to learn more about evaluation activities and constraints. The researcher, at the request of the NGOs, was thus an active participant, although a typical CHAT change lab was not possible. No formal intervention schedule with regular sessions comprising diverse stakeholder groups were possible, due to the global dispersion of said stakeholders. Further details of the advice, suggestions and techniques used are provided in the research design, sections 4.4.3 and 4.4.4 below. This was an important component of the study and the particular critical use of CHAT, which deviates from formal change laboratory processes. It involved careful testing of critical observations in the wilds of NGO evaluation practices, without involvement of all key stakeholders from different countries and socio-economic strata, but drawing on critical literature insights and how they may or may not add value to volatile development contexts (Picciotto, 2013) where NGOs are normally sensitive to outsiders (Wallace et al, 2006; Shivji, 2006). Critical engagement thus required the offer of specific and local consultancy support to NGOs on evaluation knowledge to secure research access, but also critical questions, suggestions and provocations to understand NGO needs and power/data/knowledge relations within wider development chains and networks.

This section has detailed how CHAT informed the study's engagement with research participants in order to make clear NGO participant views, opportunities for learning, and transformation, the supply of consultancy style evaluation advice, and the probing of critical issues during research interactions. This constitutes the methodological element of the critical engagement approach. The next section looks at how CHAT was used to understand practice and power, and how it operationalised the 6P sensitivities.

4.3.3 CHAT and practice: activities and contradictions

This section describes in detail CHAT's view of practice as "activities". This articulates how practice and power are conceptualised in the study, and thus how the 6P sensitivities are

operationalised using particular CHAT concepts. Activity systems, activity networks and associated contradictions are the core CHAT components in this study.

In CHAT, in-situ practice is traditionally understood as activities (Nicolini, 2012; Miettinen et al, 2009: 1317). Activities are analysed through activity systems and activity networks (Engeström, 1987; Karanasios & Allen, 2013: 300), which are changing over time due to historically and culturally facto. Activities involve subjects (e.g. people, groups), tools (concepts, technologies), and objects (purposes, goals) of activities (Engeström, 1987; Blackler & Regan, 2009). One account of CHAT's evolution posits an expansion of the notion of activity over three generations. The first generation drew on Vygotsky's (1978) concerns with how an individual subject's consciousness develops through tools and concepts used in everyday learning activities. The second generation grew through Leontyev's (1978) emphasis on collaborative group activity. And the third generation is associated with Engeström's (1987) discussion of organisations as activity systems and networks.

Vygotsky (1978) explored the social and interactional basis of individual child language development within a Marxist psychology, radical at the time, which attempted to unite material and mental perspectives in a non-dualistic approach. Vygotsky understood meaning as formed through subjects, objects (or objectives), and mediating tools e.g. linguistic concepts or material artefacts (Figure 4.3). In this view, the individual learner, their goals, and available physical or knowledge resources contributes to outcomes, learning, change and wider culture.
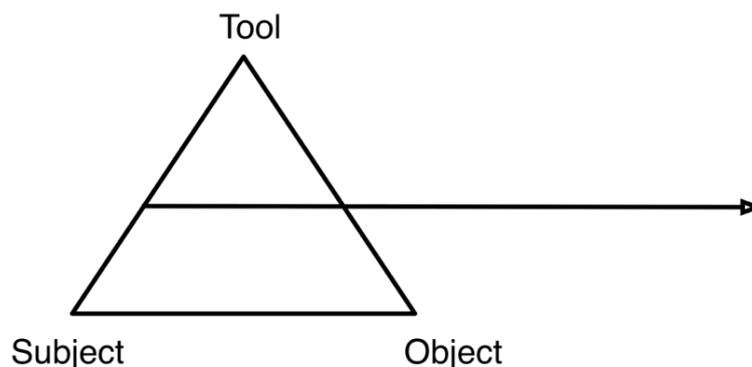
Figure 4.3: Vygotsky's model of mediated activity

Also important is Vygotsky's view of activity as the methodological unit of analysis, which contrasts with the positivist "*accepted scientific paradigm*" according to Holzman (2006: 112a). In positivist thinking, method is understood as a tool separate to and yielding of representational results. In contrast, Vygotsky saw methods and results as elements within developing activities. In this view, evaluation methods are considered elements within evaluation activities, "*simultaneously prerequisite and product, the tool and the result*" (Vygotsky, 1978: 65; Holzman, 2006b). In short, methods and results are parts of ongoing practice, not external representations of it.

Vygotsky's student and colleague Leontyev emphasised how group activity and interaction contributed to individual psychological development. A canonical example in Leontyev's work is the hunting expedition as an illustration of the division of labour. For Leontyev, when primitive humans hunted game for food, individuals would take on different roles such as maintaining a fire, cooking, pursuing the game, or waiting to ambush it (Leontyev, 2009: 186-187). This role specialisation contributed to diverse psychological development across populations and diverse possession of specialised knowledge. For Leontyev, this knowledge was relational and intricately tied to ongoing activity: *"Knowledge of a thing is only possible in its relation to other things, in reciprocal action with them, in motion" (Leontyev, 2009: 25)*

In today's data/knowledge intensive development sector, Leontyev's insight into *knowledge in motion*, is important for understanding the many locales of knowledge labour required to construct and circulate impact data, claims or narratives. Combining Vygotsky and Leontyev to look at aid sector evaluations implicates individuals in multiple groups that are constructing impact data, knowledge, methods, and results across dispersed organisational activities. Engeström's work explicitly moves to this larger organisational scale.

To analyse organisational activities, Engeström (1987: 78) developed an activity system model for incorporating rules, social norms, communities and the divisions of labour that shape activities (Figure 4.4). Activity systems take shape and are developed over "*lengthy periods of time*" (Engeström 2001: 136). They are best analysed in terms of their historical development; that is, the *"history of the theoretical ideas and tools that have shaped the activity"* (ibid).
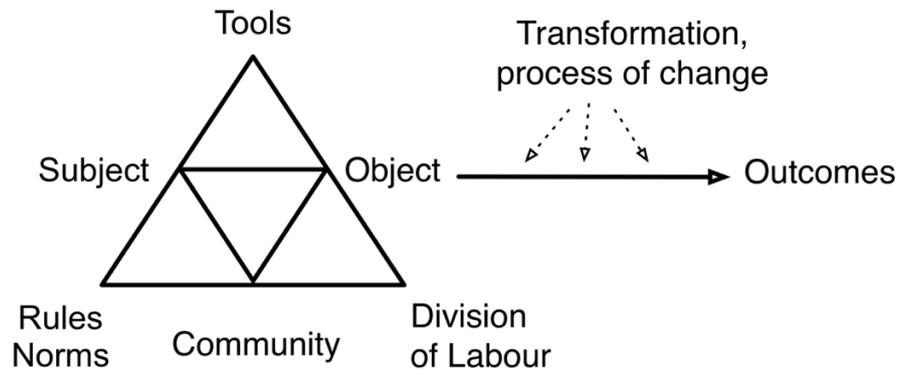
Figure 4.4: Activity systems including social relations (Engeström: 1987: 78)

CHAT acknowledges multiple, intersecting or interwoven activities in networks (Engeström, 1987: 103; Karanasios & Allen, 2013: 300), for example in Knowledge Management (Blackler, 1995; Worthen, 2008) or Development Studies (Karanasios, 2014; Kontinen, 2007) (Figure 4.5).
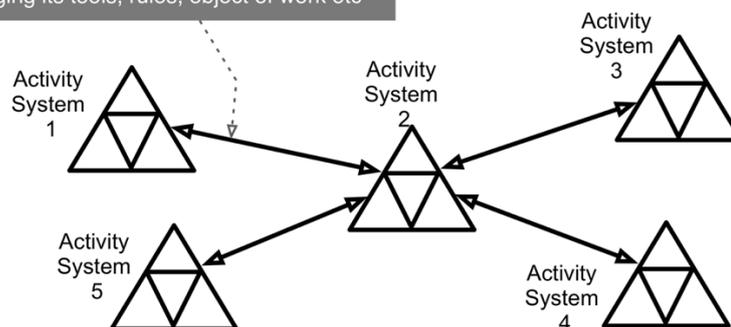
Figure 4.5: Multiple activity systems in networks of relations

Fundamentally important for this study is the ability to analyse how one evaluation activity produces knowledge or data that is then used in follow-on activities, or how knowledge is "*in motion*" in Leontyev's terms. Worthen (2008) discussed similar parallel processes of knowledge production in everyday work, and this motion is well-illustrated by Miettinen et al (2012) in relation to the design of physical products: "*a product designed & fabricated becomes a tool for use in another activity*" (Miettinen et al, 2012: 11).

Conceptual products are considered outcomes of activities, but during their production - within and between activity systems - tensions and contradictions arise. These manifest to researchers as localised "*problems, ruptures, break downs, and clashes*" (Kuutti, 1996: 37), and can arise as symptoms of deeper activity use/exchange value contradictions. Engeström (1987) describes four kinds of contradiction within activities and between activities (Table 4.1).

Contradiction	Description
Primary (1)	Within single activity system elements, e.g. within a rule or a tool
Secondary (2)	Between activity elements, e.g. between a tool and the community
Tertiary (3)	Between an activity and its earlier formation, e.g. old/new farming methods
Quaternary (4)	Between activity systems in a network, e.g. marketing and production

Table 4.1: Description of kinds of activity system contradictions

As we have seen in section 4.3.2, CHAT interventions support dialogue and collaboration with participants in organisations to resolve work contradictions and explore opportunities for expansive learning and new ways of working. Kontinen (2007) has identified six such contradictions directly relevant to and likely to be experienced in development and NGO work. These contradictions generate conflicts and dilemmas for practitioners, that become embedded into "*historically formed institutions, discourses and ideas*" in development (ibid: 9-10). Kontinen advises researchers to be aware of these contradictions, but to also be open to novelty and unexpected findings (ibid: 13). The contradictions are listed below and inform the case analysis and discussion in later chapters. Development work contradictions: development

discourse versus development practice; rationalised input/output project blueprints versus messy, complex social processes; equal versus unequal ownership of initiatives; external/distant control versus immanent control by local communities; technical and professional development logics emphasising solution management versus moral development logics emphasising social justice and political change; and, critical analysis as "*a commitment to improving practices*" (ibid: 12) versus research that is critical of development itself.

Activities are traditionally understood in activity theory as part of institutional work. Historically, such institutions have dominated activities, producing stable activity systems such schools or clinics (Hedegaard et al, 1999):

> "*Different institutions - home, daycare, school - were dominated by different activities; the dominating activity in an institution acquired the role as leading activities in different periods in a person's life". (Hedegaard et al, 1999: 14).*

However, views of activities and their boundaries have expanded in CHAT. Hedegaard et al. acknowledged Elkonin's (1972) role in diversifying the view of traditional and stable institutional activities. Engeström's (1987) early work on activity systems and networks added to this expansion. In his more recent work on the fluidity of contemporary activities (Engeström, 2009) and how they can be conceptualised as organic patterns (e.g. mycorrhizae process trails) or runaway collaborative objects that go viral, Engeström shows this to mean that activities do not necessarily have a centre or a controlling core. They are globally diffused, as in the case of Wikipedia. In a similar way, knowledge construction today is less likely to be purely local or occurring in small groups and bounded activities. NGO impact evaluation is subject to what Blackler & Regan call (2009: 164) "*distributed agencies*", where methods, demands, subjects, data, and results are globally distributed. Describing distributed activities across systems, networks or chains supports a more nuanced understanding of power, data and knowledge relations relevant to the 6P sensitivities as: "*Human activity does not exist except in the form of action or a chain of actions (Leontyev, 1978: 64; Engeström 2000a: 307)*

In summary, a key aspect of the approach builds on how CHAT has over time broadened its view of how situated practices are formed, sustained, stabilised and changed, from the level of individuals, to groups, systems, networks and globally distributed activities. In the next section, the novel notion of CHAT-based temporal sequences of distributed activities is introduced as a key analytic for understanding activity sequences and power, data, knowledge relations.

4.3.4 CHAT and power: temporal chains, submerging and elevating

Building on CHAT's activity systems and network concepts, in order to articulate power relations and the 6P sensitivities, the study must demonstrate how power/data/knowledge silences, configurations or omissions are generated during evaluation practices. How do rationalised knowledges subjugate local or other knowledges in practice (Avgerou, 2002)? How do governing processes or TIEK prescriptions marginalise other ways of knowing impact? This section describes temporal activity chains and processes of editing along the chains, by which aspects of data/knowledge relations become submerged or elevated over time.

In ICT4D, Hayes and Westrup (2014) show, for example, how the qualities of efficiency and effectiveness are elevated in technical evaluation methods designed by consultancy organisations, who strategically paint local NGOs or competitors as incompetent, and thus submerge their voices. In a similar vein, Mukute & Lotz-Sisitka (2012: 361-364) examine governing controls and local responses using CHAT and critical realism. They show how South African farmers learn sustainable agriculture, but simultaneously resist governing influences including: government farming policies; commercial promotion of unsustainable agricultural technologies; divisive gender impositions (male land ownership, female labour supply); and university/education sector promotion of "modern" or "scientific" agricultural knowledge over farmers' own knowledge (ibid: 358). They use CHAT to account for what "*research participants ... say and do*" (ibid: 364), and critical realism to critique subjugating structures.

Such subjugating influences can be silent. Blackler (2011) depicts organisational power relations in four quadrants, across two dimensions. The first dimension is from personal to collective, and the second from overt to unobtrusive power. Power relations in the current study lie in Blackler's unobtrusive dimension, in terms of how TIEK shapes NGO evaluation activities and how evaluation processes become normalised in everyday work as part of what people "*think of as ordinary*" (Blackler, 2011: 732).

If subjugation is normalised and silent, then how can it be critically approached in the study? Normalised power relations embedded in TIEK and DIKW assumptions and omissions mean the process of knowing impact becomes less contestable and less subject to dialogue or political negotiation. Critical attention is therefore required as problems may not surface in open dialogue between participants. Kontinen's six contradictions may not be explicit in daily

practice. For this reason, Blackler (2011: 733) and Mukute & Lotz-Sisitka (2012) recommend an awareness of power and politics beyond local client or local partner activities. In this way, the critical approach uses CHAT to question silent, unobtrusive, normalised TIEK and implicit DIKW influences, and elevate the 6P sensitivities.

It is worth revisiting Engeström's (2008: 22-47) work with a TV production crew to elaborate on this aspect of critical engagement. In his study, there were few conflicts or disturbances. This could have been a sign of skilled performance (Dreyfus & Dreyfus, 1986) by the TV crew. Well-managed problems or disturbances can lead to innovation and collaboration, or badly-managed disturbances can prolong stagnation, and conceal deeper systemic contradictions (Perrow, 1984). On analysis though, Engeström suggests disturbances, misunderstandings, dilemmas, conflicts and problems in the case were not allowed to surface; they became "masked". Management withheld information about program production revenue in order to cut costs and potentially terminate the show. Marketing failed to pitch the program accurately, and the TV crew were defending an old way of making the show, rather than experimenting with new ideas and technologies. The masking of disturbances ignored systemic contradictions, constrained learning, collaboration and innovation, and silenced power relations. The case is insightful because diverse activity systems and agencies (management, marketing and production) were all involved in masking the conflicts and disturbances.

Such networked mediations are common in CHAT, where outcomes produced in one place and time become are changed, mediated, used and adapted in another place (Miettinen et al, 2011). This means, for the current study, that specific impact data/knowledge is developed and transformed in a flow or motion (Leontyev, 2009) through time-sequential activities. In much Knowledge Management literature, what becomes focal or peripheral depends on the need for innovation and effectiveness (e.g. Nonaka & Takuechi, 1995; Tuomi, 1999). In TIEK, what becomes focal depends on applying an appropriate method to deliver results that will elevate project efficiency and effectiveness and produce robust scientific narratives of cause and effect. However, in CHAT, these goals, methods and results are viewed as elements of activities (Holzman, 2006b: 112). In a distributed way, beyond the NGO's themselves in the study, TIEK goals, methods and prescriptions mask or submerge alternative ways of knowing impact.

Having isolated TIEK methods and techniques as specific elements of activity, one can identify how impact data is mediated; as Tuomi (1999: 111) suggests, how it is decontextualised in one activity and decontextualized in subsequent activities. What is left in as impact data/knowledge

is repeatedly transformed in temporal sequences of activities. But what is left out of this temporal chain? CHAT can register such mediations as locally incremental changes that transform the data, the knowledge, the activities themselves, and the power relations between agents. Therefore, by paying attention to what is edited in or out, that is, what is "submerged" or "elevated" along the temporal chain of impact data/knowledge construction (Figure 4.6), one can describe the genesis of impact end-products and power relations occurring over time.
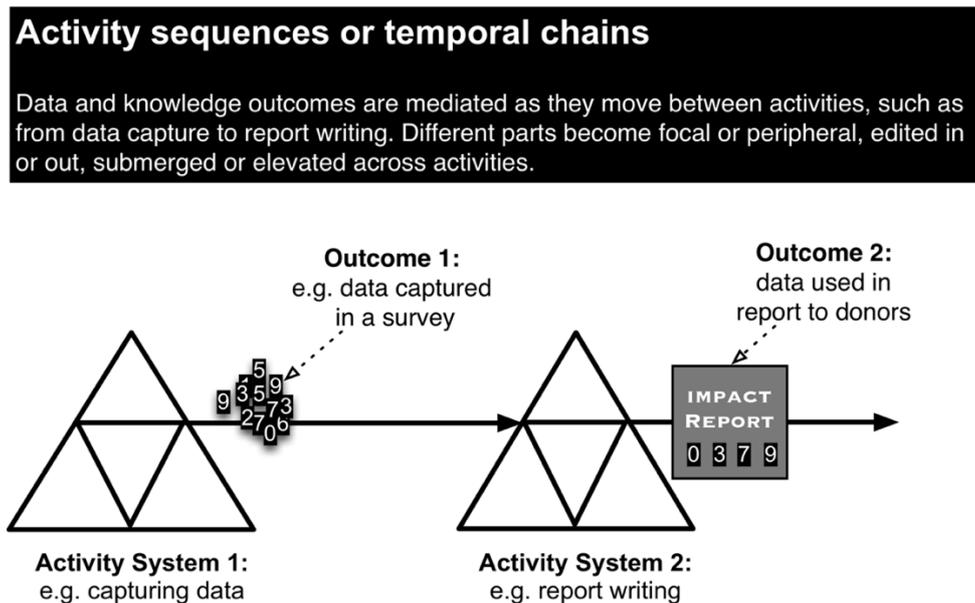
Figure 4.6: Data is mediated between capturing and reporting activities

4.3.5 CHAT's role in the study

The CHAT analysis of the case requires two steps in order to understand how data and knowledge may be subjugated or rationalised (Avgerou, 2002). These concern how micro-encounters between prescriptive evaluation methods and data/knowledge models are enacted in situated processes of knowing (Blackler, 1995), and how this generates power relations between stakeholders (e.g. Hayes & Westrup, 2014).

- Step 1: Identify the activities involved in impact data/knowledge construction; and
- Step 2: Identify how power relations are generated during the construction process, which elements are elevated or submerged incrementally

These steps follow the study's research questions, in firstly identifying activities and secondly identifying power relations in impact data/knowledge construction processes. The third research question, exploring how NGOs reflect and respond to critical insights and views about their impact evaluation work, involves the engagement outlined in section 4.3.2 above. CHATs resources for critical engagement and analysis of the 6Ps, constitute an approach to learning about impact evaluation power, data and knowledge relations with development NGOs, although not a process of engaged scholarship guaranteeing progressive transformations.

4.4   Research Design

This section describes the research design. Section 4.4.1 clarifies the study aims and briefly re-affirms the research philosophy. Section 4.4.2 describes the case design and the criteria for case selection. Section 4.4.3 describes each NGO case in terms of their contextual background, researcher access to participants, methods for generating data, and lists the data generated. Section 4.4.4 explains how the data was analysed using CHAT and the 6Ps, and section 4.4.5 reviews design limitations, ethics process, security, confidentiality issues. Finally, section 4.4.7 reflections on the use of CHAT within the critical engagement approach.

4.4.1 Research aims and the CHAT perspective

The research aims to interrogate the socio-political relationships between power, data and knowledge in development impact evaluation activities. This entails a research design that is critical, qualitative, and practice-oriented. Although CHAT focuses on activity as the unit of analysis, rather than individuals or macro social structure, it is still difficult to make definitive statements about research paradigms or philosophies because disputes between schools of thought (e.g. positivists, phenomenologists, critical realists or social constructivists) are "*inextricably interlinked and mutually constituting*" (Spender & Scherer, 2011: 15).

With this caveat in mind, CHAT's philosophical perspective on investigating the activities of development evaluation involves a historical and cultural view of research design ontology and epistemology. This is best conveyed through a short story. In the story, three philosophical positions are represented when three baseball umpires call foul balls (Engeström, 2000b: 301).

The positivist / objectivist umpire says: "*I whistle the ball foul when it is a foul ball*". The subjectivist / interpretivist umpire says: "*I whistle the foul ball when it seems to me that it is a foul ball*". An, the social constructivist umpire says: "*The ball is foul when I whistle it a foul ball*". Engeström explains that CHAT does not adopt the first objectivist position, nor the second interpretivist position, nor the final umpire's constructionist position. Instead, Engeström argues that each of the umpires is committed methodologically to individualist views of knowledge and reality. CHAT would require a fourth umpire to say: "*Given the current rules of the game, lines, ball, bats, and roles played out collectively (umpires, batters, pitchers, spectators etc), it is my role to call a ball foul now*". This interpretation foregrounds the historical development of baseball and the collaborative construction of the rules, tools, roles etc. It does not isolate the umpire as the sole arbiter of the rule. Rules and roles develop in collectively and historically formed activities.

This story illustrates CHAT's perspective on reality as historically developing activities, and its epistemological view of knowledge as constructed and negotiated by diverse stakeholders. This philosophy scaffolds CHAT's commitment to multi-voiced views of reality, knowledge and power relations and makes it valuable for a critical engagement.

4.4.2 Research questions, NGO cases, and case selection criteria

Initial research questions developed between August and October 2013, as the first case study began, focused on what factors shaped NGO impact evaluation in practice. However, as the first case study progressed, the specific relationships between data, knowledge and power became increasingly important. This theme, the literature review and theoretical development all helped mature the enquiry, leading to central three research questions.

RQ1: How is impact evaluation data/knowledge constructed at small development NGOs, in practice?

RQ2: Are / how are power relations generated during impact evaluation data/knowledge construction at small development NGOs?

RQ3: Can / how can power relations be addressed and responded to in impact evaluation data/knowledge construction practices at small development NGOs?

A case approach was adopted because it supported a critical analysis of qualitative data on the silencing of power relations in NGO impact construction processes. Case collaborations are common in CHAT studies of organisations (e.g. Mukute & Lotz-Sisitka, 2012; Engeström, 2008; Karanasios & Allen, 2013). Mukute & Lotz-Sisitka draw on Yin's (2003) observations, firstly that case studies are appropriate when time and space configurations cannot be manipulated, and secondly, that multiple case studies support comparative and contrastive analysis (Yin, 2009). Also, case studies are valuable for influencing practices at the time, and policy or research over the long term (Merriam, 2001; Mukute & Lotz-Sisitka, 2012: 347).

Critical engagement, CHAT resources, and the case design all feature researcher participation, observation, and collaboration, in this case when the researcher offered technical or critical advice about impact evaluation in emails, conference calls or workshops. Offering support enabled trust and access, because NGOs are hard to research (Wallace et al, 2006), prefer to prioritise their day-to-day work rather than granting access to researchers (Lewis & Kanji, 2009: 3), and are sensitive to scrutiny from the media, governments, and academia (Roche, 1999:1-2). Access was fragile and required the researcher to balance critique and consulting.

Regarding case selection, an opportunity arose in August 2013 to work with Rural India and their philanthropic funding partner, the Imagine Foundation, on their upcoming 2013-14 impact evaluation. More details about the engagement process and access are described in section 4.4.3; however, in terms of case selection, this early opportunity initiated the empirical engagements. Rural India was important for three reasons. Firstly, it provided access to senior management at the NGO and at the donor organisation. This was a rare opportunity to research the micro-interactions of key management-level agents, pivotal in designing, shaping and managing evaluations. Secondly, the fact that the funder was a private philanthropy using evaluation techniques from private sector modes of evaluation meant the case offered insight into contemporary evaluation design. Thirdly, as the case progressed it became clear that there were not only shaping pressures upon Rural India's evaluation activities, but power relations were being enacted through everyday, in-situ impact data/knowledge work. These were perceived as unproblematic parts of normal professional practice by the managers. This insight helped mature the research questions and guided the selection of a second NGO case when the Rural India collaboration began to wane in March 2014.

At this time, the author contacted forty UK and internationally based NGOs, drawing on personal contacts in the sector and email enquiries to identify an NGO interested in an impact

evaluation collaboration. Seven organisations agreed to collaborate, leading to research proposals, field trips and collaborations in the UK and Asia. Each collaboration had a focal point in relation to impact evaluation. One project focused on how evaluation was constructed in funding bids for a UK NGO working with street children in Nigeria. Another focused on how evaluation was used to align a consortium of NGOs and research organisations working on environmental sustainability in southeast Asia. Another explored how evaluation of impacts was embedded into daily judgements and decisions between designers and clients as part of work with an NGO in Myanmar bringing design thinking approaches into development spaces.

However, one NGO stood out as an interesting contrast to Rural India. This was the Hill Tribe Support Group in Thailand (HTSG). Where Rural India had displayed a level of evaluation expertise, experience and an instrumental attitude to producing impact data and knowledge, HTSG were novices, lacking capacity, technologies, experience or awareness of the process of constructing impact data and knowledge. The similarities and differences between the two cases provided a rich vein for critically analysing development NGO impact evaluation, how it is done successfully or learned and rehearsed. As a result, the research utilised a contrastive case analysis of two NGOs, one an expert and one a novice evaluator, to better understand power, data and knowledge dynamics in development NGO impact evaluation.

Both NGO cases also exhibited power/data/knowledge relations that were not manifest in overt conflicts, but were mundane, normalised and relatively "*unobtrusive*" (Blackler, 2011). These silent power relations were unlike "*critical events*" (Long, 2004: 60-61), such as the Union Carbide Bhopal chemical disaster in India in 1984, which represent a loud rupture that illustrates development power dynamics. Problems were not part of open conflicts or ruptures (Kuutti, 1996: 37). They appeared instead silent or masked (Engeström, 2008). In both cases, impact construction generated power inequalities quietly and incrementally, making for an interesting contrastive study that could shed light on configurations of evaluation NGOs, and the mundane governing processes[24].

---

[24] On reflection, in both cases, the empirical data afforded an analysis of power, data and knowledge which contributed to scholarly deliverables. However, NGO field success was less clear, even though the "tales from the field" (Van Maanan, 2011) supported scholarly knowledge production. This mixed success help to dispel assertions that CHAT studies overemphasise progressive change and transformative success (Avis, 2007).

4.4.3 NGO Case 1: context, access, timeline, methods and data

This section introduces the research context, how access evolved, a case timeline, methods used to generate data, and details of the data used in the study. The first NGO case features interactions between the small NGO based in central India, "Rural India", and their philanthropic funder based in the UK, "The Imagine Foundation" or "Imagine UK" for short. Rural India has been working with female farmer groups on agriculture and livelihoods projects since 2006. Imagine UK awards grants for "*efficient and impact-driven models of development*", actively supporting NGOs to help them achieve "*measurable results*" (Imagine UK, 2014). Their partnership has lasted over 10 years. More detail about the partners is included at the start of each case chapter.

In terms of access to participants, the research began in August 2013 when the author's supervisor received a request for impact evaluation advice from Imagine's Project Manager (Leonard). Leonard was aware of the supervisor's research into impact assessment, and they set up an initial discussion with the philanthropy CEO Vijay, in which he explained how funding NGOs supported his desire to give back to society and reduce poverty in his home country, India. The supervisor informed the author (his PhD student at the time) about the case, with a view to it being developed as part of doctoral research into NGO impact evaluation. This led to a joint collaboration between August 2013 and April 2014 featuring a series of online group conference calls, the co-design of evaluation tools with Imagine UK and Rural India's director, and plans to support Imagine's ongoing development of an impact evaluation management information system (MIS) for the Indian NGO sector.

Collaboration centred on the conference calls between Leonard (project manager for Imagine based in London), Chandan (NGO director of operations, based near Delhi), Paul (author, PhD candidate) and Nigel (professor, Paul's supervisor). Visits to Imagine's headquarters in the UK and to India were planned for Spring 2014, but in the end did not transpire. In March 2014, Leonard informed Paul he was being transferred to another project and would be unavailable to continue with the collaboration. By April 2014, Chandan had become less committed to the collaboration plans and visits, and the case study stopped generating new data. In retrospect, and for ease of communication, the study can be conceptualised as occurring over three phases, as shown in Figure 4.7.

Research Timeline
NGO Case study 1 - Rural India & Imagine UK

PHASE 1	PHASE 2	PHASE 3
Initial contact from Imagine requesting advice; phone calls to scope project	Main data generation via group conference calls, reports, statistics, and discussion of co-designed research tools e.g. new survey, new qualitative interview designs	Plans for field visits to India and UK HQ fail to materialise; Leonard confirms he is leaving role; case winds down

7/2013   8/2013   9/2013   10/2013   11/2013   12/2013   1/2014   2/2014   3/2014   4/2014

Phone call 1 — Group call 1 — Group call 2 — Group call 3 — Group call 4 — Group call 5 — Group call 6 — Phone call 2

Figure 4.7: Research timeline for NGO Case 1 Rural India and Imagine UK

As can be seen, the methods or techniques used to generate data with the partners depended very much on the above opportunities to engage with the funder and NGO managers, which allowed the author to offer advice on and suggest critical views on impact evaluation. Group conference calls, telephone calls, document analysis and data spreadsheet analysis constituted the kinds of data generated. Probing and critiquing evaluation processes involved the co-design of evaluation tools (e.g. a survey, an interview plan, and an ICT feedback sheet) alongside other suggestions by the author, such as running in-depth interviews with farmers to generate qualitative data and learn their views and needs. In this sense, the suggestions and co-designed tools acted as "cultural probes", designed to provoke responses from participants, but not to prescribe or dominate the research process (Gaver et al, 1999: 2). They allowed the researcher to study organisations without asking for formal accounts or prepared answers, enabling the research participant to "*determine the response*" to teasing stimulations (Salancik, 1978: 638). These techniques were part of the critical engagement approach and aligned with CHAT principles that foreground dialogue, multiple perspectives, and a focus on participant practices.

Effectively, the engagement with Rural India and Imagine was one of exchange: the author offered advice on evaluation, and the participant NGO and philanthropy allowed access to reports, data, and their management staff during their 2013-2014 evaluation process. The

collaboration involved both critical research and applied consulting. Empirical data generated during the collaboration came from: phone calls, online group conference calls, and documents such as project reports, statistical evaluation data in spreadsheets, co-designed evaluation tools, author field notes, and email exchanges. Key participants in the study were: Vijay (the Imagine founder and CEO; Leonard (Imagine Project Manager) and Chandan (Rural India Director). Table 4.2 summarises and the empirical interactions. Detailed information is in Appendix 3.

Type of data	#	Details and focus
Phone calls	2	Initial scoping call and termination call
Group conference calls	6	Previous evaluations and current 2013-14 plans
Documents	2	Livelihood project report (2013), beneficiary profiles
Spreadsheets	10	Data on 4,000+ farmers; demographic & agricultural focus e.g. crops, income
Co-designed documents	3	3 research tools: 1) new survey for quantitative data capture 2) new interview script for qualitative data capture 3) new feedback survey for staff/volunteers using house-survey ICTs/software
Notes	50	Number of pages of notes taken by researcher (notepad & digital)
Emails	55	Call arrangements, attachment sharing, evaluation planning

Table 4.2: Overview of Rural India / Imagine Case 1 data, August 2013 - March 2014

This data, together with the CHAT analytical concepts discussed earlier, supported analysis of how senior managers at the philanthropy and the NGO designed evaluation processes. A description of the data analysis process follows in section 4.5. below.

4.4.4 NGO Case 2: context, access, timeline, methods and data

This section introduces the research context, how accessed evolved, a case timeline, methods used to generate data, and details of the data used in the study. In this case, research centred on Hill Tribe Support Group (HTSG), a local NGO in northern Thailand with normally around 20-25 staff. HTSG were founded in 1996 when two staff left a large international NGO during a period when it was withdrawing funding from Thailand. HTSG provides healthcare and education services for minority ethnic communities and children in isolated rural settings that lack transport or telecommunications infrastructure. Unlike Rural India, HTSG's funding comes from diverse sources, including direct public fundraising, local and national government, national and international private foundations, international NGOs, the UN, and the EU.

In terms of access, the study began when HTSG responded to an email enquiry about collaborative research on impact evaluation in May 2014. A series of email exchanges between the author, two office staff (Susan and Orr), and the Director of HTSG (Khun) led to a one-week site visit to HTSG in Thailand in August 2014. This visit included discussions and conversations with key office and field staff, group meetings, formal and informal social activities such as lunches and car rides, a presentation by the Director, a meeting with a board member, and an evaluation needs analysis workshop run by the author.

Between August 2014 and February 2015, discussions continued over email and included a proposal for further evaluation research and advice as well as a second trip to build HTSG's impact evaluation capacity. The second trip was for 10 days in February 2015, and featured further conversations, group discussions with field and office teams, meetings, and a series of workshops lead by the author that focused on evaluation approaches and implementation. Collaboration continued through 2015 and 2016. During this time, it focused on distant support for the design of an evaluation baseline study as part of a healthcare project extension to two new villages, advice and support for a new Migrant Project funding application to international donors, and a baseline evaluation to start the Migrant Project.

Figure 4.8 shows the research timeline through four phases from May 2014 to December 2016.

Research Timeline:
NGO Case study 2 - HTSG Thailand

PHASE 4

Migrants Project funding application, project design, evaluation advice & support

PHASE 1

Emails, reports
Field visit 1
Meetings re projects, partners, evaluation needs & collaboration

PHASE 2

Field visit 2: needs, capacity building, workshops & co-designed tools for Health Project team

PHASE 3

Health Project new village evaluation advice & support

| 4/2014 | 8/2014 | 12/2014 | 4/2015 | 8/2015 | 12/2015 | 4/2016 | 8/2016 | 12/2016 |

Field visit 1: 8/2014

Field visit 2: 2/2015

Group Call: Health & Migrant Projects

Data analysis: Health Project

Group call: Migrants Project

Funding bid: Migrants Project

Survey analysis: Migrant Project

Figure 4.8: Research timeline for NGO Case 2 HTSG Thailand

In terms of both data generated and methods, this case was more extensive than case 1, involving two field trips to Thailand by the researcher. The methods and techniques used to generate empirical data during the study included: one-to-one conversations, small group discussions, large group meetings, presentations, workshops, field notes, recordings, email exchanges, reports, spreadsheets, photographs, diagrams, online conference calls, and the co-production of evaluation documents and plans created during workshops and online collaborations. Conversations, meetings and workshops generated data about specific NGO activities, not an "*entire culture and social life*" (Hannerz, 2003: 208). Conversations were not interviews structured in advance, but open-ended to encourage participants to discuss their contexts and needs (ibid: 209).

As with case 1, CHAT informed the interactions and engagements. Probes or stimulations by the researcher were used to explore how evaluation was perceived by or could be arranged by the NGO, to further understand participant views, plans, goals, and limitations. A key feature of the research was the workshops held during visit two, in which the author led sessions on

evaluation approaches and techniques, and learned about the novice status of the NGO in terms of their impact evaluation experience, expertise, resources, tools and attitudes. The workshops themselves had a dual focus – to offer consulting support about TIEK, but also to test and probe NGO concerns, expectations and needs. In this sense the researcher was not a "*passive recorder who avoids provoking responses*" (Salancik, 1978: 638), but was an active researcher "*playing the role*" (ibid: 639) of a consultant. This kind of probing admittedly followed the author's interests (Gaver et al, 1999: 29) in testing critical insights, but it also allowed the researcher to build dialogue and opportunities for learning around the limits and problems faced by the NGO. Examples of this included how needing to do marketing with impact stories surprised staff, or how exploring evaluation voices and networks in the first field trip prompted the NGO to develop data brokering and sharing relations with partners. These responses and others are described in more detail in the case chapter.

The author provided advice and support for HTSG on how to plan, approach and conduct impact evaluations. In return, HTSG allowed access to their staff, work processes and project information. In retrospect, the author influenced HTSG's approach to impact evaluation. As with Rural India and Imagine, only by offering a "service" as a contribution, and by engaging with the NGO's needs, was access and the testing of critical insights possible.

In terms of data generated during the case, Table 4.3 gives an overview. Appendix 4 adds detailed information on all empirical interactions and sources. Key participants included: Khun (NGO CEO), Susan (Donor liaison and US intern), Orr (Finance Officer), Pang (Healthcare Project Manager), Cherry (Evaluation Specialists and US Intern), and Mia (Migrant Project Coordinator and US Intern).

Type of data	#	Details and focus
Emails	100+	Arrangements, requests, advice; often multiple-page messages
Site visits	2	Visit 1: August 2014 (5 days). Visit 2: February 2015 (10 days).
Documents	34	E.g. annual reports for funders
Photographs	43	From field trip visits, of groups, documents, workshop materials

Spreadsheets	7	Project data on beneficiaries, staff and volunteers
Co-designed documents	8	Evaluation plans, brainstorms, tools (e.g. survey or interview questions designs), funding application sections
Group meetings	12	In offices, meeting rooms, restaurants; 3-10 people, 30-120 mins
1-1 conversations	10	With one staff member or director, 15-25 mins on average
Presentations	2	By director (1) and by researcher (1)
Workshops	6	Focus on learning aspects of impact evaluation
Group calls	2	Online group conference calls for evaluation advice/planning
Website pages	20+	HTSG homepage, about, project pages
Field notes	100+ pages	Hand-written in notepads, digital on laptop, text & diagrams

Table 4.3: Overview of HTSG Case 2 data, May 2014 - December 2016

Study participants included HTSG's director, one board member, office and field staff, the healthcare project team, and three US interns who volunteered at the NGO for one to two years each during the period of the research. Twelve out of 25 staff members took part in the research.

The author speaks Thai to work level, although with some grammatical inaccuracies, and a staff member / translator was available for all group events. Recordings were made of workshops. Field notes were taken throughout, using a notepad and pen or a laptop during or shortly after events.

Interactions centred on HTSG's impact evaluation goals, plans, knowledge and capacity, with the content of conversations, group calls, and workshops oriented around supporting their evaluation learning and a funding bid. The author's input included: asking if and how evaluations were being carried out already; asking what resources were available (e.g. human, technological, time); probing how to improve the NGO's awareness of and performance of

impact evaluation; and probing how to use impact and evaluation knowledge in planning, communications and funding. Overall, the empirical data and analysis are used to describe HTSG's evaluation work and how their learning about impact evaluation was implicated in changes to their organisation and work.

4.4.5 Data analysis

Analysis of the data generated in both development NGO cases utilised the CHAT framework described earlier (Engeström, 1987; Karanasios, 2014) to answer the three research questions. The first question required identification of evaluation activities, using CHAT's conceptualisation of practice as activity systems. This focused on who was doing what, with which techniques or technologies, rules, roles etc, and what the outcomes were in terms of data/knowledge products. The response to research question two included an analysis of how power relations were generated during impact data/knowledge construction activities, if specific activities featured use/exchange value contradictions in activity systems, how data/knowledge was edited or mediated, submerged or elevated in sequences of activities, and how these processes changed or transformed the NGO's work. The third research question involved exploring, through critical engagement, if and how power relations could be elevated, or if alternative evaluation practices could be considered, to amplify the 6P sensitivities around power and practice. Together, the questions, analysis and responses support the conceptual critique which is part of the critical engagement approach adopted in the study. Figure 4.9 shows the trajectory of this analysis in four steps, moving from the research questions on the left (column 1), through CHAT's concepts (column 2), to examples of potential data (column 3), and how the analysis relates to the 6Ps (column 4).

Data analysis steps - from CHAT to the 6Ps

① Research Questions	② CHAT concepts	③ Data examples	④ THE 6Ps
RQ1: How is small development NGO impact data/knowledge constructed in practice?	Identify & analyse activities, activity systems, outcomes, activity sequences & networks	Show activities in processes e.g. from designing questions, to capturing data and report writing	How can the CHAT analysis support wider views of impact data/knowledge using 6Ps, in contrast to TIEK/DIKW?
RQ2: Are / How are power relations generated during impact evaluation data/ knowledge construction at small NGOs?	Identify contradictions Identify how impact data/ knowledge are mediated or shaped by elements such as tools, concepts etc Identify activity sequences of submerging & elevating	Submerging or elevating specific: - voices (e.g. villagers, staff) - outcomes (e.g. data types) - models / methods (e.g. stats) - resources (e.g. staff skills, ICTs) - transformations in NGO work	How can the power relations identified using CHAT, align with the 6Ps? Can CHAT support a 6Ps view? Can CHAT articulate the sensitivities, omissions, governing processes clearly?
RQ3: Can / How can power relations be addressed and responded to in impact evaluation data / knowledge construction practices at small NGOs?	Show how impact is viewed from multiple perspectives, what constrains or promotes reflection and responses to power relationsMultivoiced view of impact	NGO acknowledges or acts to: - hear diverse voices in network - share data/knowledge - reflect / experiment with tools, methods, processes - elevate submerged contexts, voices or political dialogue	What lessons are learned from the cases about noticing, reflecting and acting on 6P sensitivities or silences in evaluation? Can 6Ps or critical configurations be elevated and responded too? Can NGOs respond to the 6Ps and critical silences?
	Analysis in Case Chapters		**Review in Discussion Chapter**

Figure 4.9: Process of data analysis using CHAT and the 6P sensitivities

Finally, the analysis steps one to three in Figure 4.9 are described in the case study chapter analysis sections. However, a more comprehensive review of the 6Ps and CHAT, and how they each inform critical engagement is reserved for the discussion chapter.

The analysis of evaluation as activities avoids the methodological individualism of Engeström's (2000b) three umpires, replacing individualism with activity as the unit of analysis. This move also incorporates the TIEK and DIKW marginalisation of power and practice evident in the impact evaluation meta-model that reduces evaluation to the inputs/outputs "methods plus results" frame.

4.4.6   Limitations, generalisations, ethics, security and anonymity

Four limitations in the approach and design invite a degree of caution and care when relating the findings or generalising to other settings in the development sector. Firstly, the two NGOs are small service providers, not large international NGOs. Secondly, the evaluations were

conducted by NGO staff, not external consultants. External evaluations are common in development (Picciotto, 2013: 20). Thirdly, the participants were NGO managers and staff, meaning a broader spectrum of participants, for example beneficiaries or donors, would have supported a more extensive analysis. Fourthly, the research questions, methodology and use of CHAT (Karanasios, 2014) support a micro-analysis of evaluation as practice, not macro-level samples of a sector or a national policy. Therefore, lessons from the study are limited in how much the analysis resonates or can be generalised to specific or wider sector experiences.

In terms of ethics, the study was approved by Lancaster University's ethics committee, and three supervisors, including two professors. The ethics review featured a research plan, participant information sheets, and consent forms. Both NGOs agreed to collaborate on the research, and participants at each event where given the option of joining activities, not participating, or providing comments off record. A participant information sheet and a consent agreement are included in the appendices.

In terms of security and anonymity, data was stored on a single secure, password-protected, digitally encrypted computer. All names and references to people, projects, organisations and sites were anonymised. "Rural India", "Hill Tribe Support Group", "Leonard", "Khun" etc are all pseudonyms. Anonymity and confidentiality were important for establishing trust, because as mentioned earlier, NGOs are difficult to study and face much scrutiny (Wallace et al, 2006; Lewis & Kanji 2009; Roche, 1999). One UNDP organisation in South Asia initially agreed to take part, but pulled out of the study in mid-2014 because management were uncomfortable with the focus on evaluation practices. This shows how anonymity, as well as contributing to the NGOs' work, was an essential, yet fragile, ingredient in sustaining researcher access.

4.4.7 Reflecting on CHAT as part of critical engagement

There are three aspects of the study related to the application of CHAT, which require self-critical reflection. Firstly, how the author was an agent in the study, a CHAT "subject" influencing the cases. Secondly, how the use of CHAT did not successfully incorporate a full expansive learning cycle, or follow a traditional interventionist process of progressive transformation. And thirdly, how specific concepts and tools were found more useful than others from CHAT's conceptual repertoire.

Firstly, the author cum researcher[25] was an active participant in case one. From a CHAT perspective the author occupied a "subject" positions in a CHAT activity system, working to produce scholarly outcomes, such as a PhD thesis. In case one, this influence is acknowledged when the researcher acting as a consultant offering advice and stimulating critical responses from the NGO, suggests redesigning the quantitative evaluation survey, adding new qualitative interviews to the plans, as well as an ICT feedback questionnaire. Ultimately, the NGO adopted the quantitative changes, but rejected the qualitative interviews and ICT feedback tool. In doing so, they exerted their control over the evaluation process, and simultaneously articulated the ways in which they valued certain aspects of evaluation (statistics), but not others (richer understandings of farmer voices or ITC usage).

In case two, workshops and evaluation plans constituted key influencing tactics on behalf of the author. This influence is best viewed as an outside consultant or expert bringing technical impact evaluation knowledge (TIEK) into the case, albeit at the request of the NGO. This constituted a form of power over the NGO as they were novice evaluators. In this way, the author was entangled in the power relations and potentially could have framed what impacts became legitimate or not, a core topic in the study. However, the second NGO, although adopting collaboratively designed plans, could not fully implement them. Their lack of capacity, staff, data, systems, attitudes and skills to do so became a key result of the study. This shed light upon the very problem of expert control, which the researcher was entangled in.

As such, in case one the NGO was expert enough to reject the researcher's suggestion of hearing broader voices framing impact. In the second case the NGO power relation with the author, a novice NGO trying to absorb TIEK from an expert outsider, became a key concern of the study, articulated in the chapter 6 and chapter 7 as an evaluation capacity gap. From a CHAT perspective, in case one the researcher's subject position "power" was rejected, but in case two it was illustrative of broader inequalities in evaluation knowledge production capacity.

Secondly, why was the study not set up, as many CHAT studies are, as a formal intervention, adhering to a complete expansive learning cycle, to create transformative change? Firstly, the

---

[25] The author's doctoral supervisor joined the first phone call with case one participants, and was a participant in two early group calls, albeit rarely speaking. His role was supportive for the author initiating the case, and an influence upon the start of the case itself. He played no part in the remainder of the study, and no part in case two.

NGOs, funders, and beneficiaries were dispersed an unable to come together for change lab sessions, as discussed earlier. Second, there was no mandate for reflective transformation, only technical consultancy in exchange for single researcher access to NGO staff, reports and data. Thirdly, CHAT resources were used to critique, stimulate or probe NGO activities to further understanding how power, data and knowledge were related in impact evaluation work. CHAT was thus used as part of critical engagement, not transformative change, which fits with the wider literature on power and knowledge relations, audit culture and managerial controls in the aid sector. These are issues which are wider than single locales of transformative change, and the study contributes to critical understanding these wider demands and practices. In particular case one illustrates why and how an NGO may reject progressive change involving farmer agencies, voices and narratives in contemporary aid markets. Case two shows how an NGO may lack resources, capacity and attitudes to catch up with contemporary technical evaluation knowledge, and thus be left in an ongoing knowledge capacity gap. A critical use of CHAT, rather than a change lab perspective, thus informed the development of the research approach (i.e. "critical engagement"), and the lessons learned in the study. This is reflected on further in in chapter 7, particularly in relation to contribution three, on critical engagement.

Thirdly, the Change Lab approach used in much CHAT research makes use of tools and devices such as video recordings, CHAT terminology for social scientific concepts, diagramming activity systems with participants using the canonical CHAT triangles, and double stimulation loops. Given the mandate from the participating NGOs to provide evaluation consultancy advice in exchange for research access, there was no authorising environment, resources, time or staff availability to conduct a series of CHAT change lab sessions or workshops. There was no mandate to teach CHAT terminology, particular when working across languages in case two. With the literature suggesting that power/data/knowledge problems are distributed over global aid sector markets and bureaucracies rather than configured in one local site, and the lack of research resources as this was a self-funded PhD project, not an institutionally funded intervention, a critical probing approach to NGO practices was deemed more valuable than what would have been very limited attempts to convene dispersed stakeholders, teach CHAT terminology, and stimulate a transformative process. Therefore, instead of focusing on the multi-voiced account of inaccessible dispersed stakeholders, the historicity of a single NGO locale as opposed to global aid networks, Leontyev's hierarchy of a single activity, or the fragmented local object, the study focused on how dispersed voices were elevated or submerged by the NGOs in their evaluation activities, how broader sector history has led to an

audit culture which constrains evaluation practice (see chapter 2 and 7), and how a wider global network of agents and relations is deemed significant, in contrast to a localised activity hierarchy. These issues are reflected on in chapter 7 and chapter 8.

4.5   Learning about critical engagement

This chapter has advocated critical engagement as a way of exploring the spaces between scholarly critiques and unreflective interventions. The chapter has outlined an approach to researching development NGOs, and a methodological frame to operationalise the approach. This frame draws on particular concepts and tools from CHAT for engaging with research participants, and testing critiques in the field; in other words, in the wilds of development evaluation, rather than in scholarly exchanges or debates (Law, 2008: 150). The frame also draws conceptually on the 6P sensitivities from the theory chapter to avoid the assumptions and omissions embedded of TIEK and DIKW. The research design reviewed the study aims, reconfirming its cultural historical grounding in contrast to positivist or interpretivist traditions. It also outlined the case design, selection criteria, and described how particular methods and probes were used to generate data. Lists of data, participants and activities were included, as well as explanation of how the analysis was performed, and a final review of study limitations.

In summary, development writers have highlighted the need for "engagement plus critique" in studies of power and politics over many years (e.g. Walsham, & Sahay, 2006; Gardner & Lewis, 2015). In searching for such alternative approaches for understanding impact and evaluation, Hickey's (2012: 1243) hybrid of critique, technical knowledge and engagement, incorporating a sensitivity to power, working with local institutions (even though they may not be perfect), and deploying evaluations which acknowledge history and politics, supports a conceptual shift away from "*ever more sophisticated ways of measuring and identifying*" (ibid: 1243-4). Hickey's shift is thus also towards critique and engagement. The conceptual aspect of the shift requires looking at evaluation practices that stretch beyond target communities and which permeate wider aid-chain dynamics, often through professionally unobtrusive, mundane or seemingly intangible development 2.0 data/knowledge flows, such as those this study focuses on. To engage with and critique these activities and flows as a methodological move, acknowledges the diffused politics of knowledge generation (Gardner & Lewis, 2015) and evaluation results (Eyben et al, 2015). It puts power/data/knowledge relations front and centre.

Hickey's proposals imply a need to explore spaces between scholarly critique and unreflective engagement so as to understand such relations. The spaces occur between established approaches, methods and industrial sectors such as academia and development. Power relations in these spaces are hard to study legitimately because they cross institutions, professions, and silent, normalised regimes of truth (Marcus, 1998). Critical engagement, using critical concepts in CHAT as well as the 6P sensitivities, offers a possible framework for exploring and responding to such challenges.

5 NGO Case 1: Rural India and Imagine UK, expert impact-makers

5.1   Introduction

This chapter describes the first NGO case study, that of Rural India and their funder, the Imagine Foundation UK (Imagine). The case responds to the three research questions: identifying impact evaluation activities (research question 1); the generation of power relations during data and knowledge construction (research question 2); and, how NGO understand and respond to such power relations (question 3).

Section 5.1 introduces the chapter and section 5.2 recaps the case setting. Section 5.3 is a comprehensive description of the empirical data generated. This is largely observations, quotations, processes, participant claims, and samples from evaluation data spreadsheets. The empirical evidence is used to identify different sets of activities that constitute an evaluation cycle. To make the range of activities involved in the cycle clear, the findings draw largely on how the cycle is managed, illustrated through the work and comments of Imagine's project manager, Leonard and Rural India's director, Chandan. Activities are described with reference to the relevant periods of the evaluation cycle.

Section 5.4 analyses activities identified in the previous section. Firstly, CHAT activity systems, contradictions, and two forms of impact are analysed. Secondly, the novel concept of

temporal chain of activities is used to illustrate the incremental generation of unequal power relations during impact data/knowledge construction. Section 5.5 summarises the key findings and analytical consequences of the case. Of particular significance is the story of how experts construct impacts, and how, in practice, that construction process generates inequalities of power, data and knowledge.

5.2   Case setting

The Imagine Foundation UK was started by "Vijay", an Indian national who owns a capital investment fund in the UK. Imagine, the philanthropic arm of the fund, finances NGO projects across India and uses data and lessons from their activities to develop software for the development sector. Example software applications have included a crowd-funding platform and an evaluation management information system (MIS). Imagine's grants to Indian NGOs over the last decade total in the tens of millions of dollars.

Rural India is a small NGO working on agriculture and livelihoods projects with female farmer groups in central India. Their aims are to encourage self-sufficiency through improved farming practices and to reduce rural poverty. To achieve these aims, Rural India runs empowerment and livelihood programmes in numerous villages. Target beneficiaries are predominantly poor female farmers, who enrol in Rural India's projects voluntarily.

Rural India's projects run over multiple years, and Imagine's 2013 annual report describes how typically in the first year of a new project activities include community mobilisation, formation of Women's Livelihood Groups, capacity building through training, and setting up agricultural demonstration plots. Villagers are supported to grow crops including mustard, wheat, or cotton, raise livestock, keep livestock healthy, reduce livestock diseases, and ensure livestock are vaccinated appropriately. In year two, livelihood centres are established that provide "*one stop solutions to all problems related to agriculture and animal husbandry*" (Imagine Foundation 2013: 3, 14 Annual Report). These centres offer support to farmers on crop choices, cultivation methods, marketing, branding and packaging of products, access to government services, and financial support. Rural India projects and activities have reached over 10,000 farmers since the mid 2000s.

5.3 Empirical section

In evaluation literature, motivations are important parts of evaluation processes (Roche, 1999); rigorously knowing what works and why it works is important (3ie, 2012). A critical literature a similar will to know is described (e.g. Li, 2007) which relates to the need to know impact in this study. The will or need stimulates ongoing knowledge construction and sharing in wider networks (Ilcan & Philips 2010: 861; 2008; Mosse, 2004a). As noted earlier, such networks are significant in CHAT too (Blackler 1995; Engeström 1987: 103; Karanasios, 2014). In the empirical descriptions below, CHAT network activities and elements can be identified: objects of activity (e.g. needs, goals); subjects (the funder Imagine, the NGO Rural India, managers, volunteers); conceptual / technological tools (e.g. evaluation methods, digital spreadsheets and tablet computers); rules and norms (e.g. capturing and filtering data); and roles and division of labour (e.g. decision-making, meeting villagers). The analysis of these components using CHAT terminology is pursued in section 5.4, but first 5.3 describes the empirical fieldwork itself, starting with the partners' "need-to-know" impact.

5.3.1 Early-cycle activities: the need-to-know impact

The research collaboration with Rural India and Imagine evidenced ways in which a need-to-know impact was expressed. Imagine needed to know impact and needed Rural India to capture and organise data supporting the foundation's need-to-know. In this way, over time Rural India came to share the needs of their funder. How such needs are satisfied involved assumptions, methods, technologies and relationships.

Vijay, the Imagine Foundation founder and CEO, first talked about the need-to-know the impact of his philanthropy's aid funding in a conference call in August 2013. Vijay explained how his motivation for evaluating Rural India was part of "*accounting for the impact of $10 million in annual funding*". He assumed that the need to "*measure the effectiveness of services in the financial sector, and in development*" was legitimate and desirable.

Vijay further described how he migrated techniques and technologies from his business and financial ventures, which allowed him to know by measuring. Impact then could be known through careful measurement and analysis. Knowing could be achieved by quantitative survey methods and modern digital technologies, such as database systems, mobile tablet computers,

and customised marketing software applications. Knowing impact required specific activities of capturing, quantifying and calculating. Tools, techniques and technologies were thus migrated from Imagine's commercial operations in the finance sector to their philanthropic aid programs, to measure the effect of funding upon target participants, and the effectiveness of their grantees, such as Rural India.

Knowing impact in India and representing it for Vijay and his trusted project manager, Leonard, at Imagine in London was important for the philanthropy. Impact data and knowledge supported London decision-making and understanding about what was "*really*" happening in India. The place of technologies in framing and satisfying the need-to-know over distance was crucial, as illustrated in this quote from Leonard:

> *"The first thing is just to test ... test how the technology can help. We want to ... y'know so that it gives valuable information faster, to add value and so we can understand what's really happening. ... Obviously, the technology is a tool for capturing ... better ... and faster ... and analysing the data that is the key." (Leonard)*

Documents provided by Imagine and Rural India showed that the partners had been working together and assessing impact since 2006. By 2013, the two organisations had established a regular annual assessment, analysis and reporting process. Discussing this process, Chandan, Rural India's manager, stressed the importance of long-term relationships with local aid partners and villagers. These provided a "*good grip*" to facilitate data capture and thus know impacts.

> *"The typical villagers will not open up if you haven't worked with them for a long time, so our partners are chosen such so they almost adopted these places [target community places] ... So they have been working there for ... about ... oh donkey's years now, about 8-10 years now. So they have a good grip on the villagers now. A good grip means they know the family size, they know the landholdings, these are already known ..." (Chandan)*

Such a good grip required not just business assumptions, quantitative evaluation methods, and digital technologies then. It also benefited from data legitimated by long-term and local relationships.

In summarising the activities involved in needing to know impact, it must be noted that the partners were aligning methods, technologies, assumptions and relationships with ongoing and future digital strategies. These strategies included the establishment of a crowd funding platform and an aid sector evaluation management information system (MIS), which would be available to other organisations "*for a small cost at first*" (Leonard, September 2013). This alignment and resultant MIS supported other aid actors to know impact in similar digital ways to the partners, and furthered the philanthropy's positioning in the sector. It is difficult to separate these methods, technologies, assumptions, relations, and organisational strategies from the evaluation process they set in motion.

Activities constituting the need-to-know impact

- Establishing the need(s) to know impact

- Communicating the need-to-know impact

- Satisfying the need-to-know via assumptions, methods, technologies, and relationships

- Aligning the need-to-know impact with longer term organisational strategies

5.3.2 Early-cycle activities: Assembling an impact machine

Over years of partnership, Rural India and Imagine had rehearsed their annual evaluation cycle multiple times. Evaluation activities included: establishing authority lines, designing surveys and questions, data storage, customising survey software and tablet computers, assigning NGO office and field staff, enrolling village volunteers, and training volunteers and staff.

Vijay had appointed Leonard as overall manager. Leonard's role encompassed managing project evaluations from his base in the UK. He was the lead person in terms of planning and reporting and initiated the research collaboration with the author. In tandem, Rural India's Director, Chandan, managed the evaluation process on the ground in India. He supervised NGO office staff, field staff, partnerships with local groups, recruitment of village volunteers, and the training of staff and volunteers to conduct surveys in farmer households. He also supervised the input or upload of data into a central database.

Design of surveys and questions was a necessary step which initiated and preceded the collection, analysis and writing up of evaluation documents. Chandan confirmed this:

> *"The collection of the data is the first thing to be doing, then ... then analyse, then write good documents, for the research etc". (Chandan)*

The survey question format and focus were critically important. Discrete question formats included short-response survey questions, prefiguring expected numerical, Yes/No, or short phrase responses. Questions were not designed to elicit longer responses.

In terms of topic and content, survey questions focused on female farmer agricultural practices, such as how crops were grown or how livestock were reared. Further questions related to family size and composition, livelihoods, household finances, income, expenses, health and medical services, children's education and schooling, transportation types used, utilities (e.g. water access, toilet types, electricity access), and details about other local services and facilities.

In September 2013, Leonard and Chandan shared two documents which listed examples of the questions to be used in the 2013-2014 household surveys. These included: "*Have crop insurance?*", "*Having own water well?*", "*Having toilet at home?*" (Figure 5.1).

Figure 5.1: Questions to be included in the 2013-2014 household survey

In addition to the survey method and question format, Imagine provided technologies to Rural India to help implement the evaluation process. Technologies included tablet computers (Android operating systems), desktop office computers, a business marketing application that was customized for evaluation needs by a software team in the UK and loaded onto the Android tablet, spreadsheet software (Microsoft Excel), and in more recent years, an online database application. In the past, NGO field workers had brought data back to the office and manually uploaded it onto spreadsheets on a central desktop PC. However, in 2013 the partners launched a shared web-based database application. From then on, all survey response data was uploaded in near-real-time during and after farmer household survey visits, then stored in the online database, making the process faster and more efficient.

Chandan oversaw the staffing requirements of Rural India, and enrolment of village volunteers to help deliver the evaluation surveys in village households. Both groups, the NGO field staff and village volunteers, received training before running evaluation activities either in villages or with other stakeholders such as local government or community representatives.

> *"We try to collect information there on the ground, from different stakeholders, the local government, the community, people, data from all the households. And we train the volunteers in the ways and means of how to conduct a household survey. If you see the survey and the Excel sheets, you'll see there are so many parameters, they collect that and put it up in the website spreadsheet". (Chandan)*

In summary, the evaluation financing, needs, methods, measurement techniques, question and response frames, recruitment of staff and volunteers, training processes, and digital technologies (application, database, tablets, desktops and customisations) were all sourced from the UK philanthropy and adopted by the small Indian NGO. A series of activities were required to generate data, capture it, and store it. Each activity involved certain choices, people, tools (conceptual, technological), group objects of work and outcomes, which, when viewed all together, assembled a purposeful evaluation machine.

Identifiable activities in the machine are were:

1. planning data design, capture and storage based on the need-to-know impact
2. sourcing methods and tools (conceptual / technological)
3. establishing authority lines
4. managing NGO office and field staff tasks and labour

5. recruiting and managing village volunteers

6. training staff and village volunteers in how to run household surveys

7. designing survey questions, in terms of format and target content focus

8. customising hardware and software (e.g. scripting app and Android tablets)

9. conducting household surveys in villages and capturing survey response data

10. inputting / uploading data to a central store (e.g. office PC or online database)

5.3.3 Mid-cycle activities: Making cells, making clamour

Assembling the mechanics of impact evaluation was necessary for Imagine and Rural India to capture, store, filter, analyse and transport data, in order to show what was "*really happening*" in a village project. Legitimate data was captured in cells, rows, columns and sheets. Illegitimate data was dismissed and disregarded, deemed at times "*inarticulate*", "*clamour*", or even ignored as silences left in the field. This section demonstrates how data was turned into either cells or clamour.

It is important to note that Rural India's role was prominent in the capture and collection of data. Office staff and village volunteers would take the Android mobile tablet and survey app software into villages and visit houses door-to-door doing survey interviews. In recent years, they had begun to upload collected data to a web-hosted database, which was, according to Leonard and Chandan, more efficient than the old method of returning to the office and manually inputting data into a database on a desktop machine.

Leonard saw technologies such as mobile tablets, spreadsheets, cells, rows and columns, the Android application, and the MIS in development as instrumental in facilitating the efficient transport of data from Rural India to Imagine UK.

> *"The first thing is how the technology can help. We want to ... y'know so that it gives valuable information faster, to add value and so we can understand what's really happening. Obviously, the technology is a tool for capturing better and faster and analysing the data ... that is the key." (Leonard)*

On the surface, Rural India's adoption of the philanthropy's process - capturing, transporting and analysing digital data in cells, rows, columns and spreadsheets - was not remarkable. It

was a normal aspect of technical evaluation. This process facilitated professional and distant analysis, report writing and decision-making. The data "cell" was a professional and mundane component of evaluation (Figure 5.2). Cells were very normal, bounded, square or rectangular boxes, containing single referents. Cell contents were framed by survey questions. Cell contents represented female farmers' oral survey responses during household interviews.
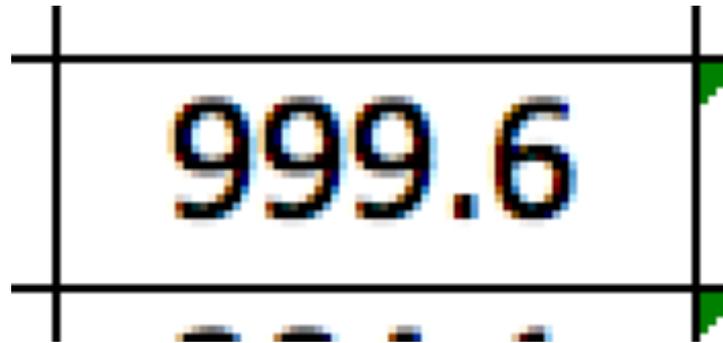
Figure 5.2: The cell, a basic and bounded container for impact data

The cell and its contents were manipulated in a number of ways. Rural India put cells into columns according to question type. They arranged horizontal rows according to an individual farmer's name, identity number, and project group number. They separated cells from other cells, and in many instances, rounded-off figures to one or two decimal places. This regimentation of cells into columns and rows allowed Rural India and Imagine to assign to individual farmers discrete values from discrete questions, such as per capita income or value of crop yield per hectare. Rows fixed discrete values to female farmers. Columns fixed discrete values to specific questions (Figure 5.3). This arrangement supported later mathematical calculations, such as simple column totals, averages, before / after intervention comparisons, and comparisons across villages.
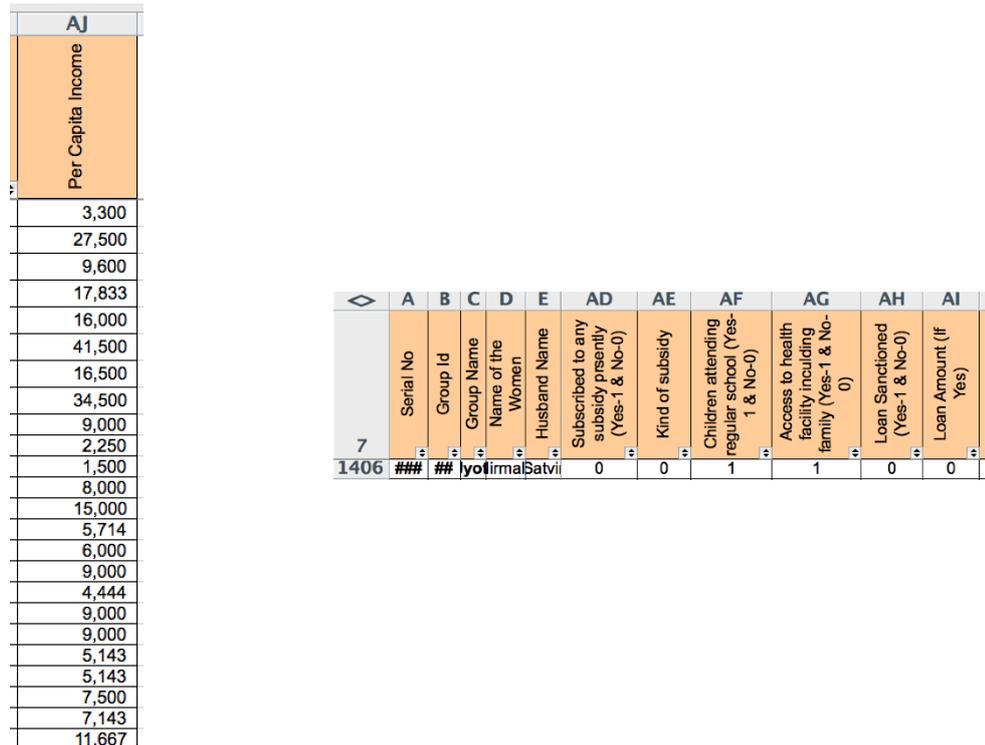
Figure 5.3: Data rearranged into mundane columns (left) and rows (right)

The path that the data travelled along firstly involved farmer-talk responses vocalised in their houses that were incrementally interpreted, captured and recorded in survey software on the Android tablet, using discrete question-answer fields in a survey app. Secondly, Rural India staff uploaded cell representations onto an internet-accessible web-page application whilst still in the field. Thirdly, the data was amalgamated into spreadsheets with many cells, rows and columns drawn from thousands of household interviews. Fourthly, spreadsheet data was analysed by Rural India's staff and sent to Chandan so that he could produce project reports for Imagine's office in the UK.

These data spreadsheets are illustrated in Figures 5.4 and 5.5 below. In total, Rural India and Imagine held data on hundreds of question items, relating to some 10,000 unique individual female farmers. This data had been collected between 2006 and 2014.

Total no. of family members	No. of members working in NREGS and having Job Cards	No. of members working as a pvt. Labourer	No. of members working in Agriculture	No. of members working in Animal Husbandary	Land (1=Yes & 0=No)	Area of Land (in beegha)	Livestock	Other fixed Asset	Type of Fixed Asset Land /House/ Vehicle/ Animal/Vehicle/	Annual Household Income (In Rs.)	BPL Family	Majority of Caste/Category (/ST/OBC/GEN)
5	1	1	1	1	0	0	1	1	House	28,000	APL	ST
4	0	0	1	1	1	5	1	1	House	36,500	APL	OBC
6	0	0	1	1	1	5	1	1	House	52,500	APL	OBC
6	1	1	1	1	1	2	1	1	House	42,000	APL	ST
4	1	1	1	1	0	0	1	1	House	61,000	APL	SC
4	1	1	1	1	0	0	1	1	House	12,500	APL	ST
5	1	0	1	1	0	0	1	1	House	61,000	APL	SC
5	1	0	1	1	1	1	1	1	House	29,000	APL	ST
6	1	0	0	0	0	0	0	1	House	50,000	APL	OBC
4	0	0	1	1	1	2	1	1	House	65,500	APL	OBC
7	0	0	1	1	1	4	1	1	House	66,000	APL	OBC
5	0	0	1	1	1	2	1	1	House	62,000	APL	OBC
6	0	0	1	1	1	2	1	1	House	52,500	APL	OBC
4	0	0	1	1	1	1	1	1	House	85,500	APL	OBC
2	0	0	1	1	1	3	1	1	House	91,000	APL	OBC
3	0	0	1	1	1	1	1	1	House	50,500	APL	OBC

Figure 5.4: Example of data captured and stored in spreadsheet software

As mentioned earlier, in one sense these data containers and the process is unremarkable; a mundane part of professional evaluation work. However, it is important to identify the cells and sheets, the steps in producing them, and the tools and activities that took place in constructing, storing and moving impact data. It helps us understand the practice and power of evaluation data construction and the contours of participation in the sequences of activities. The result of which was representations of impact for Leonard in the UK on "*what's really happening*" in India.

Farmers	Hasband	villeg	WLGs	Block	Monitoring	Area m2	Area hec.	deep ploug@80 Rs/h.	2 cross harr.@800Rs/h.	1 planking @400 Rs	Seed kg	Seed cost 30 R/kg	Fungicide	Insecticide	Azotobacter	PSB	Rs1000/h	FYM	N .Kg	N Cost 15.22 per kg	P kg	P.cost 48.4 per kg	K.kg	K cost	Ca cost	Mg cost	ZnSo4 55 p kg	Hand weeding	Herbicide cost	No.of irr.	cost of irr.2400 per irr.per hec.	Harvesting cost 5000/hec	Threshing cost 2000/hec	Total cost of cultivation	Total cost/Hectare	Total cost/Bigha
Asl	Bai	Bhoop	Be	Rol		996	0.1	79.68	79.7	39.8	15	450	0	0	0	0	99.6	0	10	154	5	242	0	0	0	0	0	0	0	6	1434	498	199	3276.24	32894	8223.5
Indraj	Gadod	Be	Rol			2500	0.25	200	200	100	35	1050	0	100	0	0	250	0	20	306	10	484	0	0	0	0	0	0	0	6	3600	1250	500	8040	32160	8040
Kul bh	Khapa	Be	Rol			2500	0.25	200	200	100	35	1050	0	0	0	0	250	0	20	307	10	484	0	0	0	0	0	0	0	6	3600	1250	500	7941	31764	7941
Ra	Sol	Tasing	Be	Rol		1250	0.125	100	100	50	20	600	0	0	0	0	125	0	12	185	5	242	0	0	0	0	0	0	0	6	1800	625	250	4077	32616	8154
Chote	Khatar	Be	Vis			2500	0.25	200	200	100	35	1050	0	0	0	0	250	0	20	308	10	484	0	0	0	0	0	0	0	6	3600	1250	500	7942	31768	7942
Ved pr	Bhood	Be	Vis			2750	0.275	220	220	110	40	1200	0	0	0	0	275	0	25	380.5	15	726	0	0	0	0	0	0	0	6	3960	1375	550	9016.5	32787	8196.8
Ratan	Chadic	Be	Rar			2250	0.225	180	180	90	35	1050	0	0	0	0	225	0	20	304.4	10	484	0	0	0	0	0	0	0	6	3240	1125	450	7328.4	32571	8142.7
Vir	Bh	Nagal	Be	Sar		2500	0.25	200	200	100	35	1050	0	0	0	0	250	0	20	304.4	10	484	0	0	0	0	0	0	0	6	3600	1250	500	7938.4	31754	7938.4
Sar	Kai	Hamid	Be	Sar		2500	0.25	200	200	100	35	1050	0	0	0	0	250	0	20	304.4	10	484	0	0	0	0	0	0	0	6	3600	1250	500	7938.4	31754	7938.4
Su	Vir	Khoha	Be	Sar		1575	0.158	126	126	63	20	600	0	100	0	0	157.5	0	15	228.3	8	387	0	0	0	0	0	0	0	6	2268	788	315	5158.5	32752	8188.1
Santos	Bantha	Ne	Raj			2500	0.25	200	200	100	35	1050	0	0	0	0	250	0	20	304.4	10	484	0	0	0	0	0	0	0	6	3600	1250	500	7938.4	31754	7938.4
Sunil	Daulat	Ne	Dec			2500	0.25	200	200	100	35	1050	0	0	0	0	250	0	20	304.4	10	484	0	0	0	0	0	0	0	6	3600	1250	500	7938.4	31754	7938.4
Bimal	Daulat	Ne	Dec			2500	0.25	200	200	100	35	1050	0	0	0	0	250	0	20	304.4	12	581	0	0	0	0	0	0	0	6	3600	1250	500	8035.2	32141	8035.2
Bhuta	Daulat	Ne	Dec			5000	0.5	400	400	200	75	2250	0	0	0	0	500	0	46	700.1	23	1113	0	0	0	0	0	0	0	6	7200	2500	1000	16263.3	32527	8131.7
Subat	Daulat	Ne	Dec			2500	0.25	200	200	100	35	1050	0	0	0	0	250	0	23	350.1	10	484	0	0	0	0	0	0	0	6	3600	1250	500	7984.06	31936	7984.1
Ram s	Daulat	Ne	Dec			2500	0.25	200	200	100	30	900	0	0	0	0	250	0	23	350.1	10	484	0	0	0	0	0	0	0	6	3600	1250	500	7834.06	31336	7834.1
						38821	3.88	3105.7	3106	1553	550	16500	0				3882		334	5083	168	8131										124650	5E+05		

Figure 5.5: Data on specific agricultural practices of female farmers

In contrast to Leonard's focus on value-free efficiency, the mundane capture of survey data from households created "*clamour*" too, via professional filtering. Capturing data converted the farmers' opportunity to talk about livelihoods and projects into a syntax. This syntax aggregated the semantic content (i.e. farmer responses) into question columns and rows. The columns indicated for example demographic details, income levels, crop types, land, and fertilisers). Thousands of rows represented farmer identities. Data was rendered legitimate and focal when put into cells; it was rendered peripheral when not stored in cells. This bifurcation od data into or out of cells unfolded in line with the earlier design of evaluation questions and expected responses. This editing of data made cells legible for later processing.

After talking with Chandan, it became clear that a wealth of other data was present during household interviews, and that while this data could be identified, it was considered undesirable or illegible. This extra data was evident in farmers' oral responses to survey questions, and included opinions, confusions, doubts, needs, concerns, and the views of other individuals who were present and responsive during household survey interactions.

This wealth of data did not fit into digital cells or question frames. For example, Chandan stated that farmers often had doubts about survey questions, and didn't fully know or use terms used in survey questions, such as "*savings*", "*incomes*" or "*expenditure*".

> *"We know the family size, landholdings but not exact figures for savings and expenditure ... because they are not very exact, they [farmer respondents] don't record their expenditure. As they sell their produce, they get the money and they finish the money in a day. They don't know how they have spent it." (Chandan)*

Household survey encounters frequently involved having multiple respondents in the room, family or neighbours, and thus multiple opinions and many responses to discrete survey questions. Chandan dismissed these as "*clamour*":

> *"... population density is high and there are a lot of disturbances, so when I am asking you a question there will be a few others, their neighbours all clamouring there, answering questions and disturbing and all that." (Chandan)*

The author proposed a set of 25 qualitative interviews with female farmers, using open-ended questions to understand this contextual data or clamour more closely. But Chandan and Leonard expressed reservations about qualitative interviews. They cited six reasons: firstly, mobile devices could not record audio; secondly, software needed difficult adaptation; thirdly, transcription from farmer dialects was arduous; fourthly, fieldworkers did not have capacity to manage open response interviews; fifthly, field bosses did not want volunteers wasting extra time in households; and sixthly, the partners were not sure how to analyse qualitative data, as shown in Leonard's comment below.

> *OK, ... so the analysing of the qualitative, is somewhat different from the quantitative. So, you know, erm ... yeah ... we have ... people for that. Not a problem. ... <PAUSE> ... 'P' [author] can also be of help ... I guess? (Leonard)*

Chandan described how the NGO used to record interviews, but stopped because farmers were "*not articulate enough*".

> *"What voice recording doesn't give us, we tried that, but one of the challenges we faced was that some people are not articulate enough to say what they wanted to say, so we lost a lot of data." (Chandan)*

Leonard and Chandan both acknowledged the need for deeper knowledge relating to farmer livelihoods, and much discussion revolved around using qualitative interview questions. For example, Leonard expressed a need to "*know more*" about farmers selling their produce, beyond metrics or discrete survey responses:

> "*… if you sell in the market, then what is the process? How do you sell? Do you, … do you yourself go there? I want him or her to kind of … say more on this.*" *(Leonard)*

However, the difficulties of doing qualitative evaluation described above trumped the potential benefits and the author's recommendation for revisiting qualitative interviews.

Chandan pointed to Rural India's existing qualitative impact data, exemplified below by an excerpt from the story of Devi, a farmer in Rajasthan, describing her participation in a Rural India project (Figure 5.6). This profile piece featured in a published report in 2013, along with four other similar profiles.

Case Story 6:

My name is ████ **Devi**. My husband's name is ████ Ram and I am a resident of ████ village of Alwar District in Rajasthan. There are 4 members in my family and the only livelihood source for us is agriculture. Even after continuous hard work, the economic condition of my family is very poor. One day, the project staffs from Farmers' Club Project ████ came to her house and informed her about the benefits of becoming part of the project. Initially, I was not convinced but observing the benefits received by other women of the community, I became a member of S████ji livelihood group.

I started following improved agriculture practices in Mustard in the area of 3,750 sq. m. under the technical guidance of HPPI. The practices involved proper seed treatment, application of

Figure 5.6: Example of qualitative profile from Rural India project report

The profile contains several indications in the text that demonstrate how Devi's original voice has been altered or mediated by the evaluators. Firstly, it was translated from Devi's local language into English by Rural India. Secondly, it switched from first person to third person mid text. Thirdly, it addressed project aims exclusively. Fourthly, it used technical or project language, such as "*livelihood source*", and "*observing the benefits*". Together these issues raise doubts about the authenticity of the text and its accuracy in portraying Devi's voice. In fact,

the text appears to have been edited for a before-after effect, to amplify farmer gratefulness and project impact.

Examples such as this show how data-capturing activities filtered data into two types. The first type was legitimate impact data: mobile, fast, efficient, mostly in cells, and legible to the evaluation machine. It showed what the partners needed to know, what answered their discrete question designs, and what could be used in follow-on activities.

The second type was illegible, clamour, waste, immobile, irrelevant, unruly, or inarticulate data in the eyes of the evaluators. It included doubts and uncertainty, ignorance of evaluator terms and concepts, data that did not fit into cells because another person in the house had offered the response, did not fit a question type, category or cell shape, or did not fit a before/after profile of a grateful beneficiary from a successful project. The bifurcation also filtered out data that was too challenging to capture, would offend bosses or required new digital recording software/hardware to be arranged. This bifurcated, illegitimate data included broader contexts of farmer lives, farmers own articulations, stories and voices, and comments from others present during survey interviews.

In dividing up the data this way, the NGO and philanthropy created evaluation silences. Good data was like mined ore; bad data like discarded residue, valueless, left in the field. Activities in this part of the evaluation cycle included:

1. capturing data from households using mobile tablets and surveys;
2. uploading data into an online database application;
3. filtering household data into survey app cells: legible, clear, definitive responses;
4. filtering household data out: doubts, inarticulate, clamour, disturbances;
5. filtering out irrelevant details, voices, contexts and editing in project successes;
6. eliding challenging methods that disrupted established evaluation mechanisms or narratives;
7. storing data for access/analysis in cells, rows, columns and sheets, or before/after profiles; and
8. transporting meaningful digital data and profile narratives to the funder in the UK, Imagine.

5.3.4 Late-cycle activities: Packaging data, bundling in expertise

Late-cycle evaluation activities packaged the mined data into more robust and meaningful narratives by adding expert forms of knowledge to the data. Whereas previous activities involved filtering, editing data in, or cutting data out, these activities involved adding expertise to the data, bundling in value-adding knowledges, using knowledge to package data, rather than data to make new knowledge, as in DIKW assumptions. Four kinds of expertise were bundled in: technology and data management; technical development discourse; scientific methods; and marketing. Chandan was also strategically aware of the lack of expertise in the wider sector.

In terms of technology and data management, Leonard stressed Imagine's strategic aims when talking about how impact data helped the philanthropy to create categories of aid effectiveness. These categories framed the design of the evaluation MIS. According to Leonard, this design process involved a "*roadmap*," evaluation "*templates*" and "*poverty alleviation verticals*".[26]

> *"Yes, that's the long-term roadmap, ... so what we would like to achieve is to create a template for each of the different verticals, like one of them being poverty alleviation or you know that sort of thing, and do other templates for other verticals so ... for the NGOs we would like to standardize the data collection procedures and make it easier to get quality data ... and it is in those templates where we can, so we can bounce ideas off you [Author], so we can get better templates, create it and so on ..." (Leonard)*

Signposted here was one of Imagine's reasons for collaborating on this research case, to secure advice for further developing their MIS designs. For example, bouncing ideas off the author helped in standardizing evaluation templates and defining verticals to structure the MIS design.

With Leonard explaining how the MIS would be marketed to other NGOs and agencies in the sector, "*for a small cost*", this bundling of impact data within technology expertise and resources shows how farmer data travelled not just to UK funding decision-makers in reports, but also how it entered other strategic considerations. Farmer data fed into impact narratives and distinct business strategies, prominently the MIS.

---

[26] "Verticals" is a commercial term used to differentiate product or service markets.

Alongside digital know-how, the partners demonstrated development expertise too. Development expertise was evident in terms used by the managers, in conversations, reports and on the philanthropy's website. For example, Leonard talked of "*poverty alleviation*", Chandan of "*capacity building*" and "*integrated development*". Phrases used in reports included "*scaling up projects*" and "*providing enablers*" for other NGOs. On their website, Imagine stressed their commitment to "*integrated development*", "*technological solutions*", and supporting local NGOs in terms of "*efficiency*", "*productivity*", "*ideas*", "*methods*" and "*impacts*". The partners used their development expertise to further package impact data as part of "*social and technological solutions*", "*lasting developmental results*", and "*transformative impacts*". Such terminology is a standard component in contemporary development discourse. Further excerpts from Imagine's website are given below.

> "*This approach addresses the social and economic challenges of a developing society by establishing a multi-level partnership between non-profit organizations, civil societies, social entrepreneurs and government bodies. It seeks to mobilize systemic transformations by integrating government policies to expedited financial, social and technological solutions and create an enabling environment to foster innovations, forge partnerships and build networks.*"

> "*Imagine catalyzes innovation, not only by providing funds but, by extending organizational support to enhance the efficiency and productivity of NGOs with new ideas, methods and models that can achieve lasting developmental results. It is also important that these innovative approaches should have a direct and transformative impact on the marginalized sections of the community and have the potential to scale up significantly in the region so that they can be replicated elsewhere.*" (Imagine website, Programs page, 21/2/16)

Together with technical development discourse and data/technology expertise, the partners also made sure that scientific vocabulary and methods were featured in their impact knowledge products. Scientific discourse is evident in the term a "*model district*" mentioned above, or references to "*solid evidence on the ground*," having "*our analysis in place*", and a "*hydro-electric*" engineering metaphor, used by Chandan.

The model district phrase was used to compare villages and use the exemplary ones to highlight where positive results and impacts had been found. These results depended on having good

data and evidence. The hydro-electric metaphor was deployed by Chandan frequently to explain how the partners' work was a dependable, scientific and mechanical process to produce favourable impacts. Scientific methods are also represented by the "*model district*" phrase and constituted by the collection of digital data on farmer livelihoods and organized in spreadsheets and cells for measurement purposes, as discussed in the previous sections. In such ways, scientific methods and measurement rationales are added during the construction of impact messages and narratives. Chandan explains:

> *"We have taken the district as a model district. We want to showcase this, that the work has changed the district vis-à-vis the other neighbouring districts. We need to present it to prospective funders, at fundraising events, so this is a kind of pitch that we are trying to do on the basis of solid evidence on the ground." (Chandan)*

> *"And we understand that if you put an x in, we expect about 10x or 15x at the other end. I also use, I often use, the analogy of hydro-electrics, so if you put x pressure at one end, the other end should give you about 10x or 15, or some multiple of x pressure at the other end." (Chandan)*

In these ways, scientific techniques, methods and discourses were incorporated into impact evaluation messages and narratives. As insertion of such explicit scientific vocabulary was less evident earlier in the evaluation cycle, it can be considered as important for bundling-in, value-adding, as part of late cycle knowledge packaging to boost impact message legitimacy.

Marketing expertise and phrases were also bundled into the partners' late-cycle impact activities. Rural India and Imagine's evaluation performance delivered impact messages able to support marketing activities. Impact was an issue at the level of specific projects, but also at the wider level of impact as an organisational capacity. Project impacts and processes were combined with various organisational expertise to form the wider narrative on impact produced by Rural India and Imagine within the Indian aid sector.

Marketing terms and phrases used by the partners in conference calls, reports, and website pages included: "leveraging expect 10× or 15× at the other end", "our marketing story". "our pitch", "a strong pitch", "put together data in a nice document", "to go fundraising", "to showcase impacts", "leverage resources", and "the need to ask others to pitch in, but for others to pitch in we need a strong marketing story".

This marketing language appears in one sense a mundane and professional part of business. Yet it is this normal, professional bundling of marketing concepts into the partners' notions of, understanding of, and evaluation of impact that is itself significant.

Finally, it is worth noting that the bundling-in of expertise in digital technology and data, development, scientific techniques, and marketing was accompanied by a dismissal of the capabilities of other NGOs in the Indian aid sector. Other NGO were "*slow"*, **"*incompetent"*** in digital processes such as data management, "*inarticulate***"**, and in need of "*smarter*" workflows.

> "*NGOs in India, most of them, the level … they are not very competent with technology, or smarter ways of capturing data … bad data logistics cause project delays.*" (Leonard)

> "*Now generally, ask any NGO in India, most NGOs in India, almost all of them, are well meaning, honest, with high level of individuals, but if you ask them what they are doing, the chance is that probably they don't know, they are not articulate.*" (Chandan)

To summarise, there were numerous activities involved in late-cycle processing or packaging of data into impact messages, such as bundling expertise in with the legitimate data captured and bifurcated earlier. Through the evaluation cycle, data was incrementally processed, packaged, and made into strategic knowledge. Packaging and bundling activities included:

1. processing or packaging data to make it into legitimate impact knowledge;
2. adding data management or technology know-how to impact data to make impact knowledge e.g. MIS templates and verticals;
3. adding development expertise to impact data to make impact knowledge;
4. adding scientific expertise to impact data to make impact knowledge;
5. adding marketing expertise to impact data to make impact knowledge; and
6. downplaying other NGOs ways of making / understanding impact.

5.3.5 Late-cycle activities: Pitching impact data/knowledge

The final set of activities evidenced in the case extends the discussion of marketing in the previous section on packaging and bundling-in knowledge. These activities are worth specific attention as they indicate how even late-cycle activities backwash through the whole evaluation cycle. These late-cycle activities revolve around exchanging or pitching impact to potential audiences and investors. The need to pitch impact data/knowledge and exchange it with other organisations, for example at fundraising events, is a key dynamic evident in, and influencing, all the previous impact evaluation activities.

In scientific evaluation, data and knowledge are constructed to find truth. In participatory evaluation, the aims foreground equality and inclusion. In DIKW-inspired data/knowledge models, the aim is to inform decision-making. However, in the Rural India case impact data/knowledge is required for pitching and exchanging for investment from donors. This is shown in Chandan's comments below relating to "*pitches*" and "*solid evidence*":

> *"We have taken the district as a model district. We want to showcase this, that the work has changed the district vis-à-vis the other neighbouring districts. We need to present it to prospective funders, at fundraising events, so this is a kind of pitch that we are trying to do on the basis of solid evidence on the ground."*

> *"So, we want to say, hey guys, with the x amount of money that you have put in we have been able to leverage 100 x."*

> *"... we have moved the community from point x to point y... we chose those villages that could demonstrate greatest impact in terms of low starting point."*

> *"So, this is a kind of pitch that we are trying to do on the basis of solid evidence on the ground."*

> *"And how do we talk to them, what is our pitch? So, to make a very strong pitch we need to have our analysis in place, ... which is often not happening. So when the NGO workers or volunteers give the data to their management, and their management collects, analyses, puts it together in a nice document and puts this together or publishes for Delhi or London or wherever, and we try to talk to different co-funders, co-investors, co-partners, and tell them that look - we need to go fundraising for this kind of program first ..." (Chandan)*

Data here is constructed to support the pitch for use at fundraising events to convince investors to fund the partnership and projects. In Figure 5.7 an excerpt from a group call shows this pitching process as a driver of evaluation activities across the cycle.

> **Chandan (Rural India Director):** So the contribution of this app is not only for a mechanical way of getting the data stored, but also to understand why this is being done. So I would say in response to your point, I would say that we'll be confident of the capacity building of the volunteers, but this would not be as much as we would like for the researchers for example.
>
> **P (Researcher):** This is a really good conversation. Let me just rephrase what we are doing again. If I am understanding you correctly, we would use the app here in the first kind of cluster or function, about collecting data in the field. And the second would be to support NGO offices in submitting their own interpretations of that data to a single repository for yourself, and for the London offices of Imagine to work with that data. So there are two inputs. One, from the beneficiaries of projects, through their volunteers or data collectors, and two, a mediating role for NGO management and office workers to apply their interpretation about what's happening, what this data shows us, like in reports.
>
> **Chandan:** Yes, and a third one, I would also like to mention and bounce this off you, is why we are doing this, is we see a roadmap, as I told you, leveraging of resources is very close to our philosophy. And we understand that if you put an x in, we expect about 10x or 15x at the other end. I also use, I often use the analogy of hydro-electrics, so if you put x pressure at one end, the other end should give you about 10x or 15, or some multiple of x pressure at the other end. So what it means for what we are talking now is that if we use this properly we will be able to have all the answers required. So, so say for example grants for a typical program like this, needs probably $100 and we have $10, and we ask others to pitch in, but for others to pitch in we need a strong marketing story. And how do we talk to them, what is our pitch? To have a very strong pitch, we need to have our analysis complete.

Figure 5.7: Transcript excerpt showing pitching activities

The marketing phrases quoted earlier (e.g. "*put together data in a nice document*", "*go fundraising*", creating a "*strong marketing story*" etc) are important in the passage above, as they are key phrases that signal value; in other words, how impact messages are instrumental in pitching activities. The bundling in of expertise previously makes sense when understood as critical architecture for impact pitching.

If the notion of pitching involved exchange and sale, accompanying it was leveraging data, which points to deployment of conceptual resources to strengthen the legitimacy of the pitch In the passage above Chandan deploys the hydro-electrics metaphor to strengthen the appeal of an impact pitch to potential investors, and here he explicitly uses the term "*leveraging*".

"I would also like to mention you know, as I told you, in our roadmap, that leveraging of resources is very close to our philosophy." (Chandan)

In this way, the pitch metaphor is a good shorthand for understanding how the whole process of Rural India and Imagine's evaluation operation functions. Pitching and leveraging speaks to all we have discussed in this chapter, from the need-to-know, to the assembling, organising, and bifurcating data and into cells and silences, culminating in packaged products for investor relations.

Despite Rural India's difficulties with more contextual or qualitative evaluations, the case suggests they had become relative experts at understanding the pragmatic elements of the evaluation cycle, and the pitching required for sustainable funding and legitimacy. This expertise was achieved through their close alignment to Imagine over many years, and their adoption of Imagine's impact mechanisms. This expertise was not in technical evaluation, but in making pitches. If we include the MIS development, the expertise and the cycles of impact knowledge construction, these lead to new organisational strategies in aid markets, particularly in the case of Rural India and Imagine, opportunities of a digital nature.

Activities involved in leveraging and pitching included:

1. packaging knowledge for exchange, producing sales pitches or narratives;
2. producing reports for funders and marketing copy to use at events, on websites etc;
3. pitching impact data/knowledge products, at fundraising events; and
4. using impact data, knowledge or narratives in new organisational strategies (e.g. MIS) or other digital innovations in the aid sector.

5.4   Analysis section

The analysis section uses existing CHAT concepts of activity systems and activity contradictions, as well as the novel CHAT concept of temporal activity chains to respond to the three research questions. Section 5.4.1 analyses the case evidence using CHAT activity systems and contradictions. A primary contradiction is found relating to not one, but two kinds of impact: one humble-form and one marketing-form. Section 5.4.2 utilises the novel concept of CHAT temporal activity chains in order to describe the power and practice relations silenced

in technical evaluation discourse and DIKW-inspired knowledge management models. Section 5.4.3 reviews what was learned from the case about reinserting power and practice into evaluation activities.

5.4.1 Activity systems and two contradictory impacts

Research Question 1 asked the following: how is development NGO impact data and knowledge constructed in practice? The concerns of impact evaluation data and knowledge construction in practice appear wider than those contained in TIEK or DIKW-inspired decision support. Practice involves more than methods and results, and more than inputs and outputs for presentation to decision-makers. In the case, data was only a part of the many foundations for making impact knowledge: sector demands framed required data, much data was discarded in households, expert knowledge was packaged in with data, and organisational strategy also moulded final knowledge products.

At Rural India, impact construction required motivation (e.g. sector demands, or Vijay's need to account for money spent); methods (e.g. surveys, collection of discrete data, YES/NO question techniques); assumptions about measuring, bifurcating data and pitching impact knowledge; tools and technologies (e.g. spreadsheets, survey software, tablets, desktop PCs); hiring and training of staff and volunteers; collaboration between Imagine and Rural India; authority lines; and organisational strategies. These were not incidentals or contingencies hanging off a more real evaluation model. Their assemblage constituted the practices and power inherent in a cycle of evaluation activities. Therefore, a first response from a CHAT perspective is that evaluations are constituted by many activities evident in practice. These activities, not a model or a prescription (technical or moral), manifest the evaluation process.

A second CHAT response to the case findings is that there is not one, but two forms of impact involved. These two forms conflict, generating systemic contradictions. The first kind of impact is not represented in documents, but is part of farmer interactions and livelihoods in the case. It is difficult to capture, and thus unclear to the evaluators. It could at times be considered as "*clamour*" or "*not articulate*". It is orally shared, uncertain, undocumented, and not data-centric. It is part of stakeholder exchanges between those relatively close to a site of supposed change, who may be directly involved with or have their livelihoods affected by local changes or interventions. A good term for this kind of impact, and in contrast with the one that follows,

is "humble impact" – personally shared, materially experienced, locally circulated, and often uncertain, partial, locally situated and negotiated by stakeholders (Blackler, 1995). This is Impact-1.

The second kind of impact in this case is impact as designed, captured, recorded, stored, analysed, packaged and pitched – impact as predominantly digital representations, or in its complete state, as "*knowledge products*" (Mosse, 2004a: 77) to be exchanged. These are produced instrumentally, for circulation within aid chains, markets, and bureaucracies, in response to sector demands (Wallace et al, 2006; Mosse, 2004a; Quarles van Ufford, 1988). They are produced for professionals trained in development or evaluation discourses, and require expert terminologies of different kinds. They align with development and evaluation prescriptions, and increasingly with marketing and digital strategies. In development 2.0, such digitised impact representations are made mobile, rationalised and transported globally between local context and expert contexts. They are global representations (Avgerou, 2002: 77) used for exchange with other organisations, decision-makers, and groups in formal aid markets and bureaucracies. Mobile impacts are strategically constructed and mobilised for exchange, and as in this case, often exchanged for funding revenue (Hayes & Westrup, 2014: 28). They can also attract legitimacy and status from other actors in the market and bureaucracies, promoting views of the producer as a reputable and dependable partner. This is "Impact-2".

Figure 5.8 below shows the activity system and contradictions evident between humble impact or Impact-1, and impact marketing or Impact-2. The diagram shows how tools (e.g. data cells, spreadsheets), rules (e.g. for designing questions, writing reports), and divisions of labour (e.g. philanthropy and NGO as designers, analysts, writers, marketers, or farmer's as sources of raw data), limit the object of knowing impact, and how they facilitate expert construction of impact marketing narratives as outcomes or products to satisfy sector demands. The activity system bifurcates impact data/knowledge, rendering humble impact (clamour, doubt, farmer views, etc) as illegible to the impact machine, and thus it and discarding it in the field. These silences are not shown, because this activity system, at Rural India and Imagine, had already become an instrumental machine for producing outcome representations for sales pitches.

In CHAT, these two impacts form a primary contradiction, based on the exchange value of representational impact to meet sector demands and secure funds or legitimacy. This contradicted the use value of humble impact in farmers' daily lives; value found in sharing,

complaining, finding solutions, building local relations - "*clamour*" from Chandan's perspective. Secondary contradictions between activity elements (roles and object; labour and object; tools and object) are included in Figure 5.8.

Activity System for Imagine and Rural India Evaluation Machine

Primary Contradiction: Knowing impact marketing contradicts knowing humble/local impact

Tools

Secondary Contradiction 1: Tools to Object
Tools limited data to expert analysis and marketing. Could not record farmer contexts or doubts in mobile app or spreadsheet

Subjects

Object: To "know" impact

Outcome:
Representations for impact marketing, "sales pitch" to secure funding, and demonstrate expertise

Transformation:
Transforming data into impact marketing knowledge

Rules & Norms   **Community**   **Division of Labour**

Secondary Contradiction 3: Labour to Object
Philanthropy and NGO control evaluation, collect / analyse data, produce reports and fundraise
Farmer's supply "raw" survey data
No labour to know farmer contexts, views, definitions. Farmers had no role in design, reporting, fund raising

Secondary Contradiction 2: Rules to Object
Rules existed for digitising and analysing data, writing reports, rounding figures
Rules also existed for removing "clamour", ignoring doubts about income/expenditure, or "inarticulate" farmer responses

Figure 5.8: Activity system and contradictions in the Rural India evaluation process

Again, tools (e.g. data cells, survey methods), rules and norms (data management, removing clamour), and divisions of labour (NGO and philanthropy evaluation control, farmer as raw data supply alone) all contradicted the potential object of knowing farmer-defined impact or impact one. The elements limited knowing to impact two, impact for marketing. Farmer voices, contexts, uncertainties, and participation were professionally and mundanely taken out of outcomes. Humble impact was submerged or masked (Engeström, 2008: 36–42).

It is noteworthy that such tensions were not considered a problem for the Rural India or Imagine management, because marketing impact had become an expert-normalised process, a problem of professional efficiencies, technical aptitude, and effective results. Therefore, in response to Research Question 1 on how impact data and knowledge are constructed, the CHAT analysis suggests impact evaluation data/knowledge is constructed in an evaluation machine assembled

over years of practice between partners, during which time partners such as Rural India and Imagine can become expert at harvesting data and constructing knowledge products to support management and marketing activities. The machine and construction process are not intended to understand, amplify or mobilise humble impact from farmer lifeworlds.

5.4.2 Temporal activity chains, submerging and elevating

Research Question 2 asked the following: how are power relations generated in practice during impact evaluation data/knowledge construction? A view of the temporal sequence of activities illustrates a relatively stable chain in this case, wherein impact data and knowledge products are edited and mediated, and power relations are simultaneously generated. For example, the findings showed that managing NGO staff requires specific early-cycle training activities, involving divisions of labour and rules, in distinction to later activities, such as data capture, storage, or analysis. The temporal view also highlights points of agency where submerging and elevating occur. These points are opportunities for change, contestation and learning, albeit with the proviso that systemic change is not guaranteed because power is diffused along the chain.

CHAT has traditionally modelled hierarchies of activities, actions and operations (Leontyev, 1978; Engeström, 1987), and focused on bounded sites such as schools or workplaces. However, Engeström (2009: 9–10) has questioned how activity systems model larger formations in society, because today these places "*are bombarded by interventions from all kinds of outside agents (e.g., consultants, administrators, customers, competitors, partners, politicians)."* NGO evaluation data/knowledge work is subject to such influences. Engeström puts forward runaway objects, knotworking, trails, mycorrhizae, and wildfire activities as new concepts for understanding the diversity of large-scale social activities today. The notion of temporally configured activity chains (Figure 5.9) follows these concepts for understanding diffuse phenomena and fits descriptions of the aid chain, in which aid data and reports flow between international stakeholders (Wallace et al., 2006: 13, 166).

> *"Forms, records, genres of email and other forms of documentation travel across activity systems and make trails that change the landscape. This is directly relevant for our attempts to understand current historical transformations in the organisation of human activities", (Engeström, 2009: 8–9).*

The Rural India activity chain incorporated diverse activities, from the early-cycle need-to-know impact and the later exchange of impact representations in funding markets. From impact demands to impact supplies, data and knowledge travel through and mediate activities, revealing a deeper, more fluid, and power-laden production process.

Furthermore, understanding impact supply and demand means recognising how evaluation mechanisms that precede data capture activities invalidate the "raw data" claims of naïve positivist models, exemplified by DIKW (e.g. Earl, 1994) and assumed in many technical evaluation methods. Impact knowledge in this case cannot be explained by TIEK and raw data alone. Various managerial, design and methodological assumptions, filtration processes, expert packaging, and market pitching activities also contribute to impact data sets and knowledge products.

At Rural India, diverse sector demands led to a data/knowledge supply chain stretched out over time and space, in which impact data/knowledge flowed and changed (Figure 5.9).
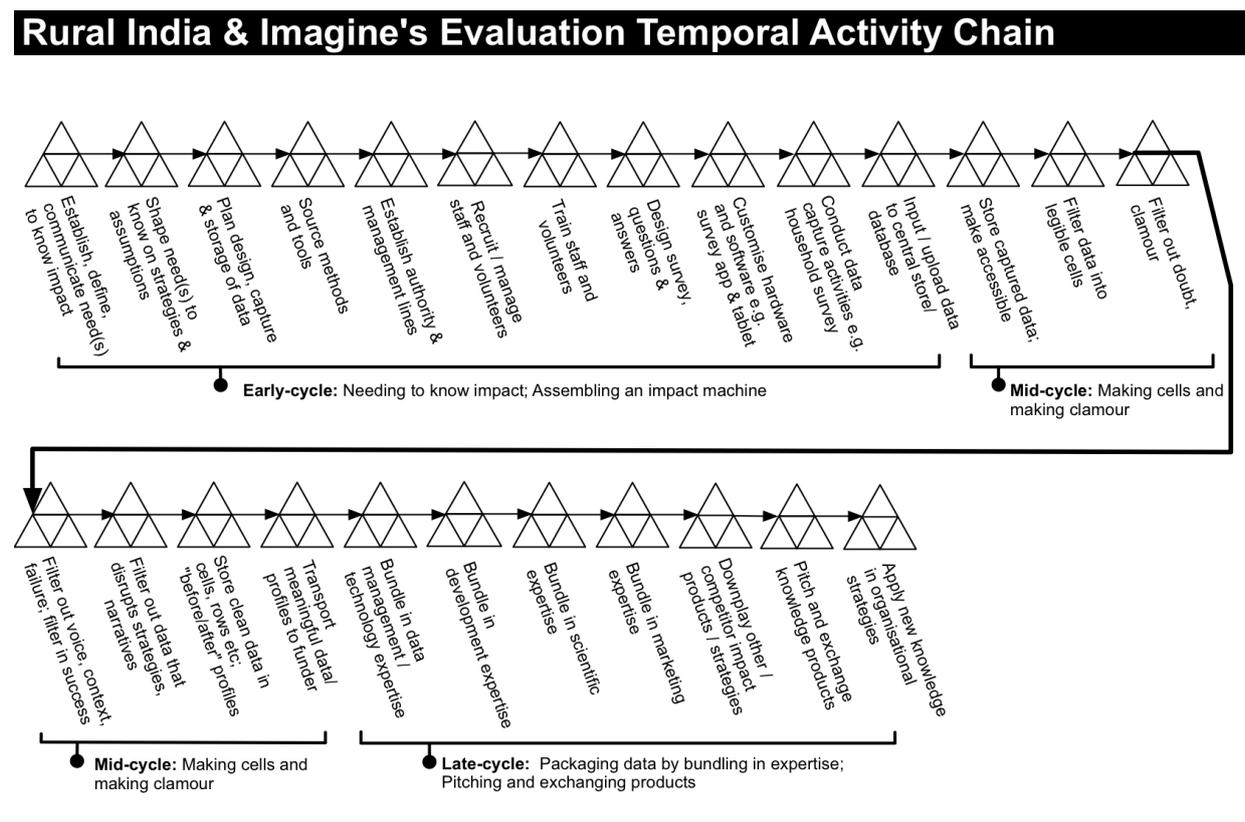
Figure 5.9: Evaluation illustrated as a temporal chain of activities

The temporal chain provides a tool for understanding power relations beyond technical evaluation or DIKW models of knowledge construction, beyond the linear rooting of impact knowledge to raw data facilitated by modern ICTs. At Rural India, bundled knowledge was not data driven. Technologies filtered or bifurcated impact, discarding some data in the field, storing and speedily transporting other data into digital systems. In Karanasios and Allen's terms (2013: 300), Rural India "*absorbed*" object of activity to market impact products provided to them by their funder Imagine, along with their technologies, rules and techniques. At the same time, the NGO benefited through funding, learning to become an expert evaluator, and gaining legitimacy as a capable deliverer of impact. By the time of this case research in 2013-2014 the partners' activities had coalesced into a stable data/knowledge supply chain.

In the supply chain, consisting of sequenced albeit often overlapping activities, representations of impact were "*redesigned to meet the demands*" (Miettinen et al., 2012: 11) of the aid sector. Such demands included selling impact, illustrating capacity, securing funding, meeting targets, and using modern ICTs.

What is important in this chain, and which speaks most directly to research question two, is how power inequalities are generated in the submergence or elevation of different pieces of data or knowledge. In other words, how pieces were kept, discarded, edited or amplified along the chain. Chandan's descriptions of farmer clamour and farmer inability to articulate their needs were normalised in relation to Chandan's own expertise. Farmer participation and voices were submerged (Figure 5.10).

Rural India & Imagine's Submerging & Elevating on the chain
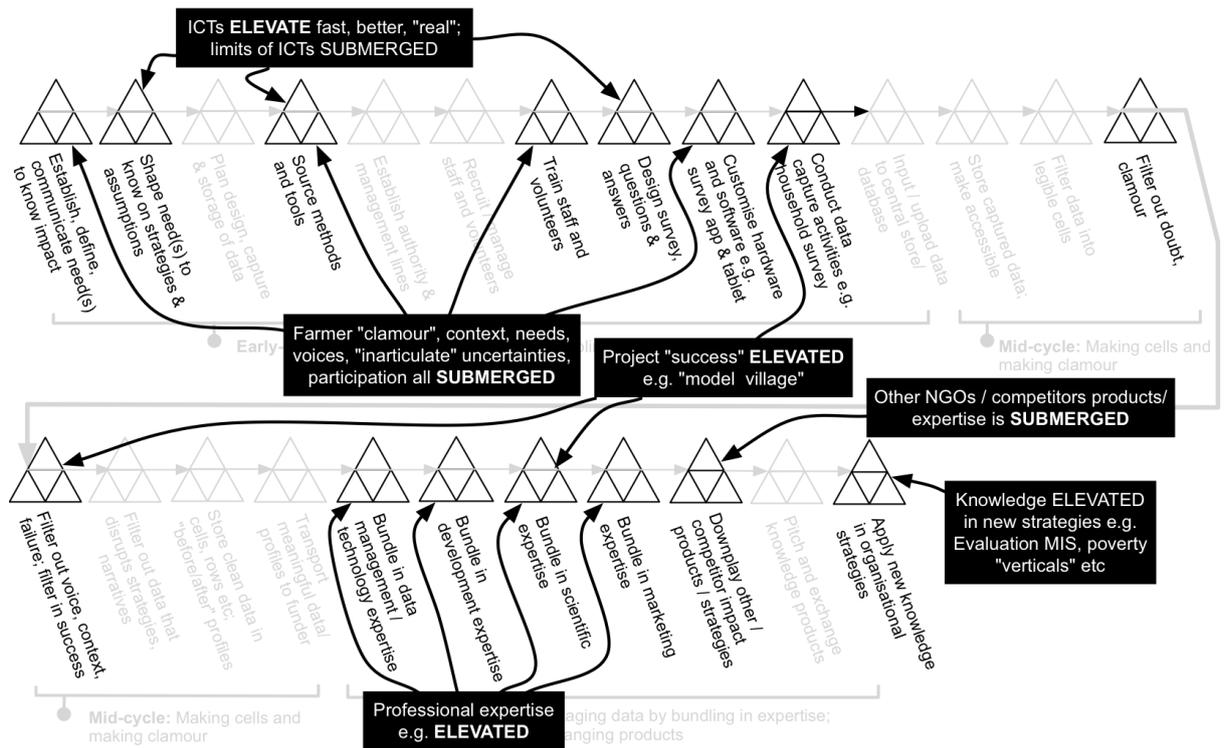
Figure 5.10: Submerging and elevating along the temporal activity chain

Complex contexts, clamour and voices seen by evaluators as inarticulate, or, in our terms, impact one, cannot be pre-designed, captured and stored by professionals in spreadsheet cells. Thus, they were discarded. However, data/knowledge that supported partner data management practices, marketing aspirations and evaluation forms of expertise were elevated. Leonard's need for fast, real data was a question of professional efficiency for him, not related to the lack or otherwise of participation by farmers.

Some knowledge was elevated and thus legitimated as part of exchangeable rationalised discourses or claims to expertise and efficiency (Hayes & Westrup, 2014). These elevations supported impact marketing and included model village successes, female farmer narratives of gratefulness, and cells which could be organised to show progressive changes. Other knowledge was submerged or subjugated in this process (Avgerou, 2002: 77), such as farmer needs, doubts about income and expenditure, group responses to survey questions, or aspects of farmer lives and lifeworlds that were illegible to evaluators. By elevating different kinds of expertise, certainty, and success, contextual knowledge, doubts about data, and farmer

participation in framing problems and changes were submerged along the temporal activity chain.

The problem of submerging and elevating mediations, and their effects on power and participation make critically important Blackler's (2011: 733) call for CHAT researchers to problematise structures beyond local interventions and confront clients with such problems. In this case, submerging and elevating occurred through diffuse mechanisms, whereby managers called participant voices inarticulate, whereby data cells ejected peripheral context, and whereby executives on the other side of the world assumed that impacts must be objectively measured using Excel sheets and business tools.

5.4.3 Re-embedding power and practice in evaluation

Research Question 3 asked the following: how can views of power and practice be re-embedded into impact evaluation data and knowledge construction? Initially, this meant asking if CHAT was able to surface evidence of power and practice in this case, and then, if participants reflected and acted upon power inequalities. The answers here are yes and no respectively.

Firstly, CHAT was able to articulate power and practice through activities, contradictions, temporal chains, and the submerging and elevating of issues in the production of impact knowledge. In CHAT, methods, data, information, knowledge, evidence, and results have more diverse genesis than in TIEK or DIKW. Therefore, CHAT is able to respond to concerns about the politics of evaluation (Norris, 2015: 136; Picciotto, 2015: 152–153), the politics of evidence (Eyben et al., 2015) and the politics of knowledge production (Gardner & Lewis, 2015: 179). The two impacts and the systemic contradictions all discussed earlier are part and parcel of a data/knowledge-intensive development 2.0 in which NGOs compete for funding and are under pressure to submerge or elevate different forms of data/knowledge. In this case, CHAT articulated these demands and the supply chain as activity systems and temporal activity chains. This contrasts with TIEK and DIKW-related models, both of which elide the diverse practices and power inequalities of knowledge production, instead focusing on inputs, outputs, efficiencies, pragmatic decision support, and privileged expert narratives or measurements. Thus, CHAT successfully renders power and practice in accounts of development evaluation data/knowledge construction.

Secondly, however, it was difficult for Rural India or Imagine to see beyond results, reflect, or act on the power relations generated by their evaluation machine. They were reluctant to adjust their evaluation practices to incorporate open response interviews, or elicit farmer's experiences beyond Yes/No questions. It was difficult for the partners to construct narratives other than success messaging, or to acknowledge and reflect on what appeared to them as "*clamour*". In this sense two of the aims of the research were unsuccessful: to encourage more multi-voiced evaluation using qualitative, more open response interviews, and to encourage reflection on evaluation power dynamics. The NGO director lamented lost data and inarticulate farmers when using qualitative interviews in the past, and cited numerous reasons for not revisiting qualitative ways of understanding farmer contexts. The philanthropic foundation's head of projects was unsure about resources for qualitative interviews or analysis. Both acknowledged problems (wanting to hear more about how farmers sell goods at market, or wanting farmers to be more articulate), but they were unable to go further than occasional comments or questions to find resolutions to these problems. Unfortunately, the philanthropy managers could not bring themselves and their evaluation machine to hear more beyond sales pitches and strategic efficiencies.

On further reflection, the CHAT analysis suggests farmer voices, clamour and more contextual narratives conflicted with the aid partners' need to show success, expertise, and certainty. More contextual evaluation would have led to questions about the lack of farmer participation and limited data validity, thus damaging impact marketing narratives. Given the two-impact contradiction, and the triple needs to digitise, transport, and pitch impact for funding markets, it was not surprising that the partners avoided richer contexts and issues of power.

In summary, CHAT supported a greater sensitivity to power and practice. However, the managers at Imagine and Rural India found elevating such sensitivities from occasional comments into reflective practices both challenging and strategically threatening.

5.5   Summary of implications

The case empirics and analysis raise various points related to development impact evaluation processes, data/knowledge construction, and the transparency of power and practice in evaluation models and results. Three implications are important.

The first implication is whether the CHAT analysis of impact contradictions and temporally diffused power dynamics in funder-NGO partnerships is unique to the case, or if the analysis resonates with broader development 2.0 processes. For example, does ICT2.0 deliver progress that is increasingly responsive to poor people's demands, as Heeks (2008: 33) states? And are data-intensive technologies "*on balance*" both relevant and beneficial to developing nations (Walsham & Sahay, 2006: 7)? Answering such questions implicates the specific data/knowledge supply and demand chains in development 2.0-related initiatives, between diverse kinds of organisations (e.g. investors, governments, IT vendors, researchers and consultants), and diverse kinds of data/knowledge intensive development processes (e.g. planning, policy development, e-development, data analytics/big data, and new innovations, such as blockchain for development [27]. The take-away problem from the Rural India analysis is: how damaging are the unequal power dynamics generated by data/knowledge intensity?

This implication points beyond the need for broad approaches and analysis (Brigham & Hayes, 2013: 127) to understand these shifts and success demonstrations. The case elevates: a need for further thinking on power and practice in data/knowledge intensive processes; a need to make the diverse organisations and processes involved more visible; and the diffused relations that generate power inequalities and push evaluation actors to ignore practice. CHAT offers conceptual tools for understanding these relations, via activity systems, contradictions, and temporal chains. However, CHAT has not always interrogated diffused networks of agents and activities beyond a small number of localised actors.

The second implication is whether and how approaches to social practice and power such as CHAT might lead to ideas, tools or methods that can be pragmatic, palatable, and useful for adoption and adaptation by development agents. In the case, the partners were open to advice on improving their annual evaluations. CHAT delivered an analysis of activity systems and temporal chains that illustrated distributed data, knowledge, and power dynamics. However, the partners found it difficult to reflect or revisit qualitative evaluation. The managers did not acknowledge problems regarding unequal power relations in their professional work, bifurcating data, packaging or pitching impact knowledge. Business needs, expert terminology, statistical metrics, evaluation concepts (e.g. before/after profiles), and digital tools supported

---

[27] See Hernandez (2017) for a recent discussion of blockchain for development.

and were elements prioritised by the managers for their evaluation machine. Reflection on power and practice was challenging and risky for them.

If sector demands, TIEK, and the DIKW legacy combine to normalise and silence power and practice, then how can more sensitive approaches be re-embedded into impact evaluation and development 2.0 more broadly? In development literature, Gardner & Lewis (2015: 180–181) implore anthropologists and aid agencies to collectively engage and critique together. Guijt (2015: 207) asks development practitioners to acknowledge the messy politics of their work or find "*space and time for more appropriate protocols and methods.*" In reviewing practice theories, Nicolini (2012: 240–241) concludes that testing practice-based toolkits in fieldwork is required. Engeström, Virkkunen, Helle, Pihlaja, & Poikela (1996) and others (e.g. Virkkunen & Newnham, 2013: 24–25) facilitate CHAT change laboratories, bringing stakeholders together to explore potential changes. These are some ways forward.

However, in this case, as in global development generally, stakeholders are geographically dispersed and inhabit different positions across data/knowledge supply chains. Participatory forms of ICT4D, KM4D, and evaluation emphasise local communities and local knowledges as a participatory response to knowing more about the local context. Yet being sensitive to power/data/knowledge relations means acknowledging wider activity chains across development 2.0, not just targeting locales of data capture, but also locales of governing, design, editing, packaging and pitching, and the points of submergence or elevation. Thus, the case has no clear antidote to the diffused generation of power inequalities in evaluation, but it does suggest ideas, methods and tools that can be adopted or adapted for use across multiple sites and professional concerns in and around evaluation. Such a multisite response can amplify the problem of silent power and practice more tangibly across more agencies than a single-site intervention to boost scientific or participatory results at the aid-target beneficiary locale alone.

The third case implication concerns the prescriptive knowledge and expert models targeting development and evaluation professionals, ICT4D and KM4D researchers, and those undertaking development 2.0 data/knowledge intensive work. This implication relates to how the data/knowledge intensive discourse and network is built, constituted, made mobile and made unequal. This involves impact evaluation models and prescriptions, data/knowledge management methods, e-development information systems, and all the sector demands for data/knowledge products which satisfy targets, populate information systems, mitigate risks, and communicate success etc. These prescriptions, products, information systems and

networks have temporally sequenced activities that generate power relations, submerge voices, and elevate marketing narratives. They dismiss clamour, generate cells, spreadsheets and profiles, and bolster the confidence to know what is really happening at target sites and in target communities in faraway places. IS and KM models do not track power or practice in these relations, nor the submerging and elevating of interests and experiences. However, approaches to social practice, such as CHAT, can and do elevate such diffused issues of practice and power.

This adds to arguments that suggest researchers and practitioners need not ignore power in mundane data/knowledge work, managerial controls (e.g. Bernardi & De Chiara, 2011: 37–38), or local knowledge generation (Walsham & Sahay, 2006: 11; Thompson, 2002). However, the danger in this case and more broadly in development 2.0 is that data/knowledge intensive processes such as impact evaluation may be captured by vested interests, bureaucratic or neoliberal market forces (Picciotto, 2015: 152). This is especially noticeable where mundane power dynamics operate expertly and silently.

This third and most problematic implication, begs the question of how to respond in ways more comprehensive than occasional reflections or doubts that can be parked and ignored by staff, professionals or managers, as evidenced in the case. Blackler (2011: 732–733) advised CHAT researchers to take heed of Hardy & Clegg's (1996) observation that power is best theorised as itself constituting "*the medium of collective action*". Practitioners and researchers may benefit from Blackler's (2011: 733) advice on how they influence clients or participants in research or evaluation endeavours. Building on Blackler's advice, this case leads to a further focus on the following four considerations when promoting sensitivity to power and practice in networked evaluation data/knowledge construction.

1. the mundane, diffused power relations in data/knowledge activity chains;
2. the submerging of doubt, uncertainty, or participation, potentially viewed as old, unclear, out-of-scope, unmarketable, inarticulate, illegible, or peripheral;
3. the elevating of success, expertise, technical methods, and scientific rigour, considered as professional, certain, modern, unproblematic, and virtuous; and
4. how knowledge and "raw data" always have histories, etymologies and provenance through which their demand, supply, and power dynamics can be articulated and critiqued.

Blackler's advice and the list above are calls to acknowledge how we ourselves and our professional data/knowledge processes produce power relations and hide micro-activities and messy practice from our legitimate work spaces. The NGO director's "*sales pitch*" comments are crucial in this. The managers in this case did not know how to reflect on or entertain an alternative way of evaluating that could acknowledge and respond to power inequalities generated by their own professional work.

The next chapter features an NGO which lacked expertise in doing impact evaluation, in order to see if they too elide issues of power and practice. The later discussion chapter then responds to the power/data/knowledge dilemmas raised by the two cases.

6 NGO Case 2: HTSG Thailand, novice impact-makers

6.1 Introduction

This chapter describes the second NGO case, HTSG Thailand. The case responds to the three research questions, which focus on identifying impact evaluation activities (research question 1); the generation of power relations during data/knowledge construction (question 2); and, how NGOs understand and respond to such power relations (question 3). Section 6.1 introduces the chapter and section 6.2 recaps the case setting. Section 6.3 is an extensive section which explores the empirical data generated and identifies the activities involved in HTSG's impact evaluation work. Activities are described in early-, mid- and late- periods of the evaluation cycle.

Section 6.4 analyses the activities in two ways. Firstly, by using CHAT activity systems and contradictions to articulate two kinds of contradiction in the case. One contradiction concerns the two forms of impact, and a second contradiction relates to the learning and transformation of HTSG into a data/knowledge intensive organisation. Secondly, the novel concept of temporal activity chains is used to map the ways in which different power/data/knowledge relations are submerged or elevated during HTSG's evaluation work and transformation.

Section 6.5 summarises the implications derived from the case. Three important implications are described, relating first to how impact evaluations are performed, second to how small NGOs like HTSG are transforming, and third to how power/data/knowledge relations are shifting in contemporary development evaluation practices. These implications are insufficiently addressed in TIEK and DIKW legacy models of data/knowledge management for development.

6.2   Case setting

HTSG is a small, resource-limited NGO in northern Thailand, with around 25 staff. They provide healthcare and education services for minority or marginalised ethnic communities. HTSG runs multiple short-term projects, predominantly funded by international foundations, INGOs, the UN, and state donors such as the EU. Projects typically last 1 to 3 years. The chapter looks at how HTSG were learning to establish impact evaluation processes, with advice and field visits from the author. The research began in May 2014 and covers a period up to late 2016.

The first project, and most discussed below, is "Healthcare Project", which provides healthcare services for pregnant women, mothers with new-born infants, and children under 10 years old. It has reached over 3,000 beneficiaries since its first iteration in 2001. A team of half a dozen staff work on this project, engaging volunteers in target villages to provide front line services. Services offered include: promotion of dental hygiene, nutrition, breastfeeding, child development, and public healthcare services.

The second project in this study is "Migrant Project". It is a relatively new project started in 2016, resulting from a funding application co-written with the author and submitted in late 2015. This project has a small team of staff working on it and focuses on identifying and improving services for Burmese migrant workers and families in northern Thailand. Specific activities include helping migrants to gain access to and utilise public or private services, such as healthcare, education, policing, legal advice, human rights advice, employment, training and community development services. The project targets low income migrants, such as domestic workers and labourers who are unaware of or marginalised in terms of access to existing services.

Empirical excerpts, quotes and reports sampled in this research illustrate three key themes from the case. Firstly, HTSG as an organisation was a novice impact evaluator, lacking resources, skills, or a clear orientation towards performing impact evaluation. Secondly, HTSG was learning to perform evaluation by itself without a guiding funder or expert staff, while at the same time struggling and transforming the organisation in order for staff to become better evaluators and more adept data/knowledge producers. Thirdly, although HTSG was changing and incrementally aligning with impact evaluation prescriptions and governing processes, it was not exclusively seeking to pitch impact to funding markets. Staff were also applying their new data/knowledge intensive capabilities in diverse ways, with other partners.

6.3   Empirical section

This section describes field data from interactions with HTSG. The empirics show an NGO undergoing significant change, and becoming incrementally more data/knowledge intensive. This change is seen as positive and rarely problematised in much technical evaluation literature (e.g. 3ie, 2012; OECD DAC, 1991; Duflo & Kramer, 2005; Dart & Davies, 2005) nor in data/knowledge management literature (e.g. Davenport & Prusak, 1998; Dalkir, 2011; Williams, 2014), but has been questioned in more critical work (e.g. Wallace et al, 2006; Ilcan & Phillips, 2008; Hayes et al, 2017).

The following section starts with a description of how HTSG staff were learning multiple framings of what impact was to them, then moves on to how they were learning but also struggling to assemble an evaluation machine, bifurcate data, package data with knowledge, and pitch to partners.

6.3.1 Early-cycle activities: Needing to learn many impacts

HTSG did not have existing impact evaluation processes in mid 2014, at the start of the research. They had no funder or expert staff to act as a guide for performing evaluations, and no expertise, norms, technologies or techniques to migrate into evaluation work. However, they did express a need-to-know impact, in diverse ways. These needs, motivations or will to know impact preceded evaluation data capture and even projects themselves.

In May 2014, the author emailed a range of NGOs to enquire about the possibility of establishing collaborative research into how NGOs performed impact evaluations. HTSG emailed back in July expressing an interest. Khun, HTSG's director, asked Susan his administrator, to respond to the offer of collaboration.

> *"I have been asked to clarify that, as of yet, HTSG does not have a full-fledged system for evaluating project impact. However, we are very interested in learning more about how impact evaluations are made at other NGOs, particularly ones whose goals are similar to those of HTSG; we also hope to eventually develop and implement our own impact evaluation strategies." (Susan, HTSG administrator/intern)*

Further email exchanges lead to a first field trip to HTSG in August 2014. During this visit the author and Khun shared early morning car rides together from the hotel to HTSG's offices. After a few minutes of small talk on the first morning, Khun raised a pressing concern: the competitive funding environment HTSG were in.

> *"We are facing a challenging time for our future. Some of our programs may need to close, they are not funded for the long-term, but often for 1 or 2 years only. The funding environment is really difficult ..." (Khun, HTSG Director)*

During a meeting and coffee at HTSG's office, Khun added more. There were growing numbers of Thai, regional and international philanthropic organisations open to funding bids, and he mentioned a foundation run by a prominent Thai retailer that HTSG could bid to for funds.

> *Author: "Surely, you could make funding bids to them?"*
>
> *Khun: "Perhaps. ... Yes, we can approach and submit applications, but they often have their own focus, or differing aims, and maybe we can't always match them. We must think carefully before bidding ... choose which ones we can have most success with. If we are rejected, it can mean we waste a lot of time in applying" (Khun, August 2014)*

HTSG had multiple funders, including five that Susan identified: a US pharmaceutical multi-national with a charitable foundation in Thailand; three large INGO's based in three countries;

and the UN. HTSG also had a local donation system with branded boxes placed around their home city, but this generated little income.

Bidding for funds implied nuanced strategies and considerations, gauging HTSG and funder agendas, and assigning staff time to writing bids. Applying for grants and designing future planned impacts and evaluation processes meant HTSG had to respond flexibly but strategically to diverse funder interests and standards.

Knowing impact involved writing narratives of planned evaluation processes and target impacts into bids and reports. Susan was given the task of writing the English language versions of HTSG's funding applications and project reports. She was an intern from a prominent US Ivy League university, volunteering in Asia at HTSG for two years. HTSG had been taking advantage of this internship program since 2011 to boost their human resources and English communications capacity. Report-writing tasks required knowledge of various funder forms, templates and processes, financial reporting standards, collecting NGO certifications, appealing to funding agency interests, responding to specific funding streams, and matching funder priorities. Susan described this work as "*kinda confusing!*".

> "*There are different application forms, different project templates, different funders, different evaluation forms, and reports, so it's kinda confusing! There are clusters of reporting periods too, like December, July, half-year reports, annual reports y'know.*" (Susan)

Khun, a development sector manager with over 20 years of experience, added that it was easier when bids and reports were carefully planned and staggered at different times of the year. Susan nodded her head, but stressed that it was still "*really tough*" to prepare reports, evaluations and bids to meet requirements. Tough or not, HTSG's need-to-know impacts, and the communication of their aims, goals, outcomes and potential impacts was shaped by diverse and differing funder niche interests, development approaches, grant standards, templates, forms and digital application systems.

In August 2014, HTSG had no guide or overall strategy for knowing impacts, no systems or processes set up. The email excerpt below points to HTSG's need to learn professional evaluation, particularly the need for quantitative measuring of outputs or outcomes. Their goal of a holistic system was self-lead, without a guide or expert staff.

"We do try to give estimates of the numbers of beneficiaries benefiting from each project, but as we mentioned to you when you visited, we still have not developed a system of measuring impact. We are very keen to come up with an impact evaluation system that will allow us to tell donors, 'so and so many children have improved in this particular area, and we know this because...'. Hopefully we will be able to work on this with your kind help!" (Susan)

Without a dominant authority or guide, HTSG were learning but struggling to define their needs and learning requirements by themselves.

Even with funder blessings and positive previous evaluations, it was questionable whether knowing impact was sufficient for HTSG to successfully accept and run projects. This was illustrated when Khun described a case where an external evaluator had recommended a project extension and the funder had guaranteed finances. However, HTSG couldn't accept the extension.

"Despite them [INGO donor] wanting to continue, it was difficult work, very taxing getting through the mountains, and we didn't have the staff capacity ... to do the participation, reporting and evaluating." (Khun)

HTSG in this case knew their planned impacts and had both evaluator and funder agreements, but they needed to avoid risks and costs they couldn't afford. As we shall see in the next section, knowing impact is not enough; there are also many organisational challenges to overcome.

This section has illustrated the early-cycle need-to-know impact at HTSG, even before projects begin, and the diverse kinds of impact that HTSG needed to know. The diversity and uncertainty contrasts with more confident evaluators or NGOs harbouring more stable funding streams. Arguably, HTSG's uncertain funding climate, and flexible but risky "needs-to-know" are more prevalent in the NGO sector than stable funder/NGO relations. HTSG had no dominant guides, limited resources, and didn't know how to create the "*full-fledged system*" they desired. HTSG's need-to-know impact was linked to their competitive funding environment, self-regulated, reactionary, and unstable.

Activities involved in HTSG's multiple needs-to-know impact included:

1. sourcing impact evaluation advice (e.g. expertise, concepts, tools, processes, staff);

2. communicating impact plans in funding applications;

3. matching specific bids to specific funder standards, niche interests and online systems;

4. strategically managing funding opportunities, project needs and NGO sustainability; and

5. defining, imagining and building a "*full-fledged*" evaluation system to know impacts.

6.3.2 Early-cycle activities: Struggling to assemble a machine

This section describes the most extensive area of learning, struggle and transformation for HTSG with regard to impact. Various transformations were necessary for the organisation to be able to produce and market impact products. HTSG were struggling to identify and assemble resources, skills, expertise, evaluation samples and questions, assumptions, methods, people, time, data cells, databases, technologies such as tablet computers, and evaluation attitudes. Their ad hoc approach to building capacity was evident during the February 2015 field trip and workshops.

Despite this, their early ideas about improving evaluation capacity required training village volunteers and establishing an "*all-encompassing database system*". In an email from late 2014, Susan informed the author about the quantitative data that HTSG had on the Healthcare Project, and cited reasons why qualitative data had not been collected.

> *"Yes, we do keep data on villages/participants in the projects in Excel spreadsheets. However, at the moment the data mostly consist of quantitative information (e.g. child's name, weight, height, number of home visits, etc.) but do not really give qualitative information (such as details on individual families' issues and needs). This is mostly due to barriers presented by parenting volunteers' lack of writing and analytical skills. Ultimately, we are hoping to improve training for the parenting volunteers (this will hopefully happen in the next project year), so that their ability to write down detailed observations about each family and child that they visit can also improve; in this way, we hope to develop a more holistic, all-encompassing database system." (Susan)*

Susan's email outlines HTSG's assumptions about doing evaluations, namely: how quantitative data can demonstrate impacts; how qualitative data could demonstrate impacts;

that basic demographic data might show impacts; how parent volunteer skill deficiencies (writing, analytical skills) limit HTSG's qualitative evaluation capacity; how HTSG can fix individual skill deficiencies through training; and how this can lead to the all-encompassing database system. This strategy of having village volunteers trained to capture quantitative measures and qualitative data on villager issues and needs, and then transferring this back to an NGO office database, revealed HTSG vision for more comprehensive evaluation.

Evaluation plans developed in the 2015 workshops focused on facilitating meetings and interviews with parents and other stakeholders, yet the plans had many facets which HTSG found challenging. Three areas - time, people and data, are detailed below.

Time was challenging for the project team and office support they already had a full schedule of commitments and in addition, evaluation work was new for them. Staff found it hard to make time for team training and learning at their main office because field work in villages was 2-4 hours away by car. Many staff were regularly in the field or lived nearer to the communities they served than the NGO's city office.

Susan's email in December 2014, four months after the author's first visit in August 2014, indicates that the Healthcare team had not met to discuss evaluation since.

> *"I just talked to Khun, and he would like to let you know that we are fully on board with your plans for impact evaluation, but as of yet we have not found a good time to sit down with the Healthcare project team and discuss how this is all going to work. If you could give us a sense of when you are hoping to visit us again, we can make sure that the Healthcare team has already been debriefed and ready to meet with you for more details." (Susan, email, Dec 2014)*

It was not only finding time for people to come together and prioritise evaluation work, but also making sure the right people were in the right roles to support evaluation activities. Two situations show how HTSG were reconfiguring their organisation, incrementally, to be more able to do evaluation and other data/knowledge intensive work. Firstly, during discussions with Orr, the finance officer, in February 2015, she explained that she was in the middle of renegotiating her role and contract with HTSG. A key part of this change was to split Orr's role to support financial management (her current work) and English language translation (her new role). The English language support would help HTSG with report writing and liaison with international organisations. Secondly, Chai, a driver and community engagement officer, was

taking on an additional role as a translator of Thai to Karen, an ethnic minority language. He was a native speaker of Karen and Thai and could also support project and evaluation work.

One further incident is indicative of how HTSG had minimal data management capacity. In February 2015, the author asked to see what data was already available, whether on computers or in cupboards. Two staff, Mai and Mee replied that data was available from health check visits to pregnant women and mothers. However, the data had been lost when a previous employee left the organisation, taking her laptop with her in 2014. In the end, it took five months before the data was recovered in July 2015.

What the loss and recovery illustrate is that HTSG staff were not highly organised data managers. They had no data collection, storage or back up processes. They stored data on personal staff laptops, not a central computer, network or cloud service. During the February 2015 workshops, time was spent on data management, how to store data, using online services and software, backing up data on external hard drives or free cloud services such as DropBox or Google Drive, and also on file and folder naming standards, which at that time were completely lacking.

Susan's earlier idea of a full-fledged evaluation system, required changes related to time, people, and data. It necessitated more diverse organisational practices, in order to support data management, training and learning, and staff roles to take up data/knowledge intensive work, including impact evaluation. Without dedicated budgets, guiding advice or parent organisation resources, HTSG took advantage of any available opportunity to develop their organisation. This included visits by professionals, external trainings, internal workshops, and free consultancy advice. The evaluation collaboration conducted by the author fit into HTSG's self-regulated mode of learning and organisational development, as an opportunity which HTSG took advantage of.

Other examples illustrate the NGO's flexible, ad hoc strategy:

- In April 2012 a retired development professional working in Thailand, Valery, offered to evaluate a HTSG project, producing a report and recommendations on running day care centres for ethnic minority children.
- In 2013, a photographer joined HTSG field trips to villages, documenting the story of one project participant. The photographer produced what Khun called a "photo-ethnography" series of images, which were posted on HTSG's website in 2014.

- In early 2015, Orr attended free UN workshops in Bangkok on NGO accounting and transparency.
- In mid 2015, Khun was awarded a place at a UN training program on education and advocacy in Switzerland. His selection was based on his work with an indigenous network of NGOs and CBOs.

HTSG was open to opportunities for professional development, and to lead and manage their own development. However, such opportunities were intermittent, ad hoc and HTSG lacked the financial or human resources to strategically plan their evaluation, learning process.

Another vital cog in HTSG's attempts to build an evaluation machine was the continuing intern program with US Ivy League universities discussed briefly above. Interns were US graduates volunteering in Thailand for one to two years. HTSG accepted four interns during the research, Susan (2012-14), Mai (2014-15), Cherry (2015-2016) and Mia (2015-17). The interns brought English language, office and communication skills, helping HTSG interface with international funders and visitors. They had high levels of computer literacy, and writing and analytical skills. Cherry had research and statistics skills from her bachelor's degree in public health and statistics. Such skills boosted HTSG's office, English, evaluation, and analytical capabilities. Nevertheless, even with this bridging tactic, HTSG lacked evaluation capacity, equipment, skills and the time to manage data, design surveys, analyse results, or market impacts.

It is noteworthy that the English language and office skills that interns brought were critical to evaluation mechanisms. English is required for assessing international funding calls (from INGOs or philanthropic foundations), writing grant applications, understanding evaluation methods, using data management or funder compliance systems, and drafting evaluations for international audiences. HTSG's local staff could not perform these tasks in English. Only Khun, to an extent Orr, and the interns had the English/office skills for such work.

During the February 2015 workshops, it was evident that the Healthcare Project had a skilled and talented team, however they had limited evaluation experience and no experience of impact evaluation specifically. Whereas the author had prepared for workshops on different methods and uses of impact evaluation knowledge, HTSG required much more support on how to identify evaluation participants, how to draft questions, how to store data, and how to plan a basic evaluation process. Workshops lasted two to three hours each, and covered: an

introduction to impact evaluation; stakeholder mapping; identifying evaluation participants; brainstorming evaluation questions; roleplaying interviews and focus groups; planning an evaluation; and, using impact messages in marketing and communications

Different evaluation methods were discussed, but not in depth, such as before/after statistical approaches, Most Significant Change and Appreciative Enquiry. The workshops on stakeholder mapping, brainstorming evaluation questions and topics, and using these to draft an evaluation plan were particularly significant in that they reconfirmed HTSG's status as a novice evaluator. Staff were unsure who evaluations related to, who should be involved, or what questions to ask. Further visuals from the workshops follow as they are integral to HTSG's novice point of departure for learning evaluation.

At a February 2015 workshop, a network diagram of stakeholders involved with the Healthcare Project developed into a discussion around the idea that different stakeholders may view impact differently, and how this could inform collaborative work on evaluation data. Diverse stakeholders are shown in Figure 6.1, including researchers, local government, HTSG, sub-district hospitals, health centres, nursery teachers, natal and post-natal services, parent volunteers in the villages, mothers and children in age brackets 0-2, 2-4, and 4-6 years old.
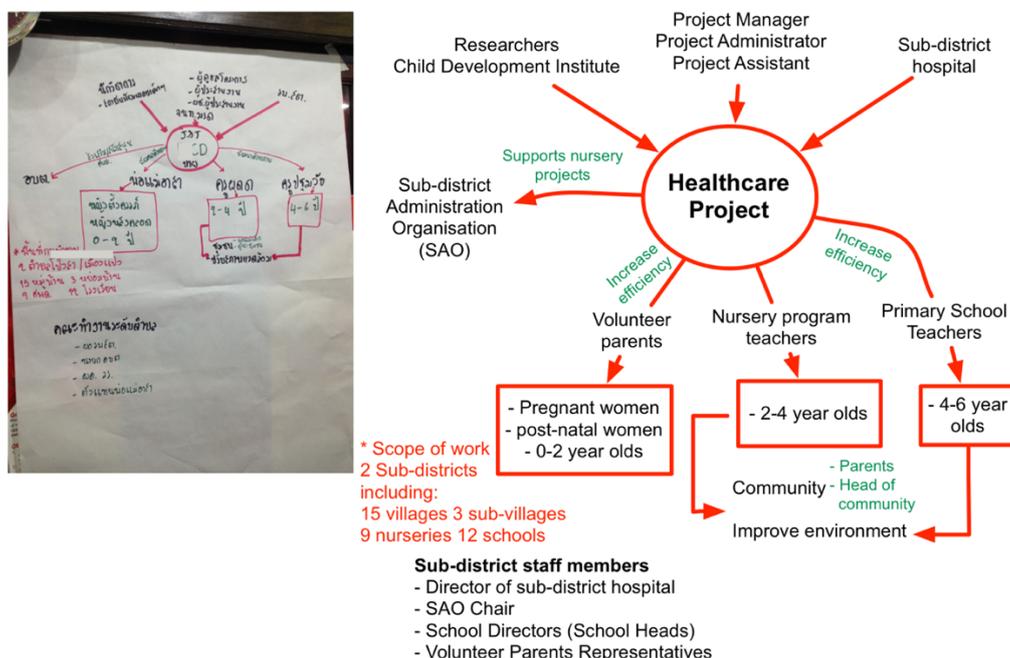
Figure 6.1: Healthcare Project stakeholder map. Thai photograph, English translation

Figure 6.2 below shows example notes from a February 2015 workshop focused on designing interviews and surveys as part of evaluation plans.
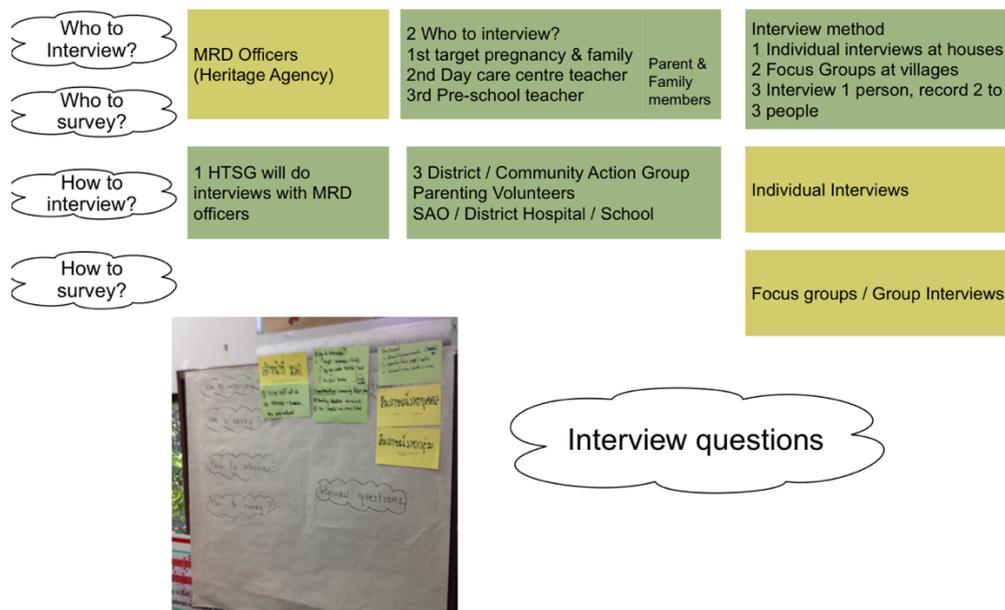
Figure 6.2: Notes on survey and interview plans. Thai photograph, English translation

Figure 6.3 below shows colour-cards used to brainstorm evaluation questions for stakeholders such as mothers, doctors, kindergarten teachers and others.

Figure 6.3: Stakeholder question cards. Thai photograph, English translation

Table 6.1 shows workshop plans for a round of evaluation activities in the Healthcare Project.

Groups to talk to	Method	Translator	When
Local Authority (TAO) Hospital Education Authority	Interviews	No	April-May 2015
Expectant mothers Port pregnancy mothers Families with -0-2 yr olds	Interviews	Yes	April-May 2015
Parent volunteers	Focus group	No	April-May 2015
Pre-school teachers Day care teachers	Focus group	No	April-May 2015
Healthcare project staff	Focus group	No	April-May 2015

Table 6.1: Plans drafted in final workshop for 2015 evaluation activities

These excerpts give a flavour of how basic training on who to include in evaluations, who to talk to, how to form questions, and how to plan activities was conducted with HTSG during the February 2015 workshops. Beforehand, there was no method in place, no database, no trained staff, no technologies, no tools, no software, and no staff assigned to impact or evaluations. HTSG had no evaluation machine.

After discussing survey design, brainstorming questions, and planning during the February 2015 workshops, the author and the Healthcare Project team agreed to run initial interviews and focus groups in April-May 2015. However, these did not take place until November 2015, when a new intern arrived from the US, and helped coordinate the activities.

One significant workshop excerpt related to the authenticity of participant responses to evaluation questions. Pang, the Healthcare Project Manager, remarked that participant mothers did not want to admit that they had not changed breastfeeding behaviour after project activities encouraging them to breastfeed. Pang suggested mothers did not want to upset themselves, volunteers, or project officials:

> *"Sometimes we should not be too direct, because they will feel bad if we ask them whether they have changed their behaviour and they haven't changed. They don't want to feel bad. They don't want to offend the project official either." (Pang, field notes, February 2015)*

Pang was reflecting on participant responses, questioning whether evaluation data was authentic, and acknowledging that participants prized positive relations over evaluation truths.

In learning evaluation mechanics, HTSG were juggling multiple, overlapping funding lines, projects, activities, and districts. Khun's email from November 2014 illustrates the juggle:

> *"Since the Healthcare Project will potentially have 2 project sites from two sources of donors from January to July, 2015 in different time frames, one from J-Corp will be its 2nd half in Muang district and one from Global Children will be its 1st half in Kweng district. Then another J-Corp project support in which I proposed to work in both districts will commence on August 2015. So we have to have a very good plan to incorporate with our new learning/developing on impact assessment research with you." (Khun)*

The excerpt illuminates how HTSG survive, running related projects, under different funders, with different reporting frames and submission periods, different extension timelines, and in different districts. Aims, approaches and projects overlap across short timescales. Overlaps, unstable funding, and relationship maintenance with district stakeholders not only require "*a very good plan*" as Khun remarks, but complicated evaluation work too. Evaluation designs, timeframes, data collection and reporting must be attuned to overlapping projects and activities.

Despite such juggling, HTSG did have some data from reporting process in previous years, but this data had problems. Some data did not focus on impact, some was lost, some misused, and some not recognised as useful for evaluations. Annual reports almost never referenced "impacts" or stakeholder voices. For example, a 2009 report makes no reference to impact or

evaluation, but has sections on "Project Goals", "Objectives", "Activities", "Planning", and "Outcomes" (Project report, 2009). The 2014 Healthcare Annual Report (Figure 6.4) contained figures that could be used to construct vaccination coverage and urgent care related impacts. However, they were used only to describe beneficiary group details.

Beneficiaries

The primary beneficiaries of the ▮▮▮ project in the 2 sub-districts of ▮▮ are pregnant woman, children 0-6 years old and their parents or guardians, who receive direct support for care-giving duties. Secondary beneficiaries consist of community members responsible for daily childcare and officers from other relevant agencies. In total, ▮▮▮ currently accommodates 1,342 beneficiaries, who can be categorized into 3 implementation levels: family, childcare center and school. They consist of expectant mothers and their unborn children, Thai and Ethnic (Karen, Lisu, Lahu and Shan) children 0-6 years old, parents and/or family providing newborn care, parenting volunteers (PV), childcare providers, pre-school teachers, educational supervisors, public health officers, community leaders, and sub-district officials and representatives. Valuable and active support from project stakeholders allows ▮▮▮ to accomplish its objectives and grow its beneficiary pool.

- 43 Parenting Volunteers (PV) from 15 villages were educated on scientific early childhood care and development regularly.

- Every newborn baby in the project area received birth registration certificates.

- Every child in the project coverage received necessary vaccinations according to the country's universal healthcare policy.

- Most children aged 0-6 months were breastfed according to parenting volunteers' suggestions.

- 9 young children received urgent care support from ▮▮▮▮ according to PV's monitoring reports.

Breakdown of each beneficiary group by project year

Group	Year 1	Year2	Year 3
Children age 0-2	212	284	179
Children age 3-4	135	201	201
Children age 5-6	207	202	202
Total number of primary child beneficiaries	554	687	582
Expectant mothers	27	121	58
Adults with 0-2 year-old children	212	284	179

Figure 6.4: Healthcare Annual Report (2014) - figures not used to show impacts

Figure 6.5 below is a graph from a 2014 annual report showing self-assessment performance metrics for parent volunteers. The bars show volunteer knowledge, skills, planning and

achievements amongst in the Healthcare Project. This is an attempt to evaluate performance, not impacts. The focus on volunteers (not mothers, babies, doctors or other stakeholders) and the lack of change indicators mean this data does not technically demonstrate impacts.

Figure 6.5: Report showing parent volunteer self-assessment performance metrics

From HTSG's minimal evaluation mechanics and the workshop activities in February 2015, it is clear that the author, as a researcher/consultant, brought impact evaluation knowledge that was lacking at HTSG. This section has illustrated how HTSG did previous evaluation and what they wanted to learn during the workshops.

In summary, the excerpts in this section suggest that HTSG, as an organisation, was engaged in a range of activities, changing and learning while creating their impact evaluation mechanisms, but struggling too. Struggle was evident in Khun's reference to the "*challenging funding environment*" and in imagining a future "*full-fledged*" evaluation system. The word "*struggle*" was used on multiple occasions.

> *"We struggle with it [impact evaluation], but I know Khun really cares about it, ... and wants it to be an area for improvement for us" (Susan, August 2014)*

"I've started brainstorming ideas for focus group/interview questions for next year's Healthcare Project, which will focus on malnutrition and dental problems in young children. I'll pass this information to the next Intern who will probably have a stronger background in public health and research methods than I do, and will hopefully be able to work on a baseline data study after I leave." (Susan, email, May 2015)

"It is a struggle, sometimes it feels like the more time you spend doing the research, collecting the data, doing impact evaluation means less time doing the programming ... especially when we don't have anyone specifically dedicated to it, and there is a lot of other things going on, and you know ... I don't think anyone at our organisation has had extensive training at it, ... but in terms of educational backgrounds, but no one has really studied anything related." (Cherry, Skype call, November 2015)

HTSG were trying to establish a full-fledged evaluation system, while searching for and not losing time, people and data, sourcing ad hoc opportunities and resources such as interns, learning about evaluation in free workshops, dealing with doubts about data authenticity, juggling multiple funders and evaluation needs, and producing and positioning data to show impacts.

Activities in these organisational changes, the most extensive evaluation area of learning and struggle for HTSG, included:

1. defining assumptions and envisaging a more holistic evaluation system
2. prioritising training, new data/knowledge intensive roles, and data management
3. capacity building and training staff and volunteers in data/knowledge intensive tasks e.g. research/analysis methods and techniques, report writing, bid writing, data management, language skills and translation (English, Thai, local languages)
4. identifying cost/resource effective opportunities (e.g. internships, training, pro-bono expert advice etc to build evaluation capacity
5. adopting evaluation attitudes e.g. navigating doubts about participant response authenticity
6. managing multiple and parallel funders, projects, reporting and evaluation processes

7. identifying, locating and using impact-related data appropriately in reports and communications

6.3.3 Mid-cycle activities: Learning to make cells and stories

This section looks at how HTSG made data cells and stories for use in building impact narratives. Empirical excerpts illustrate: how HTSG were learning, struggling and transforming; how cells were made but also lost; how numbers could submerge the dynamics of poverty; and how impact episodes could be silenced if illegible to the impact machine or too risky for the NGO.

Using data to support impact evaluation was a new process for HTSG, and one which they were struggling to master. The level of detail and range of questions used to collect data was limited. Data focused on demographic information about project populations and participants (e.g. names, addresses, marital status, place and date of birth of babies, ethnic group etc) and on project aims such as the prevalence of breastfeeding or whether infants had completed their due vaccinations. The data in Figure 6.6 were used for project monitoring, but not comparatively to show intervention impacts.

Figure 6.6: Healthcare data used for monitoring, not for demonstrating impacts

As noted earlier, HTSG recorded some statistics that could support evaluations of impacts, but this was either generic data or not placed within discussions of impact. The data in Figure 6.7

below was available for 2008-11, but was reported to funders in a section on background context, what HTSG labelled "*situation analysis*" (HTSG 2011 Healthcare Project Annual Report). Such data could have been compared over multiple years and tied to project goals in order to show impacts: for example, increasing numbers of children attending day care centres.

General Population Breakdown

Villages	No. of Families	Population			Ethnic Groups
		Male	Female	Total	
Bar ▮	103	120	117	237	Shan/N.Thai
Bar ▮	193	204	216	420	▮
Bar ▮	88	119	108	227	▮
Bar ▮	434	740	610	1350	▮
Bar ▮	95	118	119	237	▮
Bar ▮	26	60	51	111	▮
Bar ▮ Mu ▮	127	126	121	247	▮
Bar ▮ i	40	72	69	141	▮
Bar ▮	35	84	71	155	▮
Bar ▮	35	84	71	155	▮
Total	**1176**	**1727**	**1553**	**3280**	

Day Care Centers

Day Care Center	No. of Children			Day Care Workers
	Male	Female	Total	
W ▮	17	14	32	2
M ▮	10	6	16	1
Total	**27**	**20**	**48**	**3**

Schools

School	No. of Children			Pre-School Teachers
	Male	Female	Total	
B ▮ N ▮	27	18	45	2
S ▮	44	44	88	3
Total	**71**	**62**	**133**	**5**

Figure 6.7: Data in 2011 report used to describe general context, not impacts

HTSG did not know how to use this data to show either impact over time, before / after results, significant changes, or to justify activities or theories of change. The organisation was learning to position cell data and outcomes, but the outcomes listed in Figure 6.8 below, in the same 2011 funder report, show a concern with encouraging project participation (Outcome-1), participant knowledge (Outcome-2) and collaborative relationships (Outcome-4). In contemporary impact evaluations, there is an increased need to show and quantify direct impacts upon participants and how they have sustainable changes in behaviour, not changes in

understanding, knowledge or relationships alone. Technically speaking, no impacts were framed in the 2011 report, and this was typical of reports from 2009-2014.

Outcome Summary

1. ▮▮▮ has achieved a good level of coordination within the community, attracting a higher level of participation among parents, parenting volunteers, childcare providers, pre-school teachers, sub-district officials and representatives, public health officers, community leaders, and educational supervisors than in years one and two.

2. Stakeholders, particularly parenting volunteers, child care providers, and pre-school teachers have a higher quality of knowledge and concern for children and any of them can now help support or initiate activities to promote child care and development through play.

3. Four new early childhood care centers for children ages 2-4 were opened in Ba▮▮ P▮▮, and S▮▮▮ villages (Ma▮▮▮ sub-district) and P▮▮▮▮ village (W▮▮▮ sub-district).

4. Collaboration between stakeholders and local partner organizations is very strong, especially with the sub-district administration. This creates a network that will likely sustain an independent and long term ▮▮ project in ▮ district. In particular ▮▮- ▮▮ program is working closely with the ▮▮▮▮ Institute of Child Development to provide additional training and resources to the ▮▮ program.

Figure 6.8: Excerpt from 2011 Healthcare report show outcomes, but not impacts

In late 2015, Khun arranged for Cherry, a graduate from another prominent US university and major in public policy and global health, to intern with HTSG. Cherry had data management and statistical experience and Khun asked Cherry to lead HTSG's evaluation work. Cherry contacted the author via email and requested advice on three areas of evaluation: firstly, an application to the EU for a new migrant support project bid; secondly, how to use a large amount of evaluation data generated recently; and thirdly, how to set up an evaluation system for a dental hygiene and nutritional health initiative. During a Skype call, Cherry explained the problems she was facing.

"I have pretty strong quantitative background. And I've taken certificate courses and did my honour thesis, on using statistical analysis and all that, but in terms of like, ... it's very different working from just this perfect data set to try to design a whole impact study, ... and I think, I'm learning what a big job, what a big role

they've given me, and … I just have a lot of learning to do." (Cherry, Skype call,
November 2015)

"I just I think it's very different goals you know, proving the hypothesis wrong and
trying to evaluate and improve projects … and I just feel, … I feel they are very
different, at least more complex. So, I have some background, but definitely new,
new, … new stuff going on for me for sure." (Cherry, Skype call, November 2015)

In these two quotations, Cherry identified the imperfect nature of field data, and her own learning about HTSG's context. During the call Cherry was anxious, stressed, and at times demotivated and close to tears. Unprompted, the next day she emailed a thank you:

"Thank you again for taking time out of your weekend to speak with me. I so admire
your commitment to supporting HTSG's projects and truly appreciate your advice.
I feel renewed and motivated after our chat!" (Cherry, email, November 2015)

This excerpt illustrates how Cherry, a proxy for HTSG more broadly, was struggling to learn impact evaluation and apply it in front-line environments. Producing data, statistics, and indicators constituted evaluation processes that were emotionally impactful upon HTSG as an organisation and Cherry as an individual member of staff. However, such stress, demotivation, and the wider non-target impact of evaluation performances upon the organisation are not included in data cells, columns, rows, or reports. The excerpt suggests stress and emotion, elements of evaluation in practice, are illegitimate and edited out of impact construction processes.

Not only stress and emotion, but non-project activities and impacts were also edited out of HTSG's evaluation reports. In February 2015, HTSG responded to a house fire incident where two young children from a marginalised community were left homeless, injured and distressed. HTSG and the other NGOs drew on reserve funds to respond to the incident (Figure 6.9). These funds were not related to funders, projects or goals. When asked if this was a one-off incident, Khun replied:

"Actually, no, it's not unusual. Hundreds of incidents and issues like this happen,
… all the time, many times in a year." (Khun)

HTSG partner with other local NGOs to respond to such incidents, and the incidents are diverse in nature and location. However, HTSG have no specific aid or funding stream from national or international donors to respond to such ad hoc incidents. Incidents like these present problems for project planning, pre-set indicators, or theories of change which do not envisage diverse problems outside any funder's niche concerns. Impacts from non-project activities such as this and the "*hundreds*" of others are difficult to fit into an evaluation regime as they can be outside the goals, interests, planning, and evaluation apparatus of international funders.

To: Director of HTSG
From: Projects Coordinator
Re: Emergency Fund Application
Copy: Accounts and Finance Officers
Master Chai ▮ ▮, House ▮ Bann ▮ ▮, Moo ▮ ▮ a Subdistrict, ▮ Province, who is currently attending ▮ n Nursery, lost his family house, including all his personal belongings, clothes, a motorcycle, and cattle in a recent house fire.   District Field Office has therefore submitted an application for an approval for an emergency fund for him. This is in order to purchase 30 bags of cement, which costs 130 Baht per bag, and total amount is 3,900 Baht (Three thousand and nine hundred baht only). The cement will be used to re-build a new house for Master ▮ and his family to live in. I would be grateful if you could approve an emergency fund of the said amount (3,900 baht) for Master ▮ so that the project co-ordinators can carry out the work for the boy and his family. This memorandum is submitted for your consideration and approval.
Yours faithfully,
Name, Projects Co-ordinator
Approved with immediate effect.
This is an emergency case as boy is distressed.
Authoriser's name, Projects Manager

Approved for the case to be urgently submitted. K, 20th February 2015

Figure 6.9: House fire fund request memo. Site photographs, English translation

Although TIEK prescribes identifying goals and target impacts in advance, and use of project planning, log-frames, Gantt charts, and budgeting to keep aims on track, in this case, a child's welfare required an immediate response. This could not be met if local NGOs did not have responsive, local networks to assess what is happening and react rapidly.

In a final empirical example, extending the breastfeeding situation discussed earlier, Cherry described an opportunity to convert beneficiary claims into measurable quantitative data. In a Skype call, she explained how some mothers did not want to cause conflict or disappoint project officers, and therefore did not give truthful answers to questions about breastfeeding behaviour.

> *"I'm interested some of the … one of the quantitative things that I am concerned about is the exclusive breastfeeding. Basically, on all of these I can look through the home visit reports and on all of them there is a check next to exclusive breastfeeding zero to 6 months. But the story I am getting is that … that is not true. And I think both having a quantitative and a qualitative indicator that is looking at that … One of things that I have taken from a WHO … err… feeding module or something is they basically just have a list of basic things that the mother might have fed to the babies, so it's the women who have children who are under 6 months and you just go through the list and like 'Have you ever given your baby your child water?', 'Have you ever given them tea, bananas, have you ever given them mango?', that sort of thing." (Cherry)*

Cherry's strategy was not to capture an engaged rendering of participant voice as to how they fed infants, or why they breastfed or not. Rather, the strategy was an attempt to mitigate the problem of untruthful responses from mothers. It erected a quantitative frame for data. Yet, in past reports, HTSG had observed how new mothers returned to work shortly after giving birth to generate income for the family. The assumption in Cherry's strategy was that discrete food item questions produced responses with measurable data and ejected inauthentic answers. But the measurable data strategy elides the opportunity to learn more about poverty dynamics around a mother's rapid return to work after giving birth. Countable data displaced contextual depth.

In the examples above, HTSG were learning about the limits of the data they had, how to use data to build impact narratives, how to use data in the right places and ways, how to bifurcate data, how to eject personal stress or unplanned impacts from legitimate data stores, and how to capture statistics even if they displace more intimate knowledge.

Activities involved in learning to make the right kind of cells and narratives included:

1. learning how to use general data e.g. monitoring data in impact construction;

2. identifying impact related data for later analysis, positioning and packaging;

3. differentiating between data related to outputs, outcomes, and impacts;

4. assigning human resources to filter legitimate and illegitimate data;

5. editing evaluator stress, emotion and contingent organisational processes out of legitimate cell data and impact narratives;

6. editing unplanned or peripheral data out of legitimate impact construction processes;

7. converting doubts about qualitative authenticity into discrete, measurable data; and

8. submerging deeper relations and the complexities of poverty/marginalisation by elevating data that can be efficiently counted and processed

6.3.4 End-cycle activities: Learning to package for audiences

HTSG were only partially able to package data into knowledge products for different audiences, although they were learning to do so. Efforts related to the contents of funder reports, limited ability to build and market impact messages, and a lack of office and English skills. Although HTSG were experienced project planners and were increasing their advocacy work, HTSG were not skilled at crafting impact narratives for evaluation or communications. They even expressed doubts about claiming impacts.

Firstly, in the 24 project reports reviewed during the research, 20 were in English, four were in Thai. All related to projects run between 2010 and 2015 and none mentioned impacts. Statistical data was used to describe project participants and demographics, such as participant names, addresses, ages of children, monitoring data on project activities such as whether new mothers were breastfeeding, whether young children followed vaccination schedules, and the regularity of visits by project front line staff or village volunteers. Statistical data was not collected or collated to demonstrate typical forms of impact, such as before/after comparisons, regression analysis, counterfactual studies, control groups, cost benefit analysis. Project reports included sections such as "Project Background", "About Beneficiaries", "Project Activities" and "Project Outcomes". No reports had sections on impacts.

Secondly, there were some doubts, or reluctance, to claim credit for uncertain impacts. During a workshop dealing with sharing impact messages via marketing literature, public relations, press releases, or website updates, Khun and Susan expressed concern that short-term project cycles of 12-18 months, typical of HTSG projects, were too short to show real impacts. Whilst

this suggestion is supported in the literature on evaluations, impact can be identified as emerging in the short-term or sustainable over time. In the meeting, the project team were doubtful that they could claim emerging impacts. Heads nodded during the discussion, but the team wanted to be able to demonstrate real impacts in terms of fixed project successes and were doubtful that emerging impacts or impacts upon non-target beneficiary stakeholders were legitimate. Despite or because of their novice status, HSTG had a high standard for claiming impact.

Thirdly, formal project management language and activities framed HTSG's work. During the first visit in August 2014, the author was given a formal presentation lasting some two hours. The presentation contained slides on all HTSG's projects, aims, backgrounds, activities, staffing, planning, mission and vision. Neither intended impacts nor actual impacts were included in the presentation. Khun confirmed this was a standard presentation they used with all visitors. It was a formal device that stressed projects and mission statements, but did not communicate the value or impact of their work.

Of four projects running in 2014, only one had an accessible, memorable, catchy title, the "*All Children, All Rights!*" project. This name was chosen by a US intern, in 2012 by Jenny. "*I think it is ... sexier than our previous slogan for the project*", suggested Khun, knowing that this title was not the mundane kind of detailed name that HTSG projects normally had, full of specific terms, grammatical errors and noun clauses. Other project titles included:

- "Promote Child Rights to Participation and Advocacy Campaigning"
- "Project to Strengthening Community - Based Early Childhood Care and Development in Border areas of Thailand as Child Friendly Pre-School and Day Care Centres"

Formal names were generally used and Khun recognised that a catchy, sexy kind of title could support public relations. However, Khun relied on interns to be able to support such creative labelling and marketing using their native English fluency. It could be said that HTSG lacked the English and communications or marketing skills to label projects in a catchier way, but it could also be part of their project management approach and orientation. Either way, the language used did not result in accessible project titles.

Fourthly, Khun was aware of the need to package impact in attractive and legitimate messages, and of the need for HTSG to produce knowledge as a resource for other development partners, such as governments, other NGOs, funders, or the general public. He saw the future of HTSG as involving not only direct services to marginalised communities, but also becoming an important advocacy agent for children's rights in the region. HTSG were working on advocacy initiatives with the Asian Child Rights network and with the Thai government on sections for the new Thai constitution that would improve children's rights to education and healthcare. Evaluation data and knowledge is useful for such advocacy work: for writing concept notes, proposals, strengthening policy advice, and supporting recommendations. Results-based management and evidence-based policy initiatives value robust data, evaluation rigour, expert analysis, professional writing, and (often) English.

Khun described the opportunity to influence national policy:

> "It's a great opportunity. The last government was not interested, so even though it's only a two-day event, a member of the government from Bangkok will be here and they will listen to us and to other organisations in our local civil society. The municipal government is very supportive and have put 40,000 baht to helping us build the network of civil society organisations. HTSG wants to play a big part in this, and we are involved in drafting the MOU, which will hopefully contribute to the new constitution, on rights for children of any ethnicity to get proper education and healthcare services." (Khun, February 2015)

HTSG relied heavily on Khun and its interns for the skills required to produce advocacy knowledge, and for packaging impact narratives with related expert knowledge. To have a voice in advocacy and policy circles, HTSG needed to show their effectiveness, impacts and credibility not just in a project, but as a credible organisation and partner. Yet, packaging evidence and data on impacts as part of wider advocacy, marketing, or brand/reputation building was still in its infancy at HTSG. HTSG did not have regular publications, professional branding or staff assigned to such roles.

Fifthly, using qualitative data and narratives in impact construction was also challenging for HTSG. HTSG had in previous years used qualitative profiles as part of their data collection activities, however these were never deployed in reports or communications, or made publicly

available. HTSG shared 10 short profiles with the author. The profiles were based on interviews with stakeholders, including doctors, teachers, volunteers, government officials and project participants. Figure 6.10 shows a profile of two teachers.

2. Miss ▮▮▮▮▮▮▮▮▮▮ Pre-school teacher at Baan ▮▮▮▮▮ since 2003
   Miss ▮▮▮▮▮▮▮▮▮ Pre-school teacher at Baan ▮▮▮▮▮ since 2003
   They joined the ▮▮▮ program since 2005 and expressed that they saw vital project's benefit to children. They were trained on **fable telling techniques, creational activities, child rights awareness for early childhood care and participated in 'Learning Environment Reform' activities. <u>Gained capacities are</u>** skills and classroom organization techniques to help young children engaging in learning such as using fable tales, group dynamic activities.
   -Child Rights Knowledge make me awareness to participate with children activity – so children feel comfortable to talk with us more than before and told about their needs. I think parents and all of teacher should awareness about child rights, **using ▮▮▮▮▮▮▮ products as learning materials.**

Figure 6.10: HTSG project profile of two local teachers

The text in the profile above has been edited to map directly onto project goals e.g. "*fable telling*", "*rights awareness*" etc. The text uses both first person "*make me awareness*" and third person, "*They joined*". The profile does not make room for alternative perspectives or voices different to project goals. This makes little use of the ability of text and qualitative research to open issues up for new learning or hearing about the experiences of respondents. It is not clear why HTSG did not further develop or use the profiles, perhaps they had doubts about their legitimacy or effectiveness. In the end though, they were not used to claim impacts.

In sum, activities identified as part of learning to market impacts, bundle-in expertise, or cut out doubts, included:

- adding sections about impacts into reports;
- using statistics to show changes and impacts;
- highlighting specific impacts, successes and achievements;
- claiming credit for impacts, even if they are emerging and not yet long-term;
- packaging projects in presentations and through accessible language, titles etc;
- packaging evaluation evidence and impacts to strengthen advocacy initiatives; and
- using qualitative data and profiles in evaluation and marketing.

6.3.5 End-cycle activities: Learning to pitch what to who

For HTSG, "pitching" or sharing evaluation data and knowledge involved a fragmented terrain of audiences and goals; in CHAT terms, their object of activity. HTSG was shaping evaluation data and knowledge for diverse partners, not always for funding opportunities and not always for post-intervention impact depictions. They were learning to use evaluation data/knowledge in both financial and non-financial exchanges, and as part of pre-intervention grant applications. The excerpts below illustrate some of HTSG's diverse evaluation knowledge pitching and sharing activities with local healthcare providers, the Thai government, and an international grant application, as well as some doubts about impact pitching and their own humble attitudes to success.

A December 2013 Healthcare Project report (Table 6.2 below) shows interest in knowledge sharing with a district hospital and a public health authority, and a digital database to support monitoring activities with volunteers.

HEALTHCARE PROJECT REPORT, 2013 (page 22-23)
LESSONS LEARNED SECTION
With the project improving communication among local networks, Muang Amphur Hospital and Muang Amphur Public Health Offices, a more in-depth knowledge sharing was established. There is also a more streamlined communication between volunteers and hospitals, especially regarding emergency cases.
KNOWLEDGE SHARING SECTION
Create digital database which volunteers can track home visit schedules and HTSG can monitor progress. Volunteers have a heightened sense of control over their input in the project and have been more likely to participate in trainings.

Table 6.2: Excerpt from 2013 December Report

During the author's first visit in August 2014, Khun described how government healthcare data on minority groups was frequently incorrect or lacking. Discussions with the author then focused on understanding HTSG's networks and how they might extend sharing of evaluation data/knowledge with different local partners, as well as potential funders.

The Healthcare Annual Report 2014 (Table 6.3 below) described a new initiative through which HTSG was sharing a database with local health authorities. HTSG's more accurate data, collected from the field where they had close connections to project participants and parental volunteers, was being used to fix gaps or errors in a government sub-district hospital dataset.

HEALTHCARE PROJECT REPORT 2014 (page 11)

DATABASE, EVALUATION & COOPERATION SECTION

Database: According to data collected by volunteers on the project target group (expectant mothers, postpartum mothers and children ages 0–2 years old) during home visits, there is a discrepancy between information from volunteer records and the sub-district hospital. The sub-district hospital does not have information on family health history of expectant mothers, postpartum mothers, or information on child development assessment. Plans to improve database access for organizations.

Impact evaluation (external consultant) will be supporting from 12/2014 to 12/ 2015.

Cooperation on project home visits, monitoring and local hospital to share information on households and minimize discrepancy between project database and local hospital's.

Share information with doctors / ask for suggestions

Table 6.3: Excerpt from 2014 Annual Report

By the end of 2014 HTSG had begun to use their data in three new ways: 1) to support the local health authority; 2) to inform evaluations of impact; and 3) to act as an agent connecting communities, volunteers, projects, doctors and public health offices. HTSG were sharing data with stakeholders in new ways and learning how to pitch or share their data with partners.

Secondly, advice and evidence were not always straightforward for policy developers, and this made the advocacy work of HTSG somewhat complex. For example, Khun shared his concerns about how Thai birth certificates, and thus rights to nationality, were issued by the Thai government only upon receipt of a proof of birth document given by hospitals when a child is born. However, many ethnic minority mothers who lived in mountainous areas did not give birth to children in public hospitals but rather in their villages. The Thai government often refused to grant birth certificates and hence nationality to children born away from government hospitals. As nomadic mountain dwelling minorities had traditionally moved between Thailand, Myanmar and Laos for centuries, many parents had no Thai nationality papers for themselves, or their children.

As such, bringing deeper understanding on issues such as this into the constitution drafting process on children's rights was an opportunity where HTSG could progressively influence national policy using their own knowledge. Khun talked about the increasing need for advocacy and was genuinely excited that HTSG had this opportunity. His plan was to argue the case for a constitutional section that guaranteed the rights of children born in mountain dwelling ethnic groups, and he saw the collection of strong evaluation evidence and data as important for supporting his argument. This provided another reason for HTSG's desire for evaluation expertise, and another opportunity they had to deploy evaluation data and knowledge with specific audiences. In this case, HTSG's evaluation plans, objective, and the "pitch" were neither aimed at aid funding markets, nor partnership development primarily. Rather, the pitch was oriented towards advocacy and national level policy change.

A third trajectory for evaluation data/knowledge exchange was a new grant application to the Red Cross (Migrant Services Project Grant Application, 2015). In November 2015, Cherry asked for advice regarding ongoing evaluations and a new project proposal. The proposal was to set up a support centre for migrants from Myanmar who were working in HTSG's home city, to co-design, with migrant communities, a network of representatives who were trained on migrant rights and informed about government, private, and non-government organisations who could offer support. Such organisations included hospitals, lawyers, labour tribunal panels, and housing services among others.

In one email, Cherry described the current state of the proposal as "*a mess*" that she would be finalising over the weekend to submit by the funder deadline on Monday. The funding was the European Union, and the grant was administered by the Red Cross. The author reviewed the

proposal, and added two new sections: the first on project goals, logframe, indicators and target impacts, and the second on monitoring and evaluation.

In January 2016, Cherry emailed to confirm that HTSG had been successful, and won a grant for 8 million Thai baht (approximately £180,000). Migrant support services were a new area for HTSG, and it was a considerable project award for a small Thai NGO.

The added value of the authors' advice is not indisputable; however, the initial proposal draft had no monitoring or evaluation section, no indicators, no outcomes, and no mention of impacts. The fund administrator's proposal guidelines also confirmed a need for applicant organisations to meet "*narrative monitoring and reporting requirements*" (IFRC (2015: 22) Guidelines for Proposals). Therefore, one could surmise that showing evaluation knowledge and plans on the application was an important aspect of successfully securing the award.

This excerpt shows how HTSG, as a small NGO, used evaluation knowledge not only in post-project evaluation practices, but also in pre-project practices of successfully applying to international funders such as the Red Cross / EU for grants. Evaluations sections are standard in project proposals and regularly form part of a proposal assessors' scoring criteria[28]. HTSG had to pitch intended impact and evaluation knowledge to funders before a project could start.

Finally, HTSG's struggle to pitch evaluation data/knowledge to diverse audiences was also reflected in comments during a February 2015 workshop on communicating impacts about how humble they were as an organisation. When discussing the need to write stories for their website and local media outlets via press releases, Orr made two insightful remarks:

> *"But we don't have those [communication or marketing] skills, we are working here or in the field. I don't know, ... we're humble. It's difficult for us."* (Orr, February 2015)

> *"I mean when we make project plans, ... are we just imagining our impacts?"* (Orr, February 2015)

---

[28] This claim is backed up by personal correspondence with a number of UK NGO staff from 2013 to 2016, as well as by funder guidelines online, which have stipulated M&E sections in grant proposals.

One interpretation of Orr's comments is that HTSG do not wish to overstate, brag or "sell" their impacts (first comment) and that they are willing to reflect on their planning and goals (second comment). They are focused more squarely on project management and engaging communities than convincing others about their successes. However, in marketing NGO effectiveness and garnering public and financial support in a competitive funding climate, HSTG arguably needed to sell their impacts and communicate effectively. This fits in with observations (e.g. Roche, 1999) about NGOs becoming more professional, more strategic, and more like a business. Reflection is good in terms of internal learning, but impact narratives must be made clear, solid, and convincing for external audiences.

HTSG has a 20-year history of deploying sound project management to offer direct aid services to marginalised rural children and their families. Increasingly though HTSG has had to create impact knowledge and data as documented representations for public communications, funding and advocacy work. In opening topics on selling impact stories, press releases and media relations, the author had moved too far away from HTSG's traditional work and skills base, suggesting not just a need for TIEK, but a shift in attitude or culture, from humility to self-promotion. This pitching was new and uncomfortable.

Identifiable late-cycle activities included:

- identifying and engaging with local, national and international partners to develop and share impact evaluation data/knowledge;
- identifying opportunities for collaborations on service delivery and advocacy using impact evaluation data/knowledge;
- using impact evaluation knowledge in grant applications, particularly regarding goals, activities and impacts, as well as describing monitoring and evaluation approaches; and
- reflecting on organisational skills, needs, relationships, history, and attitudes to impact marketing and impact sharing with different partners, and how this can change or challenge existing organisational attitudes.

6.4   Analysis section

This analysis section uses existing CHAT activity systems and contradictions, as well as the novel concept of temporal activity chains to respond to the three research questions. Section 6.4.1 analyses the case evidence using CHAT activity systems and contradictions. A primary contradiction is found relating to not one, but two kinds of impact: one humble-form and one marketing-form. Section 6.4.2 utilises temporal activity chains to describe the power and practice relations silenced in TIEK and DIKW-inspired knowledge management. Section 6.4.3 reviews what was learned from the case about acknowledging power and practice into evaluation activities.

6.4.1 Activity systems and two contradictions

Research question 1 asked: how is development NGO impact data and knowledge constructed in practice? Two responses are offered: firstly, HTSG's lack of expertise, and secondly the evidence of two contradictions. One contradiction relates to two kinds of impact. The second contradiction concerns a current and future data/knowledge intensive version of HTSG as an organisation.

HTSG's novice evaluation status and self-directed opportunism regarding learning can be understood using the CHAT elements, tools, technologies / techniques, rules and norms, and community interactions (Engeström, 1987; Karanasios, 2014). In terms of tools, HTSG were not technology intensive, nor organised managers of data. They lost a critically important laptop when a staff member left, and it took them months to recover it and the accompanying data. They used no tablets, specialised software, file or folder standards, or a central database. HTSG used no evaluation concepts, methods or techniques at the beginning of the collaboration. HTSG had no working evaluation rules or norms. Khun wanted a more quantitative and statistically oriented set of rules and norms, but was unsure how to set this up. Doubts existed too, evident when Orr questioned whether plans and goals were real or imagined. Furthermore, the community of people involved in HTSG's work, and their divisions of labour, prioritised engagement and project management but not evaluation, communications or data/knowledge intensity. HTSG were starting to assign staff to new roles, requesting interns with evaluation skills, and beginning to share data with partners, but not yet marketing or reporting narratives of impact. HTSG's novice status and struggle to learn impact evaluation

can be seen as part of a CHAT expansive learning process (Engeström, 1987; 2001). They did not yet fully understand the scope of their expansive learning transformation, which is typical in such processes (Engeström, 2001: 139).

Impact-1 and HTSG-1

There were two CHAT contradictions evident in the case. The first was a primary contradiction between the construction of two forms of impact, here labelled "Impact-1" and "Impact-2". Impact-1 is a more doubtful, local and humble perspective on impacts as uncertain, largely oral in how they may be communicated and shared, multi-voiced, expressed in communities or to frontline NGO staff. Impact-1 involves dialogue about tangible changes within communities, within the intervening NGO, and between those two contexts. In this case, such discussions were not technically framed, based on digital data, nor clearly documented. Community members could discuss NGO advice on using government hospitals for births and vaccinating children. People could express inauthentic claims about impacts too, for example where mothers claimed HTSG breastfeeding campaigns and training were effective in altering their behaviour, but did so because they did not want to upset NGO staff.

On the NGO side, knowing Impact-1 involves staff commitment to those in need, listening, extensive fieldwork rather than office work, claims to care and honest hard work, humility, and potential solidarity with marginalised groups in specific places and times. Impact-1 is thus partial, negotiable and situated (Blackler, 1995). Impact-1 knowledge is less mobile than Impact-2 knowledge or digital data representations. In contemporary evaluation regimes, Impact-1 is devalued, submerged, or masked (Engeström, 2008: 36-42). It does not conform to objective scientific standardised assessments. It is highly contextual, changeable, and expressed through dialogue.

A second contradiction of a tertiary nature in CHAT terminology (Engeström, 1987: 103) concerns HTSG as an organisation itself. An older or more established version of HTSG, "HTSG-1", constructs Impact-1, based on staff and volunteer hard work, commitment, humility, engagement with communities, mothers, and children, and effective project management. HTSG-1 does not use evaluation models, data management processes or ICTs to create Impact-1. However, their projects do contribute to healthcare changes, increased breastfeeding rates, better dental hygiene etc from the NGO's point of view. In this older

version of HTSG, data/knowledge intensity is very limited, and there is less digital work, less office work, more field work, less reliance on English and more reliance on Thai and local ethnic languages. Field work and engagement activities take precedence as illustrated by Cherry when she said she feels like there is "*less time to do the programming",* which is her proxy for real work. There is less need to place impact narratives in the correct sections of reports, bifurcate data effectively, split out non-project impacts (e.g. the housefire relief effort), apply statistical methods to digital data, and package or pitch impact narratives in HTSG-1. Care and commitment, engagement, and good project planning are seen as more important. In Figure 6.11, HTSG-1 and Impact-1 are characterised as more direct, involving fewer objects of work and agencies, and as simpler or more "humble". Digital representations are not the medium of Impact-1 at HTSG-1. Engagement and project planning are more important than mobilising impact representations for distal exchange.

HTSG 1 & Impact 1
The older version of HTSG and their impacts relate to project management and community engagement

TOOLS
No digital tools, data management nor expert evaluation / marketing methods

SUBJECT
CEO, Project Teams
Board Members,
Office Staff

HTSG 1

OBJECT
Meet project goals, help communities, submit reports

OUTCOMES / IMPACT 1
Mothers & children change behaviour (e.g. dental hygiene)
Report compliance

RULES/NORMS
Fieldwork, engagement, project managing, phases, activities, goals

COMMUNITY
NGO funders, parent volunteers, village community, mothers & children

DIVISION OF LABOUR
CEO/intern office work,
Project team work, field work
Volunteers work with mothers & children
Mothers & children change behaviour

Figure 6.11: HTSG-1 activity system. Focus on engagement and project management

Impact-2 and HTSG 2

In contrast to Impact-1, Impact-2 involves technical representations embedded in data sheets and documents, which comply with legitimate evaluation methods for pragmatic decision-making results. These are shared with funders and other diverse partners who most often are some distance away from the target communities. Impact-2 aligns with: distanced forms of knowing and auditing (Strathern, 2000; Townley, 1995), information intensive work (Castells 2011, Ilcan & Philips, 2008), and development 2.0 digital flows (Thompson, 2008; Heeks, 2010; Quaggiotto, 2009). HTSG was trying to make diverse data/knowledge products for different consumers, each with different purposes or objects of work. These included development aid grant providers and established funders, wherein Impact-2 is a mobile data/knowledge asset, often strategically remediated in exchange for revenue (e.g. Hayes & Westrup, 2014: 28). However, consumers also included a local hospital, a public health authority, doctors, a local NGO network, and a national constitution amendment process.

As such, HTSG actually had multiple objects of work (Karanasios, 2018: 142-143; Engeström, 2008: 3; Blackler, 2009) in their data/knowledge sharing processes. These often conflicted with each other e.g. short-term results and compliance to show funders versus advocacy or long-term data-sharing relationship-building with local doctors and hospitals. Such aims contradict Impact-1 construction and displaced local participants for distal ones. These objects of work required new organisational resources, new roles, new attitudes to communications, more certainty, less doubt, and new systems and processes. These were not fieldwork activities, but office work (Strathern, 2000), which amplified plans and representations, reducing doubt, dialogue and contextual depth. Breastfeeding needed to be measurable, which is why Cherry adopted the food checklist as a strategy, rather than asking mothers about their work life and breastfeeding challenges[29].

In CHAT terms, impact activities transform the subject agent (Engeström, 1987: 103; Kuutti, 1999), producing a new subject in the activity system, transforming the old HTSG-1 into a new "HTSG-2". This is a tertiary-level CHAT contradiction and it fuels HTSG's transformation. The need to make measurable accounts of impact that could travel to partners and funders

---

[29] The checklist Cherry referred to is derived from the WHO (2010) publication, "Indicators for assessing infant and young child feeding practices: Part 2 Measurement".

required a different organisation of HTSG's resources, attitudes, skills, tools, systems and staff. HTSG-2 was assembling their impact machine. They were learning, but struggling to transform into the new data/knowledge intensive HTSG-2, which could master skilled data collection, storage and analysis, using digital processes and technologies, marketing attitudes, and effective impact packaging and pitching.

As HTSG staff were changing their goals in order to become expert impact data/knowledge producers, their organisation was transforming too. It was transforming from an organisation skilled at project management and tangible engagement towards data/knowledge intensity, data management and marketing. They were moving towards a new version of themselves, a new HTSG, that was skilled at impact construction, evaluation methods, data management, knowledge management, identifying and responding to funding opportunities, advocacy, communications, marketing and branding. HTSG-2 would be increasingly data/knowledge intensive, requiring new forms of expertise, digital technologies, and human resources. The new HTSG-2 contradicts the old one, requiring new tools, rules, attitudes, labour, outcomes, and relationships. Figure 6.12 shows the HTSG-2 activity system.
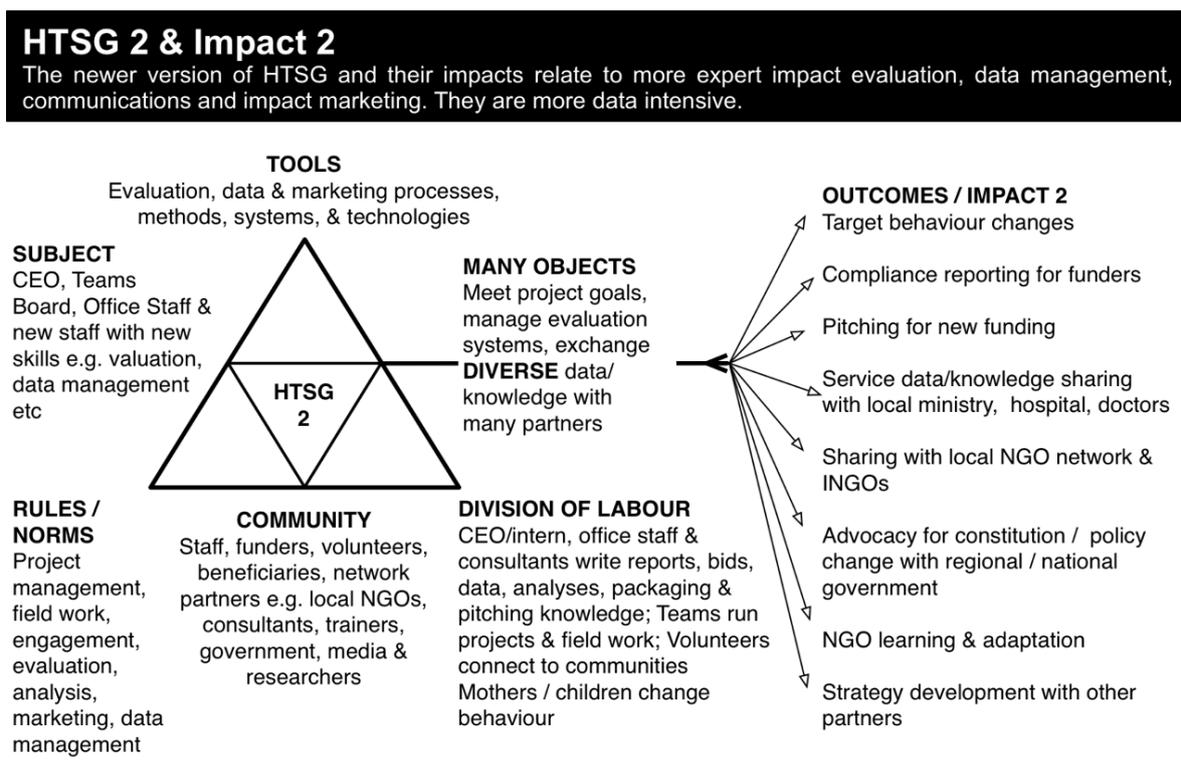
Figure 6.12: The activity system of HTSG-2, a data/knowledge intensive NGO

The objects of work around evaluation were multiple and challenging for HTSG, as they learned about evaluation data/knowledge production and sharing with diverse partners such as funders, the local hospital, or the national constitutional reform process. Each potential consumer of their data and evaluation work had particular requirements, whether for child rights advocacy, healthcare data, or demonstrations of project success, which meant, in CHAT terms, that HTSG had multiple objects of work (Karanasios, 2018: 143-144). HTSG staff had to make multiple impact products. Being novice evaluators, meeting such a variety of data/knowledge exchange needs exacerbated their struggle. They were not only making knowledge products (Mosse, 2004a: 77) for funders in the aid chain (Wallace et al, 2006; Banks & Hulme, 2015), but for a diverse range of consumers. HTSG was learning and struggling but also expressing its agency in a wider network of activities. This struggle pushed HTSG-1 from community engagement towards HTSG-2, impact marketing and data managing.

In sum, the CHAT activity system analysis shows that a primary contradiction and a tertiary contraction were evident and significant in the changes HTSG was experiencing over the course of the research period. The primary contradiction was between two forms of impact: one humble, negotiable and locally situated; one documented, mobilised, and commoditised for aid networks, markets and bureaucracies, more rigorous and made up of data/knowledge intensive digital representations. The second contradiction was between the old HTSG and the new one. The old NGO was organised around field engagement and project management; the new one around data/knowledge intensity, data management, strategic use of ICTs, and configurations to market effectiveness and success. These contradictions were relatively silent and normalised (Townley, 1995), hard for HTSG to problematise, and unobtrusive (Blackler, 2011), effectively submerged under the need to efficiently market impacts.

6.4.2 Temporal activity chains, submerging and elevating

The transformation of HTSG played out in a series of activities wherein impact data and knowledge incrementally move and are mediated. These temporal sequences offer a way to parse and understand a response to research question 2: How are power relations generated during the construction of impact evaluation data and knowledge? Three significant areas are addressed in this section: firstly, the diversity of activities beyond the TIEK meta-model of inputs/outputs or the DIKW logic of objective data for rational decision-making; secondly, the

temporal chains that highlight the times in HTSG's evaluation activities where different power relations are submerged or elevated; and thirdly, the implications from the case for how evaluation is performed, how small NGOs are transforming, and how broader aid sector power/data/knowledge relations remain silent.

Firstly, the activities of impact evaluation data and knowledge construction are more diverse than those prescribed in TIEK or DIKW decision support models. Evaluation activities in the case involved more than just the demands and supply of rigorous methods or results for decision-makers. DIKW-derived models argue that knowledge is made from objective data (Davenport & Prusak, 1998; Earl, 1994), and the results of knowledge production are better, more rational, and more pragmatic for informing decision behaviours (Davenport & Prusak, 1998; Feldman and March, 1980;). Similarly, impact evaluation models highlight key steps in their prescriptions, such as evaluation design, data collection, analysis and deliverables, prescribed through many aid industry models (UNDP, 2002; UN Women, 2015; 3ie, 2012; DFID, 2012; Stern, 2015; Groves, 2015).

In contrast, the small resource limited NGO was in the middle of diverse impact-related activities. These included sourcing evaluation methods, tools, concepts, interns, staff, and ad hoc advice. It further included promising impacts and evaluation approaches to win grant applications, imaging a *"full-fledged system",* adopting new organisational processes like data management, and impact packaging for reports. In carrying out these activities the NGO was struggling to embrace new attitudes like marketing, taking credit for impacts, expunging doubts, and brokering opportunities to share or exchange impact data/knowledge in multiple partnerships. This case shows how impact evaluation data/knowledge work goes far beyond technical evaluation or knowledge management prescriptions. Even struggling to evaluate prompts diverse transformations (Kelly, 2018).

In response to research question 2, these temporal chains exhibit locations of submergence and elevation where HTSG was learning how to produce Impact-2. Similarly, HTSG was learning how to assemble an evaluation machine and how to sequence data/knowledge activities for strategic outcomes, all the while intensifying their development 2.0 resources and adjusting their marketing orientation. These activities materially transformed HTSG.

Figure 6.13 shows the temporal chain along which HTSG were building an impact machine to construct data and demonstrate impacts. The figure shows diverse evaluation activities, from

assembling the impact machine, to making cells, to packaging impacts. Diverse objects of work (Engeström, 2008: 43; Karanasios, 2018: 143-144) are shown at the end points on the bottom right of the diagram. The multiple objects (e.g. complying with funder needs, pitching for grants, advocating change) exacerbated HTSG's orgnaisational struggle and learning processes.

HTSG's Evaluation Temporal Activity Chain
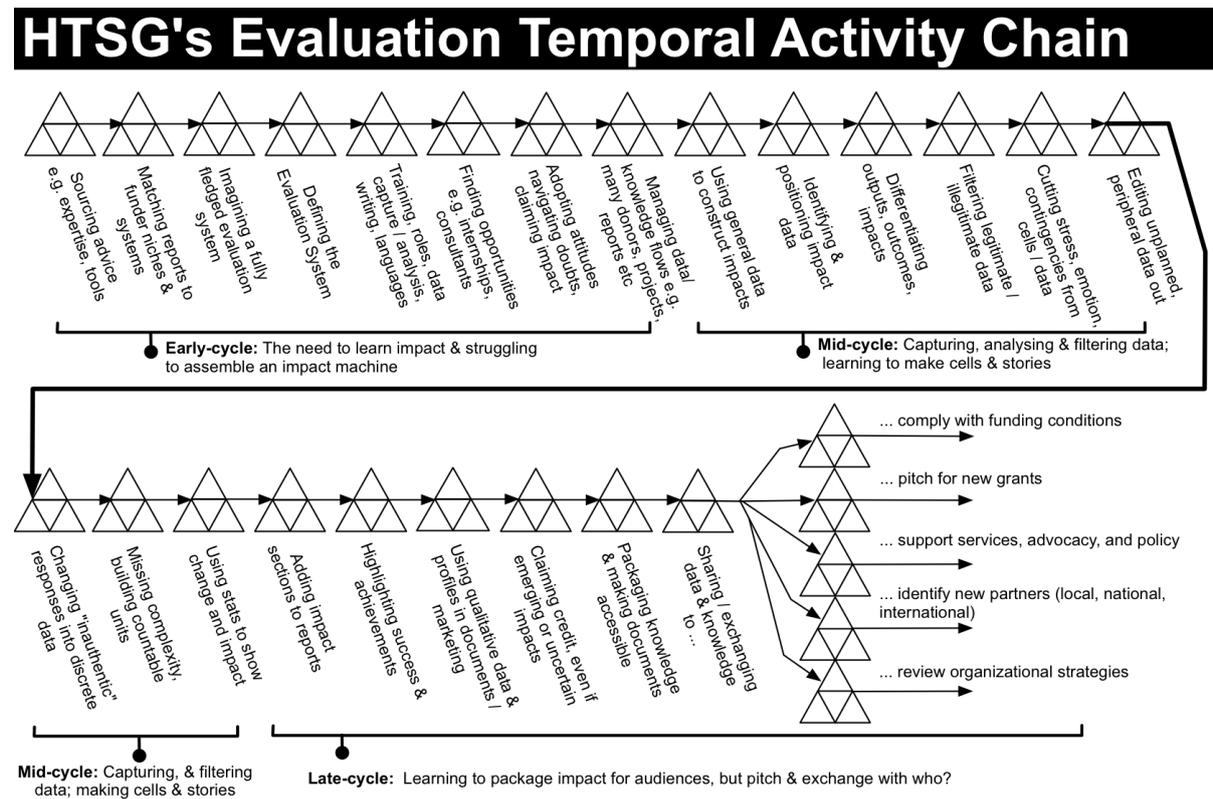
Figure 6.13: HTSG's temporal chain of impact evaluation activities

There is no doubt activities may move around, but the main insight gained from the chain is that learning and doing are intertwined and more extensive than portrayed in prescriptive evaluation frameworks, results-based management, or data/knowledge management models.

Figure 6.14 below illustrates the submerging and elevating episodes along the temporal chain at HTSG. Many issues are submerged, such as doubts, stress, lack of capability, inadequate processes (e.g. bad data management and losing laptops), impacts outside of project scope (the housefire response), or the humble attitude to impacts. These are not legitimate elements of impact evaluation, and are silenced or elided in reports, data sets, or impact data/knowledge exchanges. In contrast, other elements are elevated, such as available data sets, statistical

expertise, personal stories of successful change, new partnerships, and hopes to establish a full-fledged system.
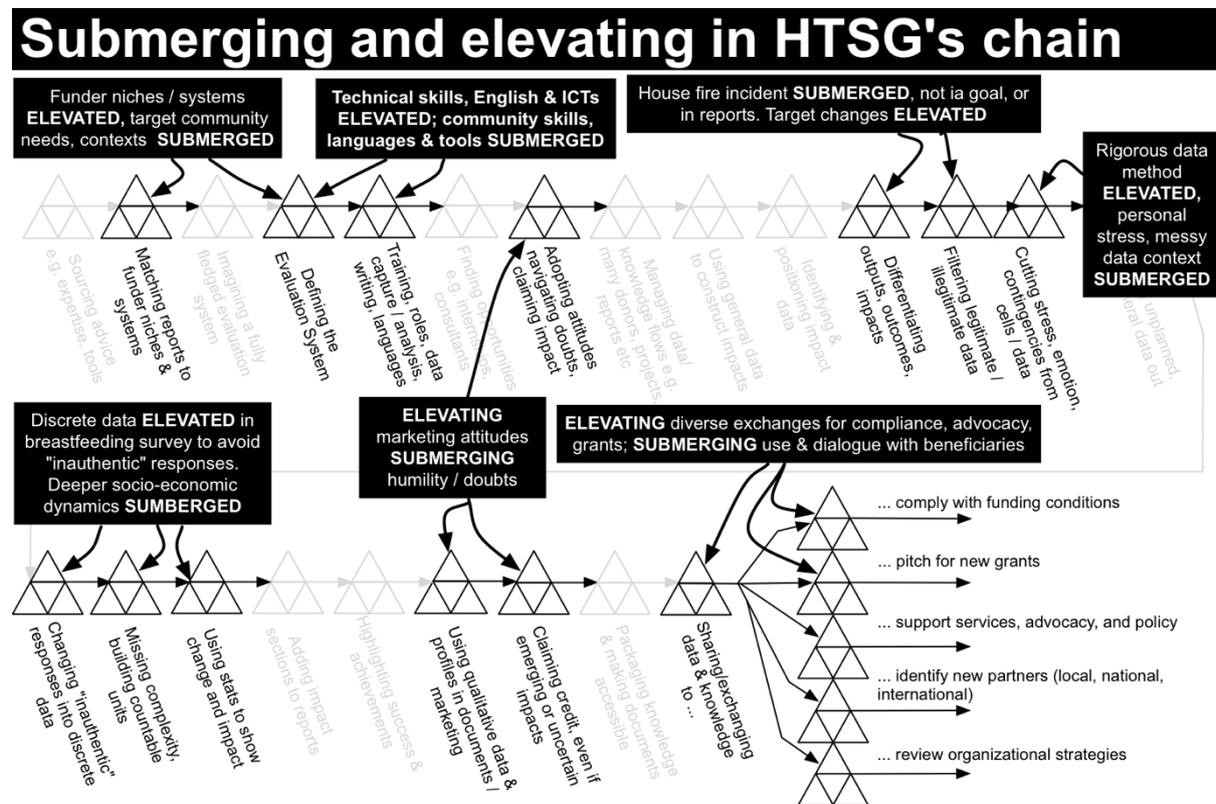
Figure 6.14: Submerging and elevating in HTSG's evaluation activity chain

HTSG wanted data to show changes in target community behaviours concerning healthcare, such as dental hygiene, breastfeeding, giving birth in public hospitals, and concerning new services offerings in the Migrant Project. HTSG were partly aware of the need to perform this submerging and elevating work, but had not yet mastered all the processes, resources, skills, people, technologies and marketing attitudes required. They were appointing new staff with new skills, but not selling or pitching or taking credit appropriately for impacts. They had begun to pitch and broker with local health authorities, but had not consistently embedded data sets, models or narratives into reports, data sharing, or funding pitches.

In summary, the temporal chain of activities illustrates the multiple objects of work, sequences of activities and elements for building an impact evaluation machine, and the process of submerging and elevating different aspects of that work. Neither DIKW nor TIEK models focus on such detailed accounts of intertwined, but temporally performed practices and power

dynamics, how certain knowledge, skills, and attitudes are amplified, and how others are elided or marginalised. HTSG were learning how to perform these elisions, omissions, agential cuts and bifurcations. But their learning was not complete, meaning they were not yet fully transformed into HTSG 2, a new version of themselves which could expertly and instrumentally market and exchange Impact-2 knowledge products. They still prized humility over marketing, lost data through bad data management, and mispresented impact accounts in the wrong places in reports. HTSG exhibited a lack of certain capabilities in the areas of data capture or data analysis, evaluation report writing skills, and software customisation. They lacked both technical evaluation knowledge and practical experience.

To be successful, HTSG had to incrementally learn to edit goals, apply concepts, designs tools, legitimate data, package knowledge, and pitch products. Such learning and change would help them to gain funds, status, capacity and new partners. It would also transform their organisation. Community engagement and project planning, areas they were already expert in, are not the same as constructing impact products for aid markets and bureaucracies (Quarles van Ufford, 1998; Mosse, 2004a), or constructing exchangeable data/knowledge assets (Blackler, 1995).

6.5   Summary of implications

The case has significant implications for evaluation, NGOs and power/data/knowledge relations. Implications derive from several areas: the activity systems, contradictions, temporal chains, and HTSG's learning and struggling. Indeed, HTSG developed new capacities (e.g. to do evaluation, to make data/knowledge for partners), but faced limitations too (e.g. their attitude to marketing, lack of methodological expertise or data management capacity). In wider development sector networks, broader patterns of power and control around financial flows, governmental inequalities (Hayes et al, 2017) and developmentality (Lie, 2015; Ilcan & Phillips, 2010) elevate certain kinds of impact data/knowledge, submerging others. These concern firstly how evaluation itself is changing, secondly how HTSG as a small NGO is transforming, and thirdly how power and politics are reconfigured in these data/knowledge intensive relations. A summary of these implications completes the case.

The kind of impact evaluation that HTSG is in the process of learning involves capturing discrete data, putting it in digital tables, sheets and narrative profiles, carefully placing these in reports, brokering data exchanges, and representing impacts to diverse partners. Although clearly present, activities involving discussion, negotiation, dialogue and engagement are less clearly represented in the NGO's learning process, in contrast to activities involving design, capture, analysis, reporting and pitching.

During the research, the author influenced HTSG's learning by encouraging the use of open questions in data collection, by foregrounding the need to talk to a diverse network of stakeholders, and by supporting the Migrant Project grant bid. These influences harboured attempts to make HTSG's evaluation work multi-voiced and collaborative. However, the evaluation process remained wedded to learning how to capture, process and exchange. The view here is of data/knowledge as products (Mosse, 2004), constructed for exchange with institutional funders or partners, not the processes, conflicts and negotiation with beneficiaries around contexts of poverty or marginalisation. HTSG saw their learning as leading towards a view of knowledge as assets (Boland, 1987) for commodification (Blackler, 1995). This underpinning perspective positions impact evaluation increasingly as expert mechanisms to make products to pitch, not evaluation as a process or dialogue, impactful itself on stakeholders and participants through meeting, listening and discussing changes or political concerns. Trading impact as products (Impact-2), not processes, appears to reconfigure small NGOs like HTSG away from negotiations and talk, but toward distal institutional marketing.

As a small organisation operating in the development sector, HTSG was transforming into a new version of itself. This transformation was in process during the time of the research, but far from complete. HTSG were becoming data/knowledge intensive, entering the development 2.0 landscape by appointing staff with appropriate skills to new roles, adopting new ways of working that incorporate evaluation design, data management, data analysis and digital technologies, and learning practices of representation, marketing and communications. At HTSG this was a move away from local community engagement (Strathern, 2000). It was also a move that HTSG's CEO saw as important for making HTSG a viable competitor in the *"challenging funding environment"*. These shifts align with the opportunities of development 2.0 (Thompson, 2008; Heeks, 2010), but also the dangers of developmentality, narrow information profiling and technical knowledge networks (Ilcan & Phillip, 2008; 2010).

If HTSG did not fully transform to data/knowledge intensity, it risked being branded as old, unprofessional or outdated, unable to capture data, analyse, market, articulate its impacts, or project digital representations of success. The activity systems and contradictions help explain the NGO's shift towards making Impact-2 products, and the organisational capabilities required, i.e. as HTSG-2. The temporal activity chains illustrate how diverse activities coalesce in the production of impact data/knowledge, and how certain elements are incrementally submerged or elevated.

What became progressively more important as HTSG established data/knowledge intensive capabilities and became more and more skilled at impact evaluation, was the rendering of data/knowledge products which could be shared or exchanged, in more or less competitive markets (Hayes et al, 2017), and the change this produced in agents performing evaluations. Perhaps surprisingly, the change in data/knowledge practices of an organisation learning to play in impact data/knowledge markets is more pronounced than the change in target community behaviours. At HTSG who was transforming more – breastfeeding mothers or the NGO learning to evaluate? Either way, the latter change was elided in evaluation representations. Evaluation changes NGOs; it impacts its agents.

HTSG's political and power relations were also impacted by their increasing yet incomplete mastery of mundane evaluation data/knowledge intensity. Such limits are normally disregarded in TIEK. They involve who is shaping evaluation activities, and who has the ability, will, and legitimacy to influence, govern or even participate in evaluations. These power relations shifted as HTSG gained mastery and influence, using impact data/knowledge to influence public health authorities, hospitals, doctors, migrant services and funders. The standard impact evaluation gaze upon target communities and their behaviour changes, in the areas of dental hygiene or breast-feeding, for example, at HTSG, precludes discussion of how increasing mastery over evaluation data/knowledge also changes broader networked power relations. An assemblage of data volume and diversity (Feldman & March, 1981), packaged knowledge products (Mosse, 2004a), expert legitimacy and reputation (Escobar, 1995; Townley, Ilcan & Phillips, 2008), and success narratives (Kelly, 2018) all constitute strategies of influencing. HTSG were learning how to desire, capture, construct, bifurcate, legitimate, package, pitch, exchange, market and strategise impact data/knowledge.

Such power/data/knowledge relations are part of institutional governing processes (Knights et al, 1993; Introna, 2003; Hayes et al, 2017), and the professional knowledge flows within and

between development organisations (Mosse, 2004a; Quarles Von Ufford, 1998). However, what is most interesting at HTSG is the moment of the research, when HTSG were in the middle of learning, mastering and struggling to make the move to data/knowledge intensity. In learning to construct legitimate impact data, and knowledge, HTSG were learning how to erase other kinds of data and knowledge, such as the housefire incident, that fell outside of the formal project's visible space. Cherry was learning to fit empirical, messy data into statistical cells, while keeping stress and deeper contexts of poverty outside of those cells. Orr was struggling to edit doubts, imaginary plans, and humility out of impacts.

HTSG was learning to treat doubts, dialogue, humility, stress, deeper contexts, and dialogue as risks, pests or vermin[30], or superfluous contingencies which could damage impact-product market value. In this way doubt, local needs, power, politics, and the representation of villager lifeworlds were submerged. They were displaced by the elevation of evaluation expertise, discrete data representations, inauthentic profiles, project goals, and successes claims. These edits, cuts and additions along the chain were mediations performed by HTSG as it learned to evaluate. Power inequalities and anti-politics (Ferguson, 1990) risk proliferation when evaluators, NGOs, funders, states and aid markets govern the science and intervention space without humility, dialogue, critique, or reflection on what their evaluations do, and what and who are submerged or elevated.

In sum, the analysis reveals three significant case implications.

I. *Shifting Evaluation:* The shifting legitimacy of evaluation methods towards technical, mobile, multiple-purpose, commodifiable, marketable, exchangeable, and digital impact products. This shift is from Impact-1 to Impact-2 in the CHAT analysis.

II. *Shifting NGOs:* The transformation of small resource-limited NGOs, and potentially other agents, towards data/knowledge intensity and development 2.0 capacity, to catch up with better resourced data/knowledge intensive competitors (e.g. global INGOs). This shift is shown in the HTSG-1 to HTSG-2 transformation in the CHAT analysis.

---

[30] In Scott's (1998) discussion of the invention of forestry science in eighteenth century Prussia and Saxony, fruits, nuts, twigs, medicines, animals and other items in forests were discounted as valuable, even when they had value for local people. The fruits and foliage came to be seen as vermin, pests, threats or as irrelevant to scientific understanding or the market value of timber.

III.   *Shifting Power and Politics:* The shift to learning, data and expert knowledge as sites of power and politics in aid markets. HTSG were trying to pivot to data/knowledge intensity to compete in unequal funding markets. However, they faced a series of silent yet mounting development 2.0 capacity gaps, governmental inequalities (Hayes et al, 2017), and developmentality challenges (Lie, 2015; Ilcan & Phillips, 2010).

As they were learning evaluation, it was not clear whether HTSG would go on to become instrumental or egalitarian in its data/knowledge production. When Orr asked in the workshop *"I mean when we make project plans, … are we just imagining our impact?",* she was critically reflecting on these shifts of power and participation. The next chapter builds on the lessons from HTSG and asks how we can respond to the bigger aid sector picture of impact participating, impact governing and impact marketing. Can there be a response to impact power/data/knowledge configurations that amplifies critical sensitivities within evaluation supply chains? What might such a response look like? This goal challenges the incarceration of evaluation data/knowledge by an older, deeper, more established audit culture (Strathern, 2000; Townley, 1995; Porter, 1996), by inventing a newer, more critical, yet mobile audit gaze; one that re-embeds power and practice intimately within data/knowledge supply chains.

7 Discussion

7.1 Introduction

In responding to the research goal of making power/data/knowledge relations more explicit in development impact evaluation, this discussion chapter offers three contributions. The contributions speak to concerns about the narrowing of impact frames and evaluation forms, which accompany development 2.0 data/knowledge intensities (Thompson, 2008; Heeks, 2010; Quaggiotto, 2009; Kelly, 2018). Each contribution posits evaluation processes as impactful in themselves, in contrast to the TIEK focus on justifiable methods (scientific or participatory) and representational results (products) for decision-makers. The three contributions thus encourage a view of "evaluation-as-practice". The practice perspective (e.g. Miettinen et al, 2009; Reckwitz, 2002; Nicolini, 2013; Spender & Scherer, 2007) constitutes an alternative to evaluations that privilege scientific truths, prolong business pragmatics, or signal participatory concern. The shift to practice illuminates power/data/knowledge relations, following diverse critical works (e.g. Foucault, 1980; Escobar, 1995/2011; Avgerou, 2002; Walsham, 2000; Hayes & Westrup, 2012; Hayes et al, 2017; Kelly, 2018).

Each of the three contributions combines provocation and pragmatism. Each is accessible, using diagrams and metaphors, but locks in critiques of power and practice, thus making them somewhat jarring. The contributions are positioned between critical scholarly work and more accessible tools that professionals in development, evaluation, ICT4D or KM4D can refer to. Each informs, probes (Gaver et al, 1999) or stimulates (Salancik,1979) reflection on power/data/knowledge relations in evaluation practices.

Section 7.2 describes the first contribution, the "Impact Iceberg". The iceberg adds to the body of evaluation literature. It is a diagnostic tool for sense-making, a big picture metaphor for making critical configurations audible and impact products questionable.

The second contribution is conceptual, responding to questions about TIEK and the DIKW legacy of eliding power and practice in data/knowledge work. The contribution is a set of "Audit 2.0" devices, labelled in reference to the evaluation trust in and awe of numbers (Porter, 1996; Townley, 2003; Miller, 2004) and the penchant for digital versioning still evident today. Audit 2.0 contrast with "Audit 1.0", where established audit cultures stress measurement, surveillance, control, and normalized products for market exchanges or governing rationalities (Hayes et al, 2017; Ilcan & Philips, 2010). The audit 2.0 devices described in section 7.3 are:

- data/knowledge chains and networks
- the 6P data/knowledge sensitivities
- the impact spectrum
- datamentality and datamateriality

Section 7.4 describes the third contribution, "Critical Engagement", which is a response to methodological questions about bridging critique and engagement. It reflects on lessons from the case NGOs regarding how critical perspectives translate or can be brokered into sites of normalized evaluation practices. Critical engagement draws on knowledge brokering (e.g. Lewis & Mosse, 2006; Shucksmith, 2016) and CHAT expansive learning (Engeström, 1987; 2001). However, a key difference from other forms of brokering is that it requires reflection and dialogue about power and practice, and also a gaze wider than typical intervention target sites. Tensions and tactics are described to inform researchers and practitioners who wish to articulate and test critiques during engagements with research partners. The contributions support collaborations between critical academia and development sector practitioners, by offering concepts upon which critical insights can be tested and trialled in the wilds of development evaluation practice [31].

---

[31] As a caveat, certain camps in critical academia and the aid industry will reject these contributions. Practitioners under demands to exchange evaluation results and deliver impact products will say the contributions are too academic, idealistic, or not pragmatic for knowing "*what to do on a Monday morning*" (e.g. Davenport and Prusak, 1998: xi). Scholars who are skeptical of collaborations with the sector will see the contributions as conformist, diluted, insufficiently disruptive (e.g. Burrell, 1993), or best kept to "*one's own society*" (Ferguson, 1990: 286).

7.2 Contribution 1: The impact iceberg

7.2.1 Configuring the iceberg

An argument has been made that the combination of increasing data/knowledge intensity in development, NGO, and evaluation work, heavily fuelled by development 2.0 (Thompson, 2008; Heeks, 2010), audit culture (Strathern, 2002) and developmentality (Lie, 2015; Ilcan & Phillips, 2010) has led to a contemporary aid chain (Wallace et al, 2006) in which demands for impact data/knowledge and technical evaluation prescriptions configure evaluation processes and products. This argument involves many moving parts, diverse stakeholders, information systems, data/knowledge products and processes, and power dynamics, which play out in contemporary development 2.0 landscapes. One way of diagnosing problems within these relations is through an impact iceberg metaphor, an accessible big picture diagnostic that can stakeholders can refer to, and which prompts them to address impact silences.

Taking critiques into the wild, so to speak, means avoiding the trickle-down notion that critiques will be welcomed and adopted unproblematically by individual evaluators, ethical, intelligent professionals, or their organizations (Picciotto, 2013). Critiques, like impacts or wealth, may not trickle down (Harris, 2016: 177). They may need pushing, questioning, amplifying, linking with established movements, rejecting, or brokering with those working in development or evaluation. The iceberg aims are critical, but also mobile in acknowledging (if not resolving) the penchant for travelling rationalities (Craig & Porter, 2016), tools and concepts which development managers, evaluation professionals and researchers can use when there is a need to deal with issues concerning evaluation power and practice.

In Figure 7.1 below, the iceberg is presented in the middle of various development impact evaluation configurations, labelled A to F. There is a circular flow around the iceberg that helps construct it, repeatedly, how audible products become elevated, and silent practices and transformations submerged. The areas are described as follows:

- **Area A:** Historical trends in the sector, which shape evaluation practices

- **Area B:** Specific sets of contemporary demands on evaluators and NGOs which pressure them to construct and exchange impact evaluation data/knowledge

- **Area C:** Agents who perform or are involved with evaluations, such as small NGOs, evaluators, donors and communities, who have unequal data/knowledge capacities

- **Area D:** Technical evaluation discourse or TIEK, which responds to demands by prescribing methods and results frames

- **Area E:** Audible, above water, legitimate impact products (e.g. efficiencies, performance, causes, participation, certainties etc) which are foregrounded in reports, arguments or infographics etc

- **Area F:** Silent, peripheral or illegible evaluation relations under the waterline, concerning how evaluation, NGOs and impact governing are configured and changing

- **Area G:** Key dynamics underpinning the impact iceberg circulations e.g. audit cultures, governmentality, developmentality, datamentality and development 2.0
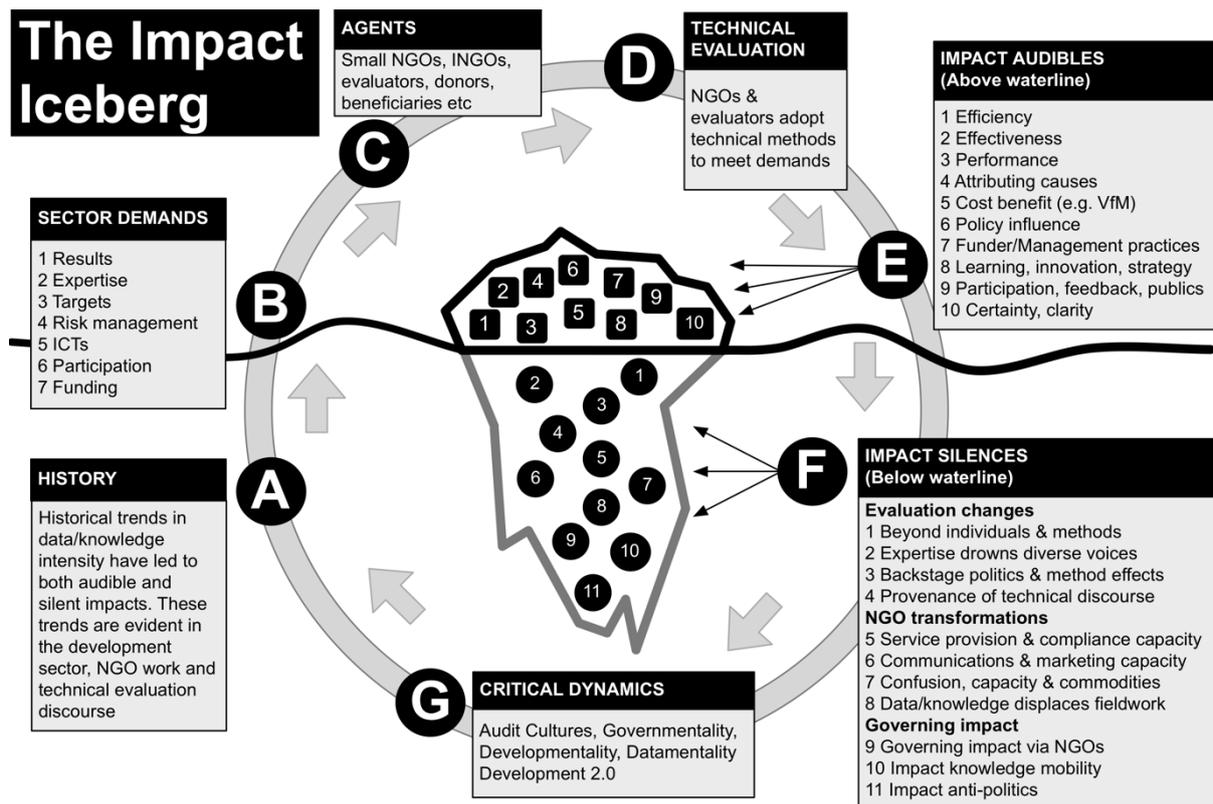
Figure 7.1: The impact iceberg and its key mechanics

Figure 7.1 area G lists some of the critical dynamics of the iceberg identified in the study. Governmentality in included, in terms of how TIEK submerges power and politics in favour of professional expertise to shape impact products for markets (Ferguson, 1990; Bächtold, 2016). Developmentality is implicated regarding knowledge networks (Ilcan & Philips, 2008) and participation/partnerships as strategies of control from a distance (Lie, 2015).

A new concept, "datamentality" is also added, which highlights the idea that objective factual data, efficient information systems and data management capacities are used to justify and produce empirical results, without accounting for power or practice. The objectivity of data claims is crucial in this view, but highly problematic. Datamentality is described in section 7.3. Suffice to say here, that digital data and information systems have become key dynamics in today's impact iceberg. Those NGOs or evaluators without a digital, data driven evaluation machine risk appearing uncompetitive and unprofessional. Such claims to objective empirical data, digital systems, and expert evaluation methods, mark control of impacts as professional, rather than community, concerns.

7.2.2 Above the waterline

On the surface, impact evaluation processes and methods are used by development organisations such as small and large NGOs (Area C on Figure 7.1) to respond to diverse sector demands for impact data/knowledge results (Area B). Thus, impact evaluation produces desired, legitimate, audible results that focus on accountability, efficiency, effectiveness, causal attribution of impact, participation and learning, organisational decision-making processes, such as funding allocation, cost benefit analysis, policy influence or program scaling. These audible evaluation deliverables feed into organisational strategy, learning, participatory development, feedback to target communities, and informing publics about impacts. These reports, data sets, charts, infographics, press releases and arguments (see Williams, 2014) are data/knowledge products which mobilise narratives of success, certainty, rigor and expertise, empirical objectivity, organisational capability and innovation. These were evident to different extents in the NGO cases. Each audible is described below.

Audible 1: Efficiency

Impact evaluation processes place value on results which demonstrate efficiency (e.g. Puri et al, 2015). New development management (Dar & Cooke, 2008), approaching aid as a business (De Haan, 2009), and auditing mechanisms (Strathern, 2000; Harper, 2000: 21-54) value efficiencies, as do accounting methods, such Cost Benefit Analysis (CBA), or Value for Money (VfM). Fast, legible, editable and easily transported statistical data, facilitated by ICTs and information systems, tablets and web databases, were prioritized by experienced managers in the NGO cases in order for them to quickly know "what's really happening". Clamour, data loss, qualitative narratives, engagement and context were expensive, inarticulate, illegible or time-consuming impediments to efficiency. Efficiency is one desired, audible evaluation good.

Audible 2: Effectiveness

Demonstrating effectiveness is another evaluation good (e.g. OECD DAC, 1991: 5), based on aligning goals and results, through plans, indicators, reports and tools such as logframes (Kerr, 2008). Although many farmers may not know their expenditures, spreadsheets columns in the first NGO case showed before/after financial impacts, demonstrating intervention effectiveness. The novice NGO wanted to demonstrate effectiveness to strengthen grant bids in a competitive funding environment, but they were unsure of how, or whether, to market their effectiveness. Information Systems, ICT4D and Development 2.0 rationales argue for the development effectiveness of new technologies (e.g. Heeks, 2008), how ICTs can make "*a better world*" (Walsham, 2012). As with efficiency, effectiveness can be elevated via auditing (Porter, 1995), but this risks damaging longer-term goods like trust building, learning, engagement (Wallace et al, 2006), and more power sensitive ways of evaluating.

Audible 3: Performance

Performance management is perceived as a third evaluation good (e.g. Roche, 1999: 7; Townley et al, 2003) following effectiveness and efficiency, and all three are instrumental in new public management (Hood, 1995). If organisations, projects, departments and individuals are performing well, then this suggests interventions are successful. However, at the expert NGO performance was not measured, and at the novice NGO only volunteers were asked to measure their own performances. Neither NGO had a performance management system.

Performance issues were evident, for example when the NGO director in case 1 did not want to waste time on qualitative evaluation, or when the novice evaluator in case 2 blamed herself for not aligning field data with statistical standards. Her wasted time and stress were not counted, measured nor usable in governing processes. Performance management was not significant in the cases, suggesting that either not all impact audibles are equally elevated, or that smaller NGOs may not, yet, feel pressured to mobilize performance metrics.

Audible 4: Attributing cause and effect

Attributing cause and effect is a scientific activity, to which impact professionals and researchers assign much credence (3ie, 2012a; 2012b; Roche, 1999, Picciotto, 2012). Cause-effect evaluation work is a key part of many methods and approaches in evaluation, but particularly in the experimental, quasi experimental and statistical genres (Mohr, 1995). Yet surprisingly, neither NGO articulated causes or effects. The first NGO and funder were expert in certain practices of evaluation i.e. designing, data management, bifurcating data, pitching to investors, but they were not expert in methods, or the science of evaluation. Their director talked about model villages and engineering a 10x impact effect, but used metaphors, not empirical statistics. What was more important to him was pitching "*a convincing story*". Such stories may vary according to consumers and may not require robust cause-effect demonstrations as prescribed in TIEK. Significantly, causal attribution may not be as influential in some funding markets or management practices, as they are with evaluation researchers or scholars.

Audible 5: Cost benefit analysis

Cost benefit analysis has been a part of evaluation discourse for over a century (Porter, 1996). It is prominent in development, with sub-discipline applications in areas like ICT4D (see Heeks and Molla, 2009) and derivative methods such as Value for Money (VfM) (e.g. DFID, 2011). In the study, one NGO was a novice evaluator and did not discuss costs or benefits, but the expert, experienced NGO did collect project and farmer income/expenditure data. Given today's skeptical aid systems and markets (Picciotto, 2012) and the vicious evaluation circle NGOs find themselves in trying to prove impacts (Roche, 1999), CBA offers relatively

unambiguous, robust and mobile accounts of impact. It is quite feasible that small NGOs will opt to learn and use such methods more in future to meet VfM agenda. Nevertheless, adopting CBA or VfM models implies training, trials, and learning, much data/knowledge work and thus the opportunity cost of not spending more time and money in the field with communities to understand what they see as costs or benefits. Global economic criteria and methods such as CBA (Porter, 1996) can submerge local criteria, methods and participation.

Audible 6: Policy influence

Data, knowledge and evidence have strong relations to the perceived good and audible result of influencing policy (Parkhurst, 2017). This is evident in much industry literature, despite anthropologist critiques about the policy-practice direction (Mosse, 2004b). The London School of Economics runs a widely read blog resource which regularly highlights the problems and connections between evidence and policy[32]. Therefore, it was a surprise to see little in the expert NGO's case regarding policy influence. In contrast, the novice NGO were using their evaluations work and advocacy with a network of local NGOs, a municipal government, and the national Thai government. The larger learning here is that more commoditized and instrumental production of impact sales pitches, evidenced at the expert NGO, does not necessarily translate into policy orientations, networks or influence. Using evaluation evidence for policy influence is a legitimate evaluation audible, but not strategically prioritized by all NGOs, based on their expertise.

Audible 7: Funder and management practices

TIEK privileges the use of evaluation data for decision-making practices involving managers and funders, around such issues as fund allocation, funding strategies, comparing projects, scaling solutions, forming best practices, and modelling (e.g. Roche, 1999; Mohr, 1995; Puri et al, 2015). Both NGOs were using evaluation findings in such organisational practices. The expert NGO was using impact data to scale solutions across scores of villages, and also to

---

[32] See LSE Impact of Social Science Blog, at: http://blogs.lse.ac.uk/impactofsocialsciences/

design an evaluation management information system (MIS). The use of evaluation data to scale, model, build poverty verticals and an architecture for an evaluation MIS is a business aim beyond a specific evaluation site or community. The novice NGO had diverse funders and short-term projects which restricted scaling, but they were sharing learnings across sites and using data to build partnerships with other NGOs, the public sector and professionals, such as doctors. What is notable is the ongoing transformation of NGOs as they take impact data and learning into these new organisational spaces and markets. As with Porter's CBA genealogy (1996), this goes beyond the initial aims of evaluation. Methods have consequences, transform evaluation as a discipline, transform NGOs capacities, and transform power relations.

Audible 8: Learning, innovation and new organisational strategies

Development evaluations prioritize learning and identification of further problems to frame and fix (Norrish & Sayce, 2006; Ferguson, 1990; Li, 2007). As with the previous audible, such learning is impactful beyond specific evaluation contexts. It influences new problem definitions and wider strategies. New development approaches, such as Problem Driven Iterative Adaptation (PDIA) (Andrews et al, 2013), Adaptive Management (National Research Council, 2004; Rondinelli, 2013) and Doing Development Differently (Wild and Andrews, 2016; Prieto-Martin et al, 2017) embed learning, monitoring and evaluation at the centre of strategy. In the cases, the expert NGO used evaluation, data management and knowledge of innovative "*poverty verticals*" to design their MIS. The novice NGO used evaluation data/knowledge in new ways, not limited to target evaluation sites e.g. brokering partnerships with hospitals, national advocacy, and designing services for marginalized migrants. Such learning, strategic thinking and innovative orientations deploy evaluation results and experiences to extend relationships, expertise and market opportunities. Feldman & March (1983) identified such diverse uses of information beyond decision-making in the 1980s. The cases suggest impact data/knowledge flows outside evaluation spaces and prescriptions, into wider organisational learning, innovation and new strategies – feeding development 2.0.

Audible 9: Participation, Partnering, Feedback, Informing the Public

Despite being largely technical products, impact evaluations are increasingly under pressure to be informative for public audiences, feedback to target groups, or to demonstrate participation with beneficiaries. Such beneficiary and public knowledge sharing varies across different evaluation paradigms (Macdonald, 1993; Norris, 2015) and methods (Duflo and Kremer, 2005; Picciotto, 2012; Groves, 2015). Above waterline audibility is also boosted if data/knowledge can be used in strategic communications e.g. branding and reputation management.

In the first case, participation, feedback and informing the public were not priorities. Impact data/knowledge was targeted toward instrumental exchanges, sales pitches and funding fairs. The partners did not feel pressure to inform publics or promote participation. In contrast, the second case NGO had published a photo-essay on a day in the life of one beneficiary mother on their website and in a local newspaper. With the author, they were also starting a participatory evaluation process with marginalized migrants, to identify service needs and solutions. Nevertheless, their CEO wanted statistical measures for grant bids, even though staff had doubts about impacts, public relations and marketing. The cases suggest that demands for participation and informing the public are linked to funding streams. The expert NGO had a long-term funder and did not need to demonstrate participation. The novice NGO, competing in uncertain grant markets, was attempting to market and promote participation, but was unsure how evaluation could support such goals.

Audible 10: Certainty and clarity

One of the most important above waterline audibles is the conveyance of certainty and clarity. This is well commented on in critical works (e.g. Scott, 1999; Law, 2004). It was also visible when the expert NGO used scientific, data management, development expert jargon, and marketing pitches to package solid results, as more objective, modern, digital, convincing, expert and robust. In contrast, many elements were submerged e.g. doubts, clamour, illegible (to evaluators) and thus inarticulate farmer lifeworlds, data loss, wasted data, competitor claims, diverse funder frames (e.g. templates, online systems), confusion (about evaluation aims, samples, techniques, questions), NGO survival needs, evaluator struggles and stress, deeper poverty complexities, humility and engagement. Although these are results of evaluation-in-practice, from the learning and doing of evaluation, most are not legitimate in grant markets, professional bureaucracies, technical prescriptions, or influential decision-

maker forums. Certainty amplifies expertise and impact governing techniques of calculation (Hayes et al, 2017), strengthening the exchange value of impact products, submerging humility and dialogue (Blackler, 1995). In making what counts count (Miller, 2004: 181), the uncertainty of the making itself is submerged, and the confidence and clarity of experts is elevated.

7.2.3 Below the waterline

Below the waterline are the impact silences. These are impacts, practices or configurations that are not easily rendered into legitimate products. They are beyond the technical targeted methods and results frame, and they challenge the demands, prescriptions and certainties of audible exchange values above the waterline. Efficiency can submerge participation; certainty can submerge doubt or complexity. Imported technical methods, such as RCTs, can subjugate local dialogue. The silences relate directly to the critical configurations encountered in the literature review, developmental and unequal partnering strategies and knowledge networks (Lie, 2015; Ilcan & Philips, 2010), and the implications at the end of the NGO case chapters. They are discussed below, under three headings: evaluation as historical practices; NGO transformations; impact governing.

Evaluation as historical practices

The literature section made the case that there is little appetite for seeing evaluation as historical practice, as cumulative activities, politics and relations, connecting some stakeholders, marginalizing others. The iceberg diagnostic suggests this is because historical practice is submerged, out of focus, illegible, in contrast to the valued claims of efficiency and effectiveness etc demanded in the sector and prescribed by technical methods. Four silences are described: evaluation as practice beyond individuals and methods; expertise drowns diverse voices; backstage politics and the impact of methods themselves; and reputational provenance of TIEK.

Silence 1: Beyond Individuals and methods

There were concerns in the literature about evaluation reform being blinkered to individual behavior, individual ethics, and method prescriptions (Bamberger et al, 2010: 2). Picciotto (2013: 18-19) argued that evaluation independence is under threat because its governance is seen as limited to individual ethics, skills, and organisational quality control. This study supports the need to look beyond individuals and methods. The NGO cases showed the diverse evaluation framing mechanisms: e.g. sector demands, funder compliance, impact marketing (Impact-2), risk management, the need for data, organisational strategies, and relationship building. At the expert NGO, methods were imported not from academia, but from commercial practices and digital vendors, spreadsheets elevated statistics and submerged doubts, tablet computers and sales pitches helped to make convincing stories. The novice NGO were learning that best practice methods could convert inauthentic villager responses into legible numbers, and that marketing could submerge humility. Both NGOs produced beneficiary profiles to highlight pre-planned successes and beneficiary gratitude.

Such bifurcating data practices were normalized at the expert NGO, and confusing for the novice NGO. At both, they were responses to development sector business challenges, rather than ethical or methodological evaluation quandaries. Evaluation methods focused on target sites do not elevate such bifurcations, nor do contractors who "*give primacy to the utilization of evidence ... for decision maker*s" (Picciotto, 2013: 19). Furthermore, at the novice NGO, because of the competitive climate, compliance requirements from multiple funders, and their ad hoc opportunism to learn evaluation, the NGO was transforming towards data/knowledge intensity, changing staff roles, learning data management, and exchanging evaluation data with other groups. These constitute diverse and unequal impacts of learning or adopting evaluation discourses, or merely mastering a method (Porter, 1996; Nichols, 1999; Mama, 2000; Picciotto 2012). Such impacts are beyond the doctrine that individual ethics and expertise will drive evaluation capacity and goods exclusively.

What are alternative perspectives that incorporate power and practice, and amplify such silent inequalities? Firstly, development and evaluation communities can acknowledge that evaluation is beyond individuals, methods and target sites, that it involves diverse practices and implies change and shifting power relations on who evaluates, who needs convincing in order to fund, and how evaluation methods elevate or submerge results. The risk is that questionable business tools and assumptions, fast digital data/knowledge, and marketing orientations trump individual ethics and method rigour as substantial drivers of evaluation.

Most small NGOs cannot overcome diffused power relations across networked development 2.0 landscapes. Evaluation lenses lack perspectives on multi-sited power/data/knowledge relations, eliding networks of professionals, evaluators, funders, and model prescribers such as researchers or INGOs who circulate critical, pragmatic and scientific subjugating rationalities (Avgerou, 2002: 77; Craig & Porter, 2006). Partnership talk or participatory intent is often used to frame local politics, needs or perspectives (Li, 2005; Lie, 2015, Bächtold, 2009). Work that questions the localized framing of participation (Cook & Kothari, 2001), the importance of judgement and divorcing incentives from measuring (Muller, 2018), and the digitization of methods and results via distant technologies or algorithms (e.g. Irani et al, 2010; O'Neil, 2016) suggest a wider view is emerging. Nevertheless, there remains a need, in aid chains and evaluation machines, to embed more sensitivity to impacts as multi-sited, multi-voiced, contested relations, even when this contradicts technical scripts and marketing pragmatics.

Silence 2: Expertise drowns out diverse voices

Technical evaluation literature elevated the expert framing of evaluations, often at the expense of diverse voices and ways of understanding change or impact. Deniston (1972; Mohr, 1995: 10) described how identifying a problem and a satisfactory intervention was an evaluation requirement. Mohr (1995: 13) added statistical methods and defining objectives as further expert framing devices. Edwards (1979) did suggest influential agents in historical roles also framed evaluations, however powerful agents, program aims, and expert voices need not monopolize evaluation frames. Alternative frames draw on local voices, for example how, monks wanted foreigners to silence their voices and how Solomon Islands officials wanted to hear about emotions, not numbers (Anderson et al, 20012). Such alternative frames, local, emotional or other, exemplify a wider array of voices, or ways of knowing impact. They challenge professional prescriptions.

The two NGO cases suggest broader views. In case one, the expert funder and NGO were skilled in the practices of building strong cases for their funding pitches. Theirs was not an academic, methodological or prescriptive expertise, but one of investor relations, business pragmatics, data management and marketing. Business tools and methods, tablet computers and spreadsheets of data on farmer livelihoods were used to design, capture and bifurcate data,

package and pitch exchangeable impact knowledge. Farmer voices were not curtailed by evaluation expertise alone, but by a cluster of disciplinary knowledges.

The novice NGO still had opportunities to involve more diverse voices, partly because they were as yet unskilled, skeptical about marketing, and lacked organisational, human and technological resources for expert prescriptions. They were trying to use evaluation data/knowledge to develop relationships, not yet expert enough to be instrumental for funding purposes, and not yet receiving excessive demands from funders for rigorous impact reports. As such, during their learning (and confusion) they made time to listen (Anderson et al, 2012) to doctors, public health officials, government, other NGOs, parents, volunteers, and migrant groups. As novices, they were seeking to use evaluation data/knowledge for funding bids, but also for advocacy and relationship building.

What can be learned from these observations? Firstly, impact demands lead to professional framing and bifurcation of data and voices. Diverse voices, critical appeals, and uncertainties remain silent in such environments. Secondly, NGOs who are not expert or experiencing heavy pressures to market impact, can explore diverse pathways for sharing data/knowledge. In both cases and in development 2.0 more broadly, NGOs are under pressure to produce representations for sharing, brokering, exchanging and pitching.

In summarizing expert discourses and multi-voiced views, there appears a compacting function, featuring demands for objective, solid impact results, using "schemata" which submerges critical reflection, diverse voices, dialogue, contexts, power, and local know-how or "metis" in Scott's vernacular (1998). Less pressure and less expert framing expands power/data/knowledge sharing options, but more pressure and expert control solidifies power/data/knowledge configurations and inequalities regarding who may trade impacts and how. Such drives in development 2.0 are likely to persist as technical discourse accelerates in a dance with skeptical demands, more targets, more risks and more digital representations. Critiques of representation and power may support shared dialogue, but not guarantee it[33].

---

[33] For an in-depth discussion of critiques and dialogue, which contrast insights from Foucault and Habermas, see Khemmis (1993) on Evaluation or Stahl (2004) on Information Systems.

Silence 3: Backstage politics, method effects

This critical question from the literature concerned if, and how, backstage political influences (Edwards, 1979) and the historical impact of methods themselves (Porter, 1996) were acknowledged in contemporary NGO evaluation practices. These concerns are related as methods develop over time, and involve a politics of representation around numbers, voices, claims of objectivity or inclusion. Were such politics visible in the cases?

At the expert NGO, backstage politics was not evident. Pragmatic business and management approaches privileged fast access to data on what was happening at target sites. Rather than a method growing in influence, it was business, marketing and technology discourses and tools that had grown to configure their evaluation machine. Visits to the philanthropy in London and the NGO in India were planned for mid 2014, but cancelled when the collaboration ended in April 2014. Thus, a deeper examination of the backstage politics and effects of the evaluation methods was not possible.

At the novice NGO backstage politics remained silent and no method was clearly established. The CEO wanted the NGO to learn impact evaluation to support their financial sustainability, but staff had doubts about marketing impacts. The NGO were learning to strategize evaluation data/knowledge, broker with partners, satisfy diverse funders, and listen to target community voices. No explicit politics nor specific method effects were significant, however the wider development 2.0 data/knowledge shift was incrementally transforming them, as a foundation for future status, legitimacy, communications, pitching, brokering and relationship building.

What do we learn from the cases? Firstly, it is not only backstage politics or methods that shapes evaluation, NGOs or power dynamics, but also the impact of changing discourses, towards business, marketing, technology and data/knowledge intensity. Backstage politics and method effects were not significant in either case. This might be a limitation of the study, or a displacement of both by broader development 2.0 shifts to digital technologies and markets (Brigham & Hayes, 2013; Hayes et al, 2017; Kelly, 2018;).

Silence 4: The provenance of technical discourse

The literature highlighted problems concerning where evaluation methods were sourced, their provenance or genealogy. The UN, World Bank, INGOs, universities and philanthropies were cited as "pumping out" evaluation prescriptions (e.g. Cornwall, 2005; Wallace, 2015; Picciotto, 2012; Duflo & Kremer, 2005; Center for Global Development, 2006; Garbarino & Holland, 2009). Prescriptions from reputable organisations become normalized, contributing to professionalized networks, impact governing, and expert problem/solution frames around marginality or poverty (Procacci, 1991). Was this concern evident in the cases?

At the expert NGO, the funder and NGO used financial sector auditing techniques and digital technologies to produce convincing impact stories, a basic before/after approach, and knowledge from marketing and data management. They used no evaluation methods which could be traced to particular sources, however their digital tools included Microsoft Excel, web databases, Android tablets, and a custom adapted marketing app. The provenance of evaluation practices here reach not back into technical evaluation, but into finance (Strathern, 2000), management (Dar and Cook, 2008) and technologies (Zuboff, 1984; Irani et al, 2009; Heeks, 2010). Provenance was more dispersed than identifying a single institution, a Harvard, a Red Cross, or a celebrity (Ponte & Richey, 2014).

The novice NGO has no impact evaluation machine, tools, rules, resources or objectives. Evaluation know-how was brought in by US Ivy League interns and the author cum consultant, from a UK university. The NGO was skilled at engaging with communities, facilitating events, producing reports and managing projects. But impact evaluation was a new set of practices for them. Through working with diverse ethnic groups and marginalized migrants, local NGOs and government levels, the NGO had opportunities to hear, adopt or adapt ideas with a local provenance. However, such locally sourced traditions or ideas were not guaranteed to be mobile or have value in international funding markets. Leal (2007) lamented these fashions, and Kothari (2005a; 2005b) highlighted the post-World War Two shift to technical development discourses in the west. Local provenance would not bridge such an evaluation capacity gap, nor help the small NGO "*catch up*" Sumner (2006) with professional evaluation. This small NGO was humble, didn't "do media", and couldn't do development 2.0.

What do these lessons teach us about evaluation practices and provenance? Firstly, the literature on provenance is credible, but not extensive enough. Adopting methods from Harvard can add legitimacy to evaluation practices, but in the cases normalized sector technologies, business and auditing assumptions, and data/knowledge intensity added more exchange value.

Development 2.0 data/knowledge intensity might be a current trend, as are Big Data for Development, Knowledge Management for Development and Adaptive Management. Such discourses carry value for NGOs adopting them as they strengthen impact messages in global markets. As such, the capacity gap will continue following different innovations in business, marketing, finance, data/knowledge management, and technology.

The provenance of evaluation prescriptions is wider than single institutes, but critical problems silent in technical debates. The iceberg helps us see how evaluation is changing. Seeing evaluation-as-practice, in order to incorporate provenance, fashions, tools and assumptions etc can help donors and evaluators see how evaluation is changing, according to development 2.0, marketing processes, and audit cultures.

NGO Transformations

The literature review showed how NGOs are being transformed by development 2.0 data/knowledge intensity, and how impact evaluation was an important part of such changes. The cases confirm this observation. NGOs provide compliance data/knowledge for aid chains, communicate and market for funding and reputation, construct impact commodities and dedicate office resources for this work. However, the cases go further and show how both expert and novice NGO evaluators are adopting new ways of working and seeking new strategic responses to development 2.0 challenges, with data/knowledge capacity a key issue.

Silence 5: NGO services providers and compliance capacity

The literature review described how since the 1990s NGOs have become services providers, working to short timeframes, focusing on results, adopting ICTs for managing data and reporting (Walsham, 2001; Wallace et al, 2006), and aligning goals, outcomes and impacts to donor compliance frameworks (Bebbington, 1997; Lewis, 2007); Banks et al, 2015; Wallace et al, 2006). The early 1970s sweetheart appeal of NGOs had been their agency i.e. intimacy with communities, political savvy, mediating roles, and ability to harness grass roots energies. However, these conflicted with 1990s moves towards service and compliance (Drabek, 1987).

In the expert NGO case, the partners assembled an evaluation machine over years, which produced audible products and convincing stories for investors, bifurcating out farmer voices, clamour, and participation. The NGO in case 2 was struggling to assemble their own "*full-fledged*" machine, not yet having expert cell data or clamour. Their organisational culture focused on community engagement and project management, not data/knowledge supply chains. Service models and compliance restrict diverse engagements and the ability to hear something new or different, to legitimately respond to a housefire incident outside project scope for example. Solid impact products showing target success and effectiveness for compliance purposes, submerging messy non-linear innovation, relationship building, experimentation and NGO agency.

In future, small NGOs will transform and learn how to produce data/knowledge products, encounter new opportunities for advocacy, partnerships, and relationship building etc. However, compliance in development 2.0 involves more and more ICTs, digital data and knowledge intensive processes (e.g. big data, the latest impact methods). This creates an innovation capacity gap, with power/data/knowledge inequalities. Small NGO service providers must learn systems, take training, procure technologies, master new forms of expertise and models etc in order to catch up and compete with large INGOs, wealthy philanthropies and international consultancies. From Big Data to digital communications, these new fashions leave small, local NGOs in an enduring (permanent?) development 2.0 capacity catch up mode (Sumner, 2006; Heeks, 2008; Heeks, 2010). In the study, one NGO had aligned to a philanthropy itself building new technologies and platforms, but the other was struggling to compete in data/knowledge intensive development 2.0 markets. This gap is an iceberg silence, and contemporary impact evaluation helps generate the silence.

Silence 6: Communications and marketing capacity gap

The literature revealed questions about how small NGOs were responding to impact data/knowledge use in communications, marketing and branding. A consequence of the need to demonstrate results, impacts, effectiveness and other audibles above the waterline is that NGOs must shift resources and strategies to focus squarely upon communications, marketing, public relations, and branding to support fundraising, reputation, and niche expertise marketing. Key issues regarding this transformation include: how the millions of small NGOs

(Lewis & Kanji, 2009) lack resources to do sophisticated communications; how they are under pressure to exaggerate results for funding, marketing or policy influence (Eyben et al, 2015; Roche, 1999: 2); and, how new actors and tactics (e.g. celebrity branding, cause marketing) (Kothari, 2014: 57), powerful imagery and campaigns become part of normalized communications (Ponte & Richey, 2014). Use of statistics, infographics, narratives, campaigns and other techniques have become tactics in effective aid communications. They are also aimed at donors, audiences and consumers in predominantly northern populations, not local actors and stakeholders engaged with the sites or communities signified in the marketing activities. Given these issues, what can we learn from the cases?

The expert NGO had a public website with information about projects and impacts. They used marketing language in sales pitches for investors at funding fares. However, there was no evidence that they engaged in cause marketing, campaigning, or other sophisticated tactics. Their marketing orientation was evident in the model village narrative, statistical before/after comparisons, and the overall need to present a convincing story to potential investors. Diverse forms of expertise supported marketing in this respect e.g. their development and data management. In the second case, there was no evidence of campaigns or strategies. Their website had broken links and old news. They had previously published articles in local newspapers about work with marginalized communities, but this was not a regular tactic. They did not hire communications staff, and although there was awareness of the need for catchy or "sexy" project titles, this contrasted with normal project long, technical titles. Staff doubted their ability to claim impacts and the value of marketing, as opposed to "humble" work.

Due to existing funding fair networks, one NGO did not need to dedicate significant resources to marketing. The other did need to market, but lacked capacity. Through the latter's transformation to data/knowledge intensity, they could over time learn to bifurcate data, package and market more. The two cases show firstly how aid markets are diverse and how funding streams configure fundraising tactics for NGOs, and secondly how impact evaluation in practice contributes to a capacity gap that encompasses not just compliance data/knowledge, but enduring communications gaps too.

Silence 7: Confusions, capacities and commodities

Maturing the capacity thematic further, what are the silent results of NGO evaluation confusion and capacity demands? Issues identified in the literature included: confusion in NGOs at the need for know-how, expertise and processes of performing impact evaluations; the hopes and fears of staff around the "*dense fog*" of evaluation (O'Sullivan, 2004: 76-78); the "*furrowed brows*", "*plethora of issues*" and need for organisational consensus (Norrish and Sayce, 2006: 4-5); the confusion of expert advisors (Stern et al, 2012: 5); the need to master theory, prescriptions, meet goals, manage data and produce scientific rigor; the staffing gap between large organisations who produce impact prescriptions and small NGO consumers of such prescriptions (Mosse, 2004a; Sumner, 2006; Brigham & Hayes, 2013); the excessive flood of TIEK produced in a market logic (Kothari, 2005b: 440; Schuurman, 2009: 834); the need not for local staff, but for staff who "*know how to measure efficiency and increase the impact of projects*" (Schuurman, 2009: 838); and finally, the need for any NGO that wishes to secure funding, gain status or get a seat at policy tables to speak in technical language or risk being marginalized financially and politically (Wallace et al, 2006). Given such a bag of confusions and capacities, what more can be learnt from learn from cases?

The expert NGO, guided by their philanthropic funder, did not display confusion or a lack of capacity to design, manage or construct impact data/knowledge. They were not expert technical or participatory evaluators, but were expert in their instrumental funding motivation, data design, capture, storage, bifurcation, and knowledge packaging and pitching. They were decisive about not including qualitative interviews, aware this could disturb success narratives. Their development approach was to bring technical agricultural expertise to participants, not to advocate political change or collaborative design with farmers. The case is informative because the partners had a business model and connections to private investors and funding fairs, potentially avoiding public skepticism about aid (Picciotto, 2012: 214; Roche, 1999: 2).

The novice NGO experienced many of the confusions, as part of their learning and transformation. They had diverse funders and partners to work with, but very limited tools, technologies, methods or skilled staff. One could say the dense fog of evaluation was very much in place and there were plenty of furrowed brows, not only about mastering a new evaluation skillset, but also about the tools, resources (e.g. English-speaking staff), priorities, and attitudes to doing evaluation. Their CEO saw the benefits of evaluation for reports, grant bids, advocacy and partnership development, but staff less sure. Arguably, this case resonates more with the millions of small NGO around the world than the expert case.

Such lessons tell us about the confusion, capacity lack and need to produce impact commodities, about how convincing impact messages work in some funding relations without the need for the floods of methods and prescriptions produced by funders, universities, the World Bank or UN. It also reminds us that many of the millions of small NGOs, with little capacity and much confusion, wish to master evaluation to strengthen their positions. Impact discourse is itself an aid sector commodity that generates gaps between producers and consumers of methods, prescriptions, standards and results. This gap is real, but not uniform. Without the right networks, small NGOs will struggle. They must be opportunistic in trying to "*catch up*" with constantly moving development fashions and technical discourses (Leal, 2008; Kothari, 2005b), and the wider development 2.0 landscape around them. This will affect their funding streams, legitimacy, reputation, status, branding and compliance capacities too. The confusions however, are by-products, waste, illegitimate impacts of expert knowledges (Scott, 1998; Escobar, 1995), uncounted (Miller, 2004: 181), peripheral rather than focal knowledge (Polanyi, 1967; Tuomi, 1999). In this sense, more silent confusion indicates more market opportunities for diverse development 2.0 data/knowledge producers – researchers, training organisations, regulators, academics, and other expert groups.

Silence 8: Data/knowledge work displaces fieldwork

The literature pointed to how NGO work involved the incremental displacement of field work by office work for data/knowledge representations, digital reports, or media for aid chain partners and local organisations. Problems here related to the demands of bureaucratic work (Quarles van Ufford, 1988; Mosse, 2004a; Wallace et al, 2006) versus listening to field staff and beneficiaries (Anderson et al, 2012; Norrish & Sayce, 2004: Wallace et al, 2006). Not just auditing work (Strathern, 2000), but evaluation work (Townley, 1995; Kirkpatrick & Hulme, 2001; Duncombe, 2009) and development 2.0 (Thompson, 2008; Kelly, 2018) risk shutting out tangible, engaged work with marginalized groups. Office data/knowledge work can directly damage field work too, if auditing destroys morale (Seddon, 2005) or breeds distrust (Demos, 2005: 7), or when performance systems amplify instrumental over communicative rationalities (Townley et al, 2003). Were such displacements or dangers evident in the cases?

At the expert NGO and private philanthropy, managers were concerned with quantitatively informed sales pitches. There was little time for qualitative data, local bosses did not want to

waste time on it, the partners didn't know how to deploy such methods, captured data had been useless in the past, and local farmers were seen as inarticulate. Authentic qualitative voices and narratives could be inefficient or damage impact sales pitches. Field work was seen to involve clamour, whereas office work engaged investors, was fast, efficient, and supported new organisational strategies like the MIS initiative.

At the novice NGO, transformation was in progress, involving more interns, more ad hoc trainings, and trying to bring teams back to the office despite the difficulties of doing so in terms of time, staff locations, and busy field work pressures. The tension of office data/knowledge work displacing engaged field work was not normalized, but resisted. Media work was not what staff were comfortable with, and evaluation work distracted from the "real work". The CEO however, was acutely aware of the precarious funding environment, the needs for more evaluation, statistics, marketing, and new office-based skillsets.

In summary, displacement problems are silent because they are either dismissed, elided, or not legitimately dealt with in technical evaluation or data/knowledge management literatures. Furrowed brows, evaluation fog, organisational struggle, the weak foundations of data as objective facts, and the potential benefits of not adopting data/knowledge intensive impact processes are largely silent problems, seen as old ways or ignorant views. Proper data management, expert evaluation systems, measuring and calculating and so forth are viewed as professional, modern processes to be adopted. The critical perspective and learnings from the cases show not that data/knowledge office work should or should not happen, but its displacements, damages, confusions and the inequalities it generates are silenced.

Impact Governing

The literature identified a third group of questions, focusing on impact governing and anti-politics. The thesis focus on power/data/knowledge, the literature, and the case lessons all make clear that these particular, deep silences are foundational for the impact iceberg. Impact governing is a critically important dynamic in the recent history of the sector, current development 2.0 landscape, TIEK, and DIKW omissions (Tuomi, 1999; Lambe, 2011; Fricke, 2009; Kelly, 2018), and the role of NGOs as vehicles for performing audits, managerial controls and calculative practices (Lemke, 2002; Introna, 2003; Townley et al, 2003; Ilcan & Philips, 2010). These foundational silences are best articulated through governmentality and

developmentality perspectives (Foucault, 1979/2000; Ferguson, 1990; Escobar, 1995/2011; Li, 2007; Lie, 2015; Bächtold, 2015; Ilcan & Philips, 2008), because they relate power, data and knowledge closely together. Three elements are discussed below: Impact governing, data/knowledge mobility and anti-politics.

Silence 9: Governing impact via NGOs

The literature looked at how governing processes transform NGOs, how they come to "*embody the values and orientations of the market*" (Dean, 2010: 201). In this view, NGOs come to align with funders, managerial controls, and regimes of practice (Dean, 2010: 27) which prioritize technical and expert views of evaluation and impact. Such alignment supports success in funding markets (Banks & Hulme, 2015) and enables NGOs to protect themselves from outside interference (Roberts, Jones III & Fröhling, 2005). The NGO benefits by freely choosing to comply with funder and expert prescriptions. In evaluation, certain claims, scientific facts and professional criteria become legitimated (Introna & Whittaker, 2004), and others such as trust, emotion, empathy, intuition or political claims, become delegitimized (Knights, 2004). With developmentality, the 1990s/2000s shift towards relationship building, partnering, participation, localization and knowledge sharing represents a new, smoother governing (Li, 2007; Lie, 2015). Were these issues evident in the cases?

At the expert NGO, spaces for diverse voices, participation, and political dialogue were strategically submerged by the partners. Orientations towards commodified impact products, digital innovations (e.g. the MIS) and funding markets were prioritized and normalized. Pitching activities backwashed through the supply chain. Opportunities for building trust and engagement were submerged, whereas opportunities for market pitches using survey data, customized technologies, and fast transmission to decision-makers were elevated. The market orientations and expert knowledge flows resonate with impact governing and developmentality knowledge networks, but not developmentality, partnerships or participation.

The novice NGO were under pressure to learn evaluation, largely to boost their competitive edge in funding markets, but also to build relationships with partners, for example through healthcare data. However, they had not yet learned how to legitimately claim impact facts, adopt expert methods, or shift to a business mentality. They remained engaged with target communities, focused on a mix of project management and community participation. They

remained suspect about impact marketing and increasing office work, but open to brokering data/knowledge. In this case, a governmental market rationality was partly in place, and the partnering and participation was not, yet, instrumental, remaining humble and engaged.

The silent implications of these cases for impact governing are twofold. Firstly, governing processes and market orientations were observable at both NGOs, albeit more mature at the expert NGO. The novice NGO was exploring diverse sharing and exchanges with local partners and for advocacy purposes. This suggests governing and participatory processes can be in parallel, for markets and for sharing as evaluation mastery grows. The cases illustrate how agents are free to choose, share or sell, and how participation and governing go hand in hand (Li, 2007; Dean, 2010), in contrast to more polemic perspectives where expertise, evaluation or development 2.0 is only good or bad.

Secondly, the capacity gap itself is a driving commodity in development 2.0 and expert knowledge networks, like evaluation. This has been termed the "*knowledge economy of capacity building*" (Phillips & Ilcan, 2004). Aid markets claim to change beneficiaries lifeworlds, however additional imminent aims involve buying, selling, constructing, consuming, learning and innovating, around impact evaluation, development 2.0 or similar fashions (Leal, 2007). Each impact audible and silence is implicit in these diverse markets, and each small NGO is susceptible to numerous capacity sinks. As free agents, each must find ways to not only consume, but also share or sell their own data/knowledge innovations, like the MIS in case one and the healthcare data in case two.

Silence 10: Impact knowledge mobility

The literature review highlighted problems around the networks and circulation of impact evaluation guides and prescriptions, management models and "*laptop consultants*" armed with participatory rhetoric (Shivji, 2006: 23); how such knowledge systems were divorced from practice (Mosse, 2004a: 85); how knowledge becomes normalized away from its local site of production; and how unexamined plans and blueprints and vague mandates serve expert producers, not sites of adoption (Sandercock, 1998: 88; Wallace et al, 2006: 36; Li, 2007). Development expertise, in books, heads, methods and processes travel from one micro setting, along global routes, to new micro-settings of application, potentially hindering participation and engagement at new sites. Such expertise defends itself and its networks (Kothari, 2005b:

439; Chambers, 1993: 5; Phillips & Ilcan, 2008). Once commoditized and mobilized even qualitative or participatory models of evaluation can become check box commodities which submerge local, critical or alternative views. Were such dangers evident in the cases?

At the expert NGO, certain knowledge qualities were prioritized: smooth, frictionless, efficient, speedy data/knowledge to be exchanged in expert networks or at aid funding fairs**.** These incorporated quantitative data and digital technologies, and contrasted with friction-full knowledge, such as local clamour, inarticulate farmers, fictional, slow or lost data, and voices in local languages. Global rationalities of business, data management and marketing were also evident, visible in the use of "*poverty verticals*" to dissect local poverty into market-oriented opportunities, part of the MIS development. A new evaluation MIS in the Indian development sector could crystallize and amplify such rationalities on a larger scale for the partners to benefit from, despite the assumptions, omissions, and silences such systems risk submerging. However, dilemmas of participation, voice, doubt and power were not concerns expressed by the philanthropy or NGO management.

The novice NGO wanted evaluations and advice from visiting professionals, the author as a researcher/consultant, and interns from prominent US universities. They grasped opportunities to send staff to training programs in Switzerland and in Bangkok to bring prestigious knowledge back to base. Staff said they were humble and engaged, but although these qualities had local legitimacy, they did not have market value globally. Listening to beneficiaries in poverty contexts, using dialogue or more traditional ways of evaluating had limited market mobility, hence the steady transformation of the NGO towards a data/knowledge intensity visible in global markets (from Impact-1 to Impact-2, NGO 1 to NGO version 2 in the analysis). There did remain spaces for participation and sharing with local service partners and public agencies, as well as using evaluation data/knowledge in advocacy work or relationship building. As such, local mobilities may not be completely under the prescriptions of global rationalities, which could lead to local innovations and opportunities,

In terms of taking the debate forward on the problem of silent global knowledge mobilities, one important result is worth highlighting. Impact data/knowledge cannot be explained solely by impact evaluation prescriptions, but it infused with business, technology, data management, software development, marketing, organisational strategy, knowledge brokering and innovation too. These accompany rationalities pointed out by Avgerou (2002: 77) and others.

Small NGOs cannot meet all these niche forms of expertise, but may focus on one or two, as in the cases.

Silence 11: Impact anti-politics

A final silence is anti-politics. How do NGOs performing impact evaluation navigate anti-politics? The literature threw up six relevant issues: how development goals expand (Escobar, 1995/2012); how failure prolongs technical interventions (Ferguson, 1990); how politics and power are obscured or black-boxed (Harris, 2002; Bächtold, 2015); how anti-politics works in specific projects and at sector scale (Scott, 1998; Gabay, 2012); how complex accounts of poverty rarely feature in technical evaluation (Banks & Hulme, 2012; Mohan, 2002; Li, 2007) and the need to acknowledge calls to re-ignite political voices using more "real" or honest" evaluation that is more sensitive to complex practices (Duval et al, 2015: 49; Sumner, 2006; van den Berg, 2004: 68; Law, 2004). In short, how can evaluation processes overcome their own power effects? What was learned from the cases?

The expert NGO massaged data cells, adapted farmer profiles to illustrate success, packaged and pitched knowledge to make exchangeable, marketable products. The partners' evaluation machine effectively silenced voices, contestations, dialogue and political representation. TIEK pressures to elide politics in evaluation were evident, but manifested through assumptions (e.g. use of discrete questions), technologies (e.g. professional spreadsheets), the need for fast data (not complex contexts) and marketable results (not honest accounts). This led to a normalized submerging of political participation, unquestioned by management.

The novice NGO engaged in advocacy with local and national actors, explored collaborative designs with migrant workers, and tried to broker data/knowledge sharing with the public sector and other NGOs. However, they also edited out the complexities of local poverty and marginalized participant voices (e.g. the housefire, new mothers returning to work, stakeholder profiles). As they learned evaluation, they were under pressure to squeeze local contexts into statistical analysis frameworks. They were troubled by marketing, but learning how to market impact. As their mastery of evaluation grows, and the organization transforms towards data/knowledge intensity, they will likely be faced with a dilemma: to perform honest evaluation, hear multiple voices and politics, and share lessons with partners, or perform impact marketing to generate funding and status opportunities.

The takeaway is the pressure to distort and bifurcate evaluation data, remove contestations, and package it for exchange in markets or with partners. This is evident in both cases, but far more advanced at the expert NGO. This impact iceberg silence illustrates the extent of submerged anti-politics, smoothed over by technical expertise, inauthentic participation (Li, 2007) and networked data/knowledge exchange requirements (Ilcan & Phillips, 2008). Any honest evaluation requires the re-insertion of politics and practice to open this black-box.

7.2.4 From critique to engagement

Many issues related to the above waterline audibles and below waterline silences have been discussed. The iceberg diagnostic also has a key argument underpinning it, which is the positioning of impact governing, unequal knowledge mobilities and anti-politics as foundational aspects of developmentality (Li, 2007; Lie, 2015, Ilcan & Phillips, 2010) structuring the iceberg historically. Such a dispersed framing of evaluation practices, evident in ongoing transformations of evaluation itself, NGOs and impact governing is in line with the diverse demands and configurations identified in the literature. Significantly, these dynamics continue remain inaudible, outside TIEK literature and the DIKW legacy.

However, if political participation and dialogue are to be re-inserted into impact evaluation practices, after decades of submergence by business pragmatics, a narrow technical discourse and DIKW omissions, then multi-voicedness, and political representation need to be re-articulated as audible and mobile elements of evaluation practice. Diagnosing silences using the iceberg is a first critical step, but responding to them requires more, ways of articulating, amplifying and mobilizing the silences within development 2.0 networks. Responses must travel beyond critics, into business pragmatics, data/knowledge models, and technical evaluation discourse itself. Evaluators, donors and NGOs have a role to play, but require accessible concepts, models, and tools to articulate and engage with the iceberg silences.

7.3   Contribution 2: Audit 2.0 devices

7.3.1 Introduction to the conceptual devices

This section describes four concepts or devices that amplify the silent impacts of evaluations, and seek to take scholarly critiques out into the wilds of practical engagement. These devices are not scientific methods for marketable results, but are devices to make power and practice more audible. They respond to the problems of auditing cultures embedded in current evaluation techniques, which are implicated in performance monitoring (Townley et al, 2003), new development management (Dar & Cook, 2008) and governing functions (Hayes et al, 2017). Such functions can be characterized as audit 1.0 to highlight the demand for results-oriented knowledge products, which deliver exchange value in funding markets (Impact-2 in the cases).

In contrast, the concepts that follow can be considered as audit 2.0 devices. They relate to Impact-1 interactions in the cases, silences, power relations, politics, and processes of contestation or transformation. Audit 2.0 emphasizes the silent power/data/knowledge relations generated during evaluation data/knowledge construction. It foregrounds multi-voiced accounts and negotiations of what happens, rather than business pragmatic, objectivist and short-term accounts of what works. Audit 2.0 highlights the impact of evaluations, rather than the evaluation of impact. It draws upon more engaged or active anthropologies (Lewis, 2005; Marcus, 1995), action research approaches, mode 2 knowledge production (Gibbons et al, 1994), cultural probes (Gaver et al, 1999; Crabtree et al, 2003), and field stimulations (Salancik, 1975). These concern processes and engagement over extracting results for products.

Audit 2.0 rejects purified accounts of impacts or sanitized success stories which are designed to boost exchange value in development networks. It re-asserts engagement, equality of representation and voice, and most challenging of all, solidarity with marginalized stakeholders or concerns. Audit 2.0 questions the DIKW knowledge production hierarchy, and the claim that data are the empirical and objective foundation for expert knowledge (Earl, 1994; Tuomi, 1999; Fricke, 2009; Lambe, 2011). Audit 2.0 questions the pragmatic business rhetoric of DIKW and the information, knowledge management and evaluation systems which draw upon its implicit legacy (Rowley, 2007; Lambe, 2011).

Four audit 2.0 conceptual devices are described below: firstly, data/knowledge supply chains and networks of activities; secondly, the 6P sensitivities; thirdly, the impact spectrum, and fourthly, datamentality and datamateriality. In terms of application, audit 2.0 devices

accompany the impact iceberg diagnostic. They are intended for use by different groups involved with impact and evaluation, who are concerned with evaluation practices and unequal power relations. These include donors, evaluators, NGOs, and researchers in evaluation, ICT4D, KM4D, mainstream development or critical development. Future work on the devices will lead to tools such as visuals, checklists, key questions, or agendas to facilitate discussion in evaluation arenas. The devices inform areas such as program design, negotiations between commissioners and evaluators, community responses to evaluation claims, training for professional evaluators, and agenda items for researchers.

7.3.2 Data/knowledge dyads, chains and networks

This section outlines the first audit 2.0 device, which is data/knowledge chains and networks. These networks offer an ontological view of development data/knowledge construction, mediation and consumption, in which diffused practice and power are constituent parts. Data/knowledge chains and networks address silent issues of power/data/knowledge which although widespread in contemporary development practice, have not been adequately addressed in three areas of relevant literature. The first area is TIEK, the second is the DIKW legacy in Knowledge Management and Information Systems (Lambe, 2011; Rowley, 2007; Fricke, 2009), and the third is development 2.0 (Thompson, 2008; Heeks, 2010; Kelly, 2018), a relatively new area at the boundaries of Development Studies, ICT4D and Information Systems. The device builds on the iceberg silences and case analyses of elevating/submerging temporal activity chains.

Five steps below elucidate the data/knowledge network device. Firstly, a recap of the TIEK / DIKW problem is outlined and secondly a basic evaluation dyad identified. Thirdly, the dyad is opened out to include larger aid chains (Wallace et al, 2006). Step four involves zooming in to see specific temporal chains of micro level activities. Step five zooms out to encompass the wider network of agents and exchanges.

Step 1: The TIEK / DIKW problem and rationale

In the literature, a linear and managerial evaluation meta-model was identified where evaluations capture inputs, analyse them, and produce outputs. Output results inform decision-making activities about target groups, made largely by funders and managers. In this model, evaluation is considered politically neutral, and a scientific or participatory endeavour.

Methods and results are issues of expert consideration and managerial strategy. Participatory methods generate more engaged results with ethical concerns or packaging, but focus their gaze on target sites of change, not aid chains or diffused power/data/knowledge dynamics (Green, 2009; Hayes & Westrup, 2012). This rationale constitutes a meta-model, aligning DIKW and audit 1.0 – the framing and measuring of impact from an expert, external position to generate documentable results and audible impact narratives for exchange away from target sites. Figure 7.2 shows how objective data is seen to deliver rational decisions, in DIKW and TIEK.
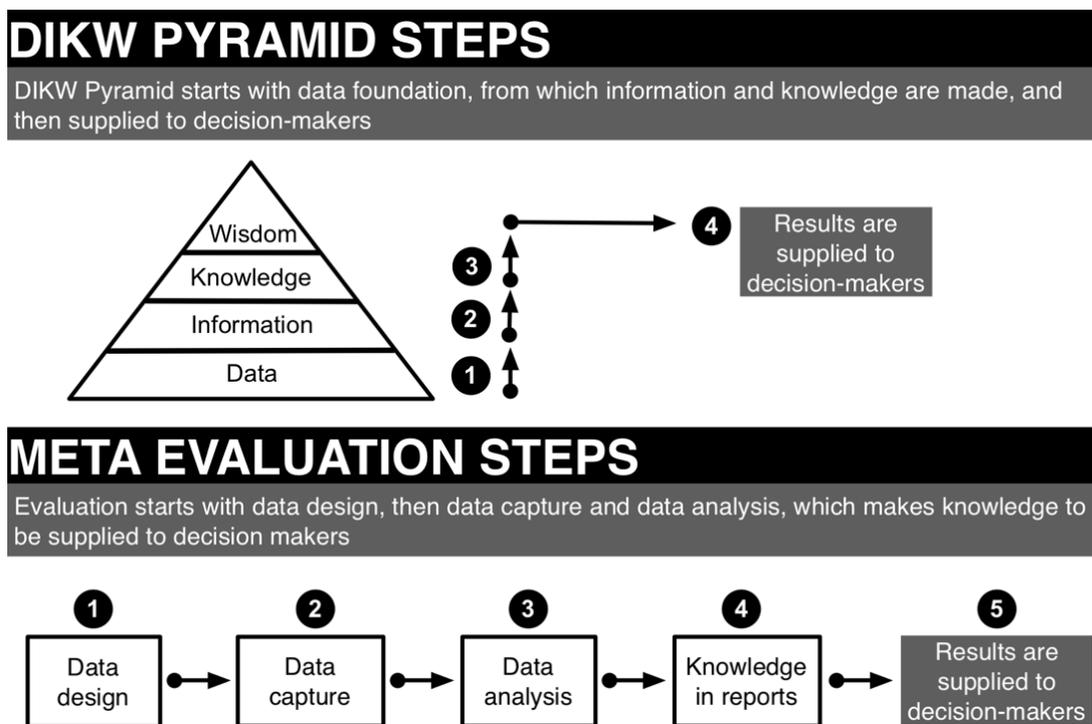
Figure 7.2: DIKW and TIEK prescribe data to make results for decision-making

Step 2: Basic Evaluation Dyad

The step 1 meta-model requires, step 2, a simple dyad with two agents (Figure 7.3). The intervention agent, such as a small service contracted NGO in country, provides inputs such as activities, finance and local know-how to produce a social or economic change in target

community sites or behaviours. Evaluations assess these inputs, changes and outputs as data which are captured, analysed, bifurcated, packaged and pitched to show changes according to intervention target goals. This is the basic dyad of two agents constructing legitimate impact data/knowledge. However, this basic dyad model is insufficient to account for broader impacts, changes, and power relations. The gaze of the evaluation process is here narrow and limited, upon target sites and direct intervention agents. In practice, evaluations are constructed in wider relations.
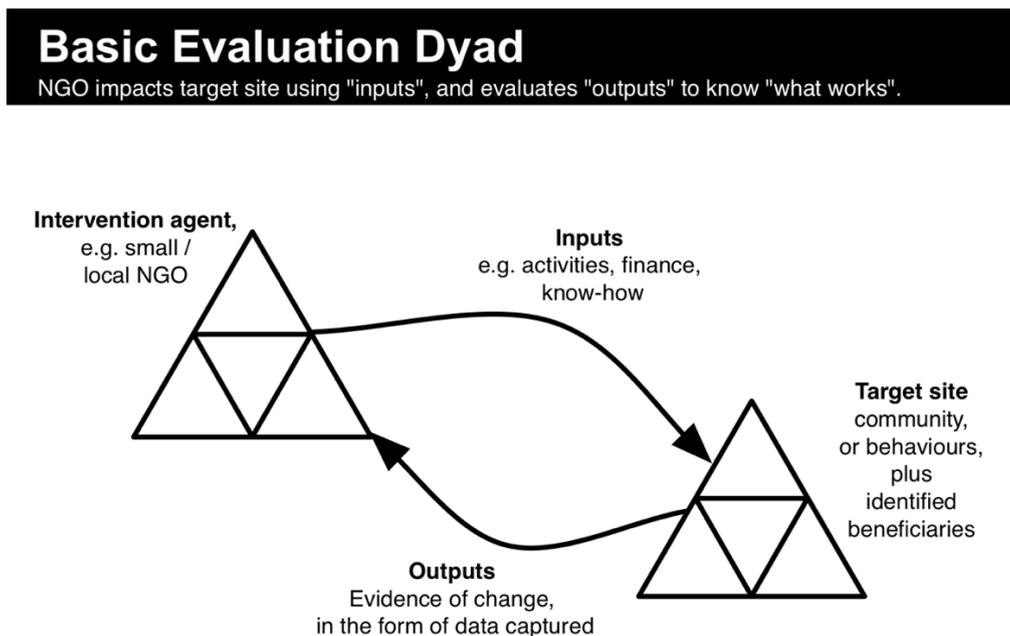
Figure 7.3: Basic evaluation dyad of two agents constructing impact data/knowledge

Step 3: The Aid Chain

In the aid chain (Wallace et al, 2006: 13 and 166), international funders, such as governments, INGOs or philanthropies, contract and finance NGOs in host country settings to provide services. The local NGO becomes a service provider and the international organization establishes a compliance framework which shape legitimate impacts (Figure 7.4).

Additionally, third party evaluator consultants are often contracted to provide evaluation expertise, independence, neutrality, ethical standards and professional legitimacy. This changes the nature of what evaluations performances and practices. Both the INGO and evaluator seek to secure their sustainable presence in aid chain markets, and evaluations

become strategic business opportunities in these ventures (e.g. Picciotto 2012; Brigham & Hayes, 2013). At this point, market governing incentives delegitimise less articulate, less documented small NGO and community ways of thinking about change or impacts.

Aid Chain Evaluation

Larger aid chain acknowledges influence of others e.g. INGO, funder and evaluator. Funders or INGOs mediate evaluations through compliance prescriptions. Evaluators mediate outputs, bringing expert knowledge and independence. Both seek market influence and sustainability.
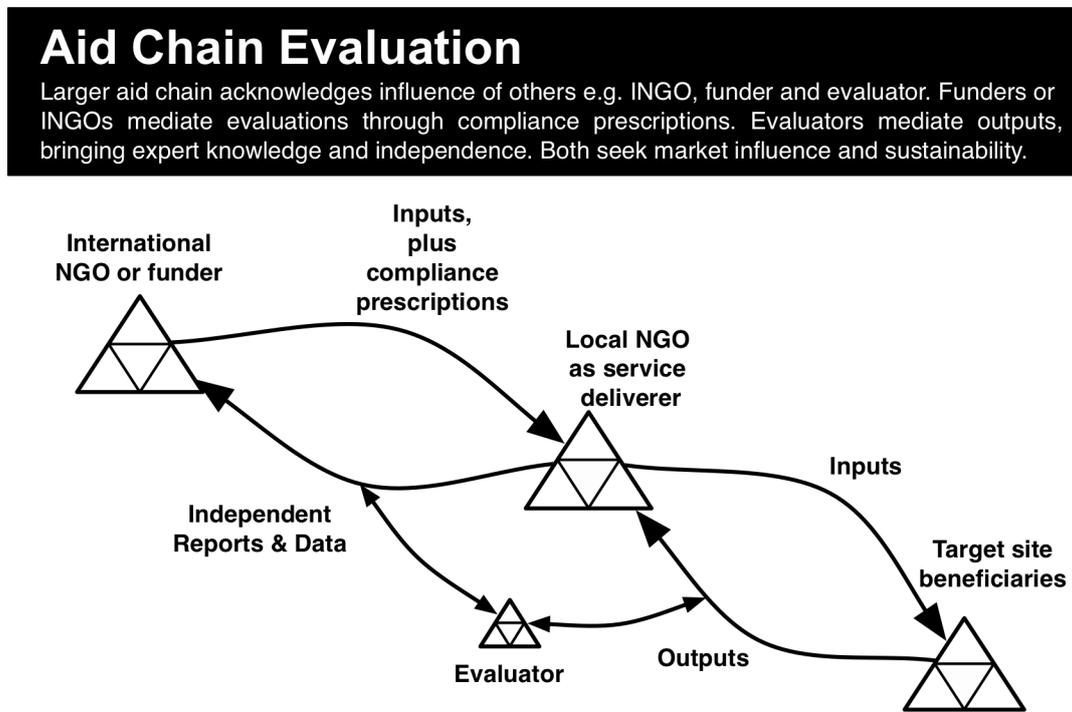
Figure 7.4: The aid chain involves further actors and activities

Step 4: Zooming in to temporal activity chains

Wallace et al's (2006) aid chain relationships are a core part of a more comprehensive impact evaluation network. However, before looking at this network, it is important to zoom in and clarify the power/data/knowledge relations in this core area.

The diverse activities in the case NGOs illustrated the submerging of target community participation and voice, and the elevation of expert knowledges, the transformation of a small NGO as it learned to practice impact evaluation, its struggles and its new opportunities. Both cases depicted how in doing evaluation, approaches, agents and governing relations were moving towards more expert, managerial, digital and marketing centric ways of evaluating. In both cases, mastering the changing transformations was perceived to bring more influence and opportunities. Case 1 illustrated this process from the perspective of partners who were skilled

and knowledgeable about evaluation. Case 2 showed a novice NGO, but one which also sought to share data and knowledge with various partners.

Early cycle activities concern assembling the evaluation machine and include: identifying the needs-to-know impact, authority lines, methods and techniques, data digitalization processes, and tasks or roles for people. The mid cycle concerns data practices starting with people's roles and tasks, and including data capture, bifurcation, storage, access, analysis, and sharing activities. The late cycle concerns knowledge mobilization and includes data sharing too, as well as knowledge production and packaging, pitching to markets, informing organisational strategies and brokering with other partners. The temporal chain, building from the cases, does not prescribe how to sell to markets or do evaluation "right", but describes more complex social and digital activities involved in evaluation practice. Other cases could identify further activities, skip specific activities, alter the sequence, or spin off in new trajectories, such as the case two use of evaluation knowledge for a grant application.

For auditing power and practice, the temporal chains support opportunities for identifying new activities, elevations, submergences, the 6Ps, and conflicts between for example participation versus efficiencies, or certainty versus dialogue in evaluation processes. These choices will depend on specific needs-to-know, organisational strategies, and the wider network around the temporal chains of specific partners, funders, NGOs and other agents.

NGOs, evaluators and other stakeholders can reflect upon Figure 7.5, and ask questions related to different activities. Who decides on the need-to-know impact? Where do tools and methods come from, from business discourse, academic models, or local communities? How is data bifurcated? Are important incidents, needs or services elided? What kinds of expertise are bundled into reports? Are critical views of the aid chain, political negotiations, power sensitivities, or collaborations built into evaluation activities, or are fast and efficient market logics dominant? And what impact audibles or silent transformations of agents and politics are produced that partners should track and be sensitive too?

The aim of the generic temporal chain in Figure 7.5 is to offer an alternative view of impact evaluation processes, which welcomes discussions of power and practice. It contests the scientific rationales, commercial pragmatics, and participatory performances of other models which risk silencing aid chain power/data/knowledge relations. The temporal chain also

clarifies activities and information points which inform Floridi's (2015) discussion of "infraethics" regarding evaluation processes, as signposted at the end of the chapter 2.
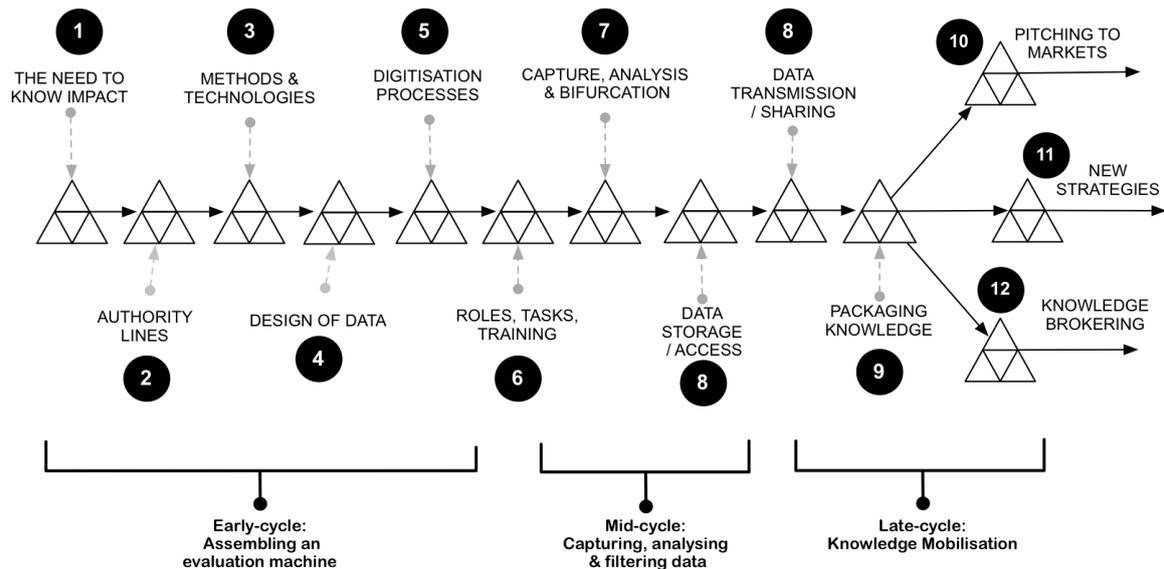
Figure 7.5: Generic temporal activity chain showing kinds of evaluation activities

Step 5: Zooming out to a wider network of agents and activities

Temporal activity chains were developed during the theory chapter and used to analyse the case NGOs. Networks of activities were also described in the theory chapter and they have an established history in CHAT research (Engeström, 1987: 103; Blackler et al, 2000: 281; Karanasios & Allen, 2013: 300). A network view is critically important to overcoming the limited gaze on target sites brought about by managerial, marketing and TIEK frames. Even participatory evaluation shares this gaze, where participation is trapped in sites of target communities and beneficiaries, blinding a wider view of power, participation and knowledge flows in the network.

"Development practice is compartmentalised to field practice, and is not allowed to permeate the organisation as a whole" (Power et al, 2002 in Rowlands, 2003: 94-97)

Network perspectives are critically important to see how local organisations might innovate in development 2.0, as did the first case NGO by learning evaluation and using data to broker new relationships. Madon & Sahay (2002) describe this networked power/data/knowledge dynamic at a local organization in Bangalore which supports slum dwellers.

"… networking logic substantially alters existing processes of production, experience, power, and culture. Prior to the establishment of Jana Sahayog, basic information about slums was produced by the government and was neither shared with other organizations nor made available to slum dwellers in a way that they could understand or respond to. In this way, power rested with the authorities and the slum dwellers had no say in formulation of policies and programs that impacted them. Since Jana Sahayog came into existence, information flow has gradually increased in the direction of slum dwellers, and vice versa, from slum dwellers to government agencies. This two-way information flow has altered the power equation in favor of the slum dwellers." Madon & S. Sahay (2002: 18)

In CHAT, networks of activities model how one activity or system influences another, shaping the elements (e.g. tools, subjects, rules, objects of work or outcomes) of neighbouring activities in the network (Engeström, 1987: 103; Karanasios & Allen, 2013: 300). However, in CHAT research such networks have normally contained relatively few activities, whereas in the earlier literature chapter and case study findings a wide diversity of uncoordinated agents was found to be influential in development evaluations. Zooming out to the network of wider influences is suggested by many practice-oriented observers (e.g. Allen et al, 2011: 781-2; Nicolini, 2013: 219-235; Miettinen et al, 2009: 1321).

This observation stretches the view of evaluation-as-practice to not just target sites and evaluators, dyads, or INGOs and funders, but a broader array of agents and activities distant in terms of geography and time. Managerial, marketing, governing and evaluation prescriptions are similarly diffused across space and time, through governments, universities, publishers, and funders for example. Thus, networks of activities encompass many data/knowledge chains, aid

chain relations and opportunities for control, participation, mediation, or influencing between specific partners and stakeholders. The earlier evaluation dyad and DIKW/TIEK meta-model of rational knowledge construction for decision-makers, cannot account for such diverse activities, agents and influences.

Therefore, a first step in understanding the wider network is to identify the array of agents which influence, mediate, produce or consume different kinds of impact evaluation related data/knowledge. This includes prescriptions, regulations, results, data, arguments, pitches and marketing materials. From the cases and earlier literature, many agents can be identified:

- donor country governments, public agencies and publics
- host country governments, public agencies and publics
- local service providers, such as schools, hospitals etc
- target participants, communities and leaders
- local civil society organisations e.g. NGOs, CSOs, MBOs
- international organisations e.g. INGOs, multi-nationals, management consultancies
- vendors e.g. technology vendors entering the aid sector
- evaluation organisations and professionals
- communications and marketing organisations, consultancies, endorsers
- evaluation and development training providers
- universities, research institutes and researchers
- influential development organisations e.g. UNDP, World Bank, large INGOs
- specific evaluation bodies or regulators e.g. 3ie or OECD DAC;
- diverse investors e.g. private individuals, philanthropies, states, INGOs, and publics

These agents are shown in Figure 7.6 below, along with an indication of their neighbouring data/knowledge exchange relationships with other agents that they may interact with regarding evaluations. For example, recipient nation governments may publish reports which publics read or media organisations write articles about, or evaluation bodies such as the 3ie may work with closely donor governments. As highlighted at the end of chapter 4 (section 4.4.7), researchers and universities are part of this network too. The network, in Figure 7.6, shows the diffusion of data/knowledge exchanges which shape contemporary impact evaluations, however it does not highlight particular activities, the objects of work shared or contested by neighbouring

partners. As with all diagrams, much is left out, however the network is more representative of evaluation data/knowledge flows than the basic evaluation dyad in step 2.

Evaluation network of agents

A network of agents (e.g. NGOs, donors, governments etc) shapes impact evaluation practices and data/knowledge products. Each seeks to sustain and extend their own influences. Each exchanges data/knowledge with various other agents, and shapes evaluation over time through their activities. Agents are show below as triangles with descriptive labels.
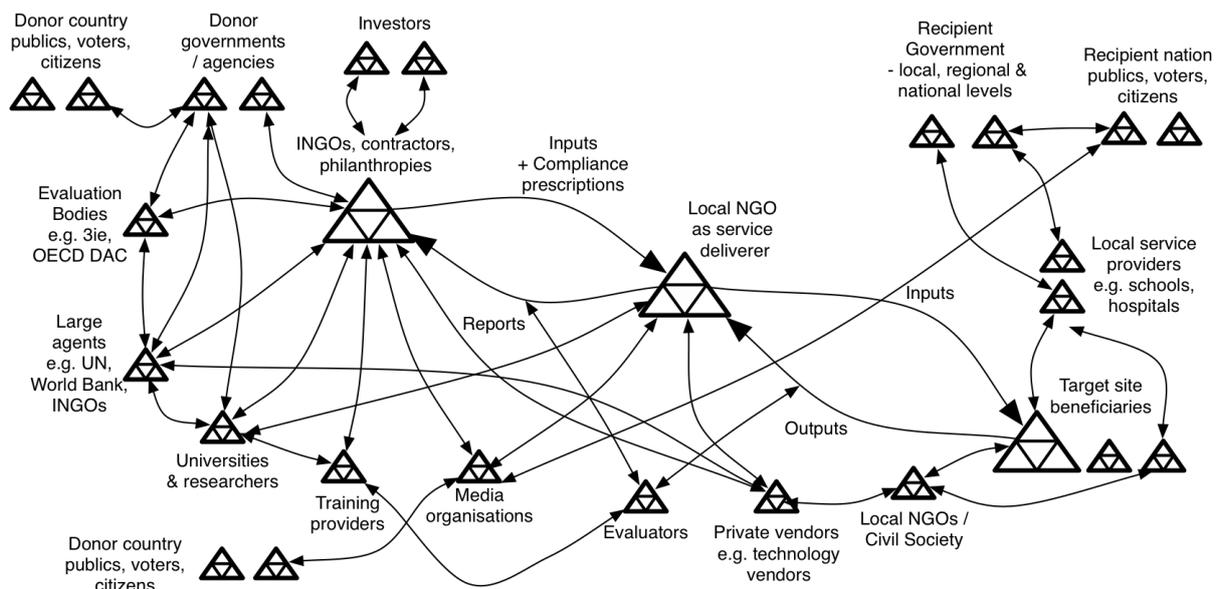
Figure 7.6: Network of agents related to development impact evaluation

In order to approach a better understanding of power and practice between specific agents in the network, there is a further need, informed by CHAT, to identify shared objects of work (Engeström, 1987; 2009; Blackler et al, 2000; Allen, Karanasios & Allen, 2013), overlapping goals or strategies. This points to how and in what ways specific agents shape evaluation practices, small NGOs, and governing processes. This shaping is diffused, networked across diverse situated agents. Example objects of work by different agents are indicated in Figure 7.7 below, however the zoomed in temporal chains for all agents cannot described in this study as the focus is on small NGOs. Future research would have to look at specific areas, such as between global regulators (e.g. 3ie, OECD-DAC) and influential governments (e.g. USAID, DfID UK), or training organisations and individual professionals.

The result is a complex shifting network of data/knowledge production, mediation and consumption activities, wherein specific agents increase their expertise in niche areas. The aid

chain (Wallace et al, 2006), is a core part of a wider network of activities that shape evaluation practices, power relations and data/knowledge products. Examples network activities include:

- Global evaluation bodies set standards to influence evaluation agents

- Universities seek research funds and training opportunities, producing prescriptions, critiques, and commentaries which amplify or delegitimize ways of knowing impact performed by competitor agents

- Governments draw on experts to establish evaluation criteria to regulate the sector, and importantly seek the support of citizens through public statements and ministry policies. Currently, value for money (VfM) is prized in these spaces.

- Policy makers assess evidence on how it supports, changes or rejects current policies and positions, whether it offers opportunities or pitfalls for new policy directions.

- Technology vendors and training providers seek to create markets for their services and products, promoting best practice, sub-domain expertise, data management systems and digital tools as needed to legitimately know impacts

- Beneficiary communities exchange knowledge about local change and views about interventions, but do not necessarily strategize, digitize or solidify impact messages for governments or international funders.

Evaluation network activities / outcomes

Agents perform diverse activities to exchange data/knowledge, which make "outcomes" that influence evaluation. Governments make policy and seek public legitimacy. Investors seek value for money. Media agents broadcast stories and ads. Vendors create markets. Universities seek research funding. Rudimentary examples of these exchanges are shown in grey boxes between agents below. The agents themselves are shown as labelled triangles.
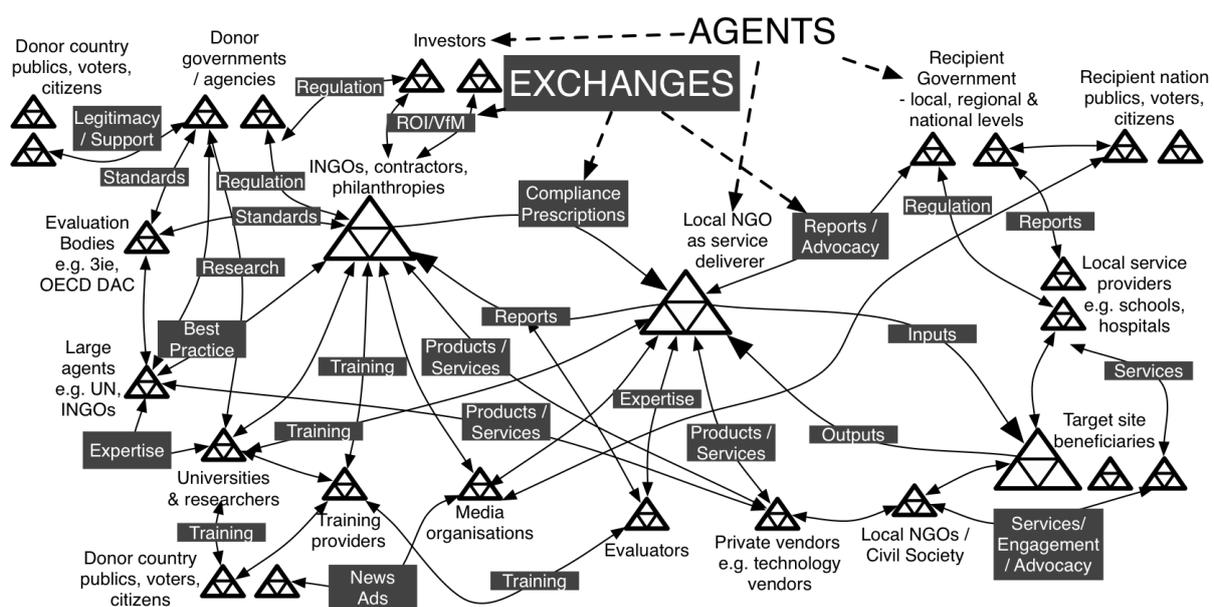
Figure 7.7: Network of agents and indications of their activities and outcomes

Audit 2.0 data/knowledge networks share affinities with stakeholder mapping or outcome mapping (Earl et al, 2001). These techniques, established in development evaluation already, extend network views beyond the evaluation dyad. However, there are five differences between audit 2.0 data/knowledge networks and other evaluation tools.

Firstly, the networks articulate overlapping practices for discussion, rather than support a single agency's strategic positioning. Secondly, they draw on what agents do, and can use terms or vocabulary from approaches to social practice, such as CHAT. They do not commit research to serve aid sector practitioners in their existing tasks, a use of the term "practice" that is common in aid sector contracting. Thirdly, data/knowledge networks focus on power and require dialogue about who exerts power over who, particularly regarding data/knowledge work. Fourthly, the networks focus on not only stakeholders and interests, but data, evidence, concepts, prescriptions and methods – how these can elevate or elide participation, certainty, authorities, multi-voicedness etc. Fifthly, critical networks expose conflicts and dilemmas for evaluation dialogue and discussion, for example where marketing pitches distort scientific rigour, or where participation is submerged for business efficiency.

Thus, this audit 2.0 device legitimizes issues that are developmental between stakeholders, that configure their ongoing relationships, and lessons that concern a wider terrain than TIEK prescriptions or DIKW rationales. Data/knowledge networks provide tools for reflection, question market values and governing controls, expand limited accounts of participation imprisoned in target sites, and help locate spaces to contest impact silences.

7.3.3 The 6P sensitivities

The 6P sensitives were outlined in the theory chapter, as an alternative to TIEK and DIKW legacies which both stressed analysing inputs and delivering outputs as representational results for decision-makers. This input/output mode (Hayes & Westrup, 2012) was surmised as a meta-theory of evaluation. The 6Ps described an alternative way of conceptualising and re-inserting practice and power into evaluations. They extended the study concerns with data/knowledge

work to incorporate data/knowledge products, processes, power, political participation and practice. This section reviews the justification for the 6Ps from the theory chapter and asks two questions. Firstly, how effectively did CHAT render the 6Ps in the case NGOs? And secondly, are the 6Ps useful in addressing the marginalisation of power and practice in impact evaluation data/knowledge construction?

In terms of a brief review, the theory chapter described in detail why the 6Ps were important. Products are included because they are real and tangible responses to aid sector professionals. Practitioners are under pressure to deliver impact arguments, reports, charts, and data products. Products can therefore function as sites, tangible to practitioners, where critical reflection can begin. Processes are important because they are also explicit enough for professionals to consider critical insights or reflections e.g. around knowledge socialization (Nonaka & Takeuchi, 1995), or where data is organised into information (Davenport & Prusak, 1998; Dalkir, 2011). Power is critical to understanding data/knowledge construction in development, it is "*endemic to all human activities*" (Walsham, 2001: 56), both productive and destructive (Foucault, 1991: 194; Gaventa, 2003: 2), inside and outside of target intervention sites. Political participation was considered important to re-insert into evaluation because technical models, DIKW legacies, governmental (Hayes et al, 2017) and developmental strategies (Lie, 2015) smoothed over it. Participation and politics are important for inclusivity, but also effectiveness, and as such support the idea of knowledge management climbing out of its malaise (Lambe, 2011: 191). Finally, practice was considered a significant required component of the 6Ps, as a baseline to interrogate what people do, how they interact, construct, consume, share or sell data/knowledge. Without a practice-based lens and vocabulary, such as in CHAT or other accounts of organisational practice (see Spender & Scherer, 2007: 14; Nicolini, 2013; Reckwitz, 2002; Miettinen et al, 2009), development evaluation would appear as a normalised, un-problematic, scientific, commercial, or egalitarian endeavour. Practice helps us pay attention to politics, power, participation and mess (Law, 2004: 18-19), rather than reducing impacts to legible, audible results.

Firstly then, did CHAT effectively account for the 6Ps in the NGO cases? The answer here is a clear "yes". Products are found in the tools or outcomes of activity systems, as spreadsheets that hold data, or data that is captured and bifurcated, or reports that are packaged as outcomes in the temporal chains. Processes are understood as transformations which lead to outcomes, for example the incremental implementation of a survey through sequenced activities from

design, to question making, to analysis. Power relations are understood in CHAT by analysing the use/exchange value inherent in contradictions, or in conflicts and dilemmas experienced by stakeholders and communities. In case one, the contradiction was between Impact-1 and Impact-2, but in case two the whole activity system of the NGO as an organisation, being transformed towards data/knowledge intensity, illustrates the power of sector governing and market demands, and diverse prescriptions. Political participation in CHAT is understood through multi-voicedness, and how the voices and concerns of different stakeholders, subjects or communities are heard or not, mediated by others or not. In the cases, the clamour of participant framers and the doubts of staff regarding marketing authenticity were relatively submerged, whereas the voices of managers wanting data to share or pitch to investors was elevated. With activities themselves being the CHAT equivalent of practices, it is clear that CHAT renders the 6Ps with clarity. The result is informative for CHAT researchers promoting research into power dynamics (e.g. Blackler 1995; 2011; Simeonova et al, 2018b; Kelly, 2018)

Secondly, are the 6Ps useful for auditing impact evaluation data/knowledge? The answer here is "yes, but". The "yes" refers to the value of using the 6Ps in contrast to TIEK models or DIKW legacies. The "but" refers to the difficulty of legitimating a 6Ps approach, or audit 2.0 devices more generally. Firstly, the 6Ps do have value as an alternative way of viewing data/knowledge work. Rather than surfacing results, which espouse claims of impact, effectiveness, efficiencies or policy influence, aid agent marketing pitches, consultancy reputation building (Brigham & Hayes, 2013), or squabbles over control of data (Markus, 1983), the 6Ps surface processes and practices upon which dialogue can be built. The 6Ps in the case studies were evident in a number of ways, for example:

- the emphasis on databases, sexy titles, reports and convincing stories (products)
- the need to have fast data collection, uploading, storage and transmission to the funder home country, or the incremental transformation of an NGO (processes)
- the use of cells to frame findings and send them rapidly to international funders, or the brokering of data to build shared relations between an NGO and a public service provider (both forms of power)
- the choice of statistical techniques, data bifurcation and rejection of clamour, preferred over open interviews with farmers to hear, collect, translate and analyse stories; profiles of farmers/stakeholders edited to show successes (problems of political participation)

- the use of CHAT to conceptualise and identify specific activities, activity chains and wider networks of activities (practice)

Figure 7.8 illustrates how the 6Ps are critical relations in the configuration of impact data, information, knowledge or wisdom, not in a linear process for privileged stakeholders, but as diverse intangible assets and processes. These are core elements in the development 2.0 landscape. Assets such as evaluation reports, databases, infographics, truth claims, methods etc (see Williams, 2014: 18 for a more extensive list) are continuously shaped by the 6P as they are constructed, moved, mediated, and exchanged across aid markets and bureaucracies.

The 6P Sensitivities

The 6Ps highlight how data and knowledge are not apolitical, logical products, void of historical processes.

For example, an impact evaluation report is constituted by all the 6Ps, as are evaluation methods, spreadsheets, truth claims, and software etc.

The DIKW legacy and technical evaluation models have unfortunately hidden such relations from clear view.

PRODUCTS · PRACTICE · PROCESSES · POLITICS · POWER · PARTICIPATION

Concepts · Beliefs · Software · Events · Branding · Truth claims · Values · Data sets · Reports · Methods · Facts · Designs · Evidence · Records · Technologies · Results
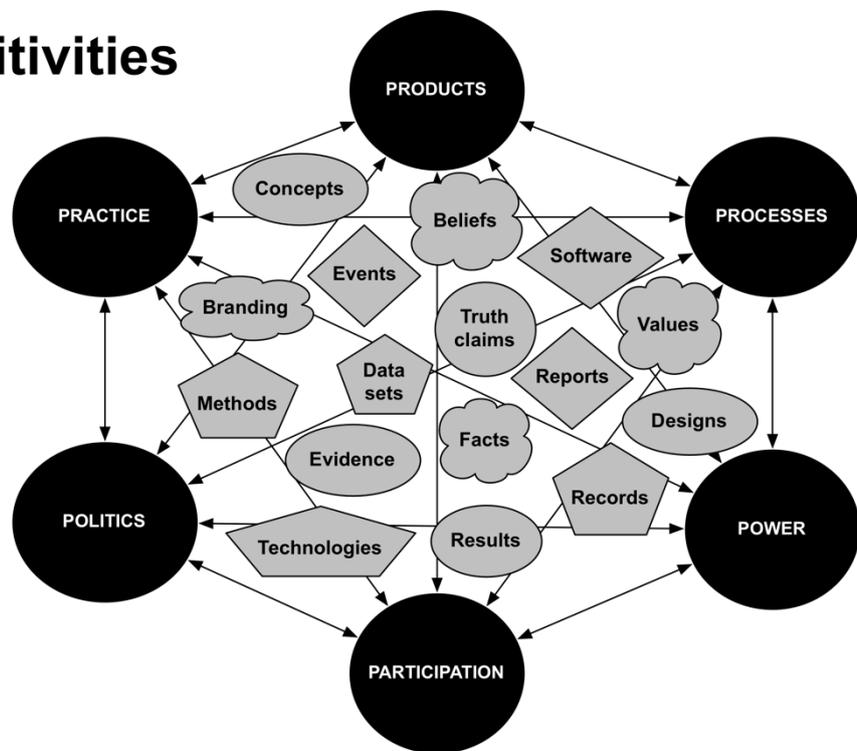
Figure 7.8: The 6Ps shape impact data, information, knowledge and wisdom

Now to the "but". Despite the 6Ps helping to make power and politics audible in the collaboration, the marketing and governing agendas of the NGOs, controlling evaluations to make sure they support funding pitches and applications, made power, politics, and practice difficult to deal with, and problematic from the NGOs' perspectives. Various literatures predict this result (e.g. Wallace et al, 2006; Banks et al, 2015). At the expert NGO, diverse expertise in data management, business strategy, software design, development and evaluation were

viewed as goods. That these goods could entail problems and ways of de-voicing others, or riddled with assumptions and omissions, was not entertained by the NGOs or funder managers. This may be a limitation of the research, but the expert NGO claimed they had tried qualitative evaluation in the past, finding it inefficient, confusing and ultimately not of value to impact marketing.

The novice NGO was more open to networked sharing with different partners, and also resistant to a marketing orientation. However, the CEO and the demands of the sector for sustainable finances had already begun to transform the NGO away from humility and engagement towards data/knowledge intensive work, evaluation and marketing. The 6Ps were evident in the cases, difficult to address, and problematic for both NGOs.

In summary, the 6Ps as an audit 2.0 device contrasts with TIEK input/output meta-models and the "methods plus results" scaffold for decision-makers. The 6Ps mature the research aim of re-inserting power and practice into evaluation (research question 3). They align with critical engagement and the lessons learned from the impact iceberg, CHAT vocabularies and concepts. However, despite being useful for understanding issues of power and practice in impact evaluation data/knowledge construction, case one suggests the 6Ps may not be palatable when parties that try to use them are not ready or willing to reflect on their own power relations, data/knowledge practices, impact marketing, or governing rationalities. Digital evaluation data may drown out inarticulate voices or doubts. New fashions, such as Big Data or Data Analytics may drown out reflection on more inclusive or older ways of doing evaluation, (Leal, 2007).

The 6Ps can be useful devices to articulate the problems of such fashions in evaluation and beyond, however much will depend on the network of agents, and whether doubt, contexts, beneficiary lifeworlds, clamour, novices, omissions, assumptions, and multi-voiced participation are viewed as part of development or illegitimate contingencies that become elided by the evaluation meta-model and method plus results scaffold.

In choosing to use TIEK or the 6P sensitivities, evaluators, researchers in ICT4D or KM4D, and donors have a choice to make – to elevate and market representations of success or to amplify the quiet omissions, networks, changing practices, and 6P inequalities. The 6Ps suggest that it's not just digital representations that are results. Impact products, new processes, shifting power, participation, politics and changing practices are real results of doing evaluation too.

7.3.4 The impact spectrum

It is not only changes to evaluation, NGOs or governing that are submerged by the sector histories, demands, TIEK, and implicit models of data/knowledge construction. It is also the qualities of data/knowledge products, processes and practices. An audit 2.0 impact spectrum, described below, contrasts more solid, audible and technical qualities, against more messy, silent, diverse and illegible qualities. This section explains what the spectrum is, describes its utility, and suggests future development.

The impact spectrum is a device to juxtapose solid, audible, rationalised and legitimised qualities of impact data/knowledge products and processes, with silenced, time and place specific, multi-voiced, marginalised qualities. The spectrum draws on how data/knowledge is constantly mediated and transformed (Engeström, 1987; Miettinen et al, 2012) and how diverse values or "goods" exist beyond scientific truths or management efficiencies (Law, 2004: 148-151). In effect, the spectrum suggests that solid impact data/knowledge products and process are firstly constructed out of complex, messy social and political practices, and secondly that deconstructing these solid products, to show, talk and respond to them, is a worthwhile venture

The spectrum shows how governing processes construct legitimate/illegitimate data/knowledge in development qualities (Hayes et al, 2017). Such processes align with neoliberal tactics, audit cultures, and the reduction of complexity, politics, voice and context which collectively narrows data/knowledge design qualities to discrete, rationalised, expert, professionally mobile and countable qualities data (Strathern, 2000; Miller, 2004). The constructed products then fit managerial and market agendas. The DIKW legacy of collecting volumes of narrowed empirical data to build pragmatic results for influential decision-makers (Weinberger, 2010; Zeleny, 1987) contrasts with seeing data/knowledge as part of contexts, doubts, alternative ways of knowing, negotiating and contesting. Governing data/knowledge may be legitimate within professional impact networks, but this legitimacy displaces other peripheral, truncated or messy qualities of data/knowledge (Polanyi, 1967; Scott, 1998; Lambe, 2011; Law, 2004). This silent mess is nevertheless diverse, meaningful and impactful.

The impact spectrum assumes that messy practices, socio-political relations, and multi-voiced accounts of changes and impacts circulating in communities and in project staff views, provide raw materials for expert led impact machines to pre-shape, extract, analyse and bifurcate, to construct robust, solid, rational, exchangeable, packaged narratives of impact. Such solid and

certain accounts can be questioned by amplifying the silent qualities and contingencies, inherent in how impact commodities were created[34]. The spectrum stimulates questions around what qualities are elevated or submerged in specific evaluation processes and products?

Such questions are important for evaluation, KM4D and ICT4D. They concern views of data, information and knowledge as related to situated activities (Suchman, 1987), negotiations, and contestations in contrast to data/knowledge being bounded things, assets or products held in brains, books, or databases (Blackler, 1995). The impact spectrum advocates a form of evaluation with more negotiation, contestation, facilitation, and sensitivity to the politics of development evaluation itself, making prominent what Law (2004) calls ontological politics. Figure 7.9 shows the spectrum.
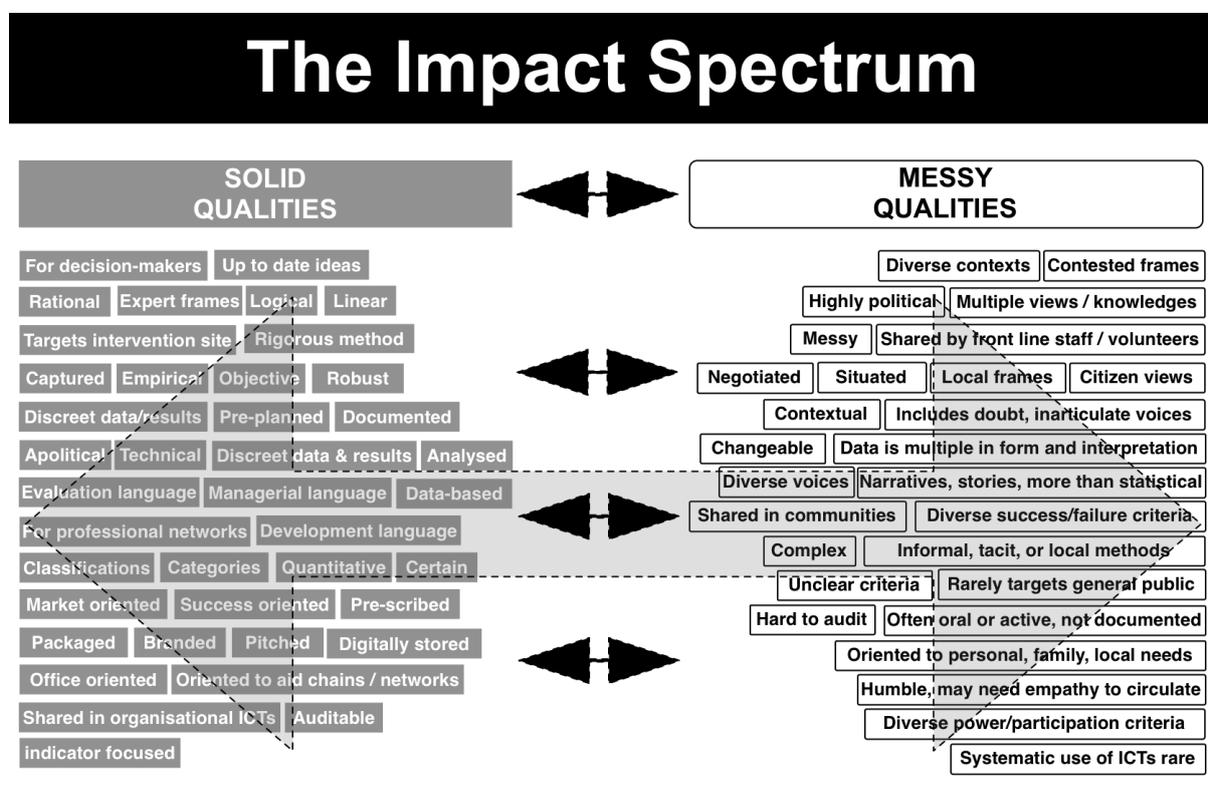
Figure 7.9: The impact spectrum showing solid versus messy qualities

---

[34] Randomised Control Trials were once conceived of as a "gold standard" of robust, solid and scientific evaluations. However, even technical evaluators are now seeing their situated contingencies. See Krauss (2018) and Narasimhan & Arun (2017).

Two applications of the impact spectrum are immediately evident. Firstly, it is useful as a tool for encouraging critical reflection in development sector practitioner spaces, evaluation designs and responses to evaluation claims. The iceberg describes a sector wide diagnostic, but the spectrum is best applied as a specific tool that can be used to tease out and reflect on the qualities of impact products under construction and their attendant processes. Stakeholders and evaluation respondents can discuss what voices, goods, doubts, or methods are being elevated or submerged? Do projects and evaluation processes subjugate traditional or community values (Avgerou, 2002: 88-89 Law, 2004: 122-139) in the name of business efficiencies or global rationalities, such as universal human rights? Do they discard inarticulate stakeholders, suppress critical dialogue (in target communities or aid chains), or elevate discrete, digital data for distant stakeholders?

Secondly, the spectrum is also useful for researchers and evaluators to extend or question their own views of legitimate/illegitimate impact products, contracted deliverables, or processes. Evaluators can use the spectrum as a reference tool to see and respond to what messy silences are being generated in an evaluation process that may demand and privilege solid products. The spectrum informs ICT4D, KM4D, CHAT or other practice-based researchers by interrogating data/knowledge flows, commodities and residues. It may be challenging, albeit revealing, to adopt or adapt the spectrum within technical evaluation frames, results-based or evidence-based approaches, as these have so often silenced the right-hand side of the spectrum. For CHAT in particular, the spectrum adds to recent work on power, mediations and transformations (e.g. Engeström, 2009; Blackler, 2011; Karanasios 2014; Karanasios & Allen, 2013; Kelly 2018). The chief insight of the spectrum is it adds to distinctions between data/knowledge criteria legitimated by market/governing logics on the right side and alternative goods, values and sensitivities to power/data/knowledge evident in the qualities of the left over, neglected residues of data/knowledge machines.

To conclude and with regard to maturing the impact spectrum audit 2.0 device in future, four issues are pertinent. Firstly, readers may recommend that specific evaluation methods e.g. RCTs or Outcome Mapping be positioned on the spectrum. However, methods and techniques, in practice and over time, may move to more or less certainty, more or less multi-voicedness in particular situations. It would be difficult to place a technique on the spectrum without investigating how it was assembled in its own temporal chain of activities.

Secondly, when impact evaluation products, reports, and data sets etc are exchanged, they can become part of further complex, political practices within aid bureaucracies. For example, despite a solid impact product with a rigorous method being in place, personal relations with funders may be more important than the rigorous report. The spectrum highlights how such interactions, here a personal relationship, remain messy and silenced, mediated out of legitimate representations.

Thirdly, the simplified binary spectrum of products and messy practices requires further work, in terms of analytical depth and multi-voiced review. However, future use and research should pay attention to the "dilution" challenge, that the spectrum remain mobile and accessible, but also critical, and the "developmental" challenge that the spectrum's role remains to stimulate (Salancik, 1978) or probe (Gaver et al, 1999) data/knowledge practices like evaluation, not to provide a blueprint prescription.

Finally, it is worth noting that the dynamics represented by the spectrum will likely intensify in future, with more extreme governmental and participatory evaluation approaches emerging. The awe of the left side in management, markets, academia and government, will be sustained in the foreseeable future. The spectrum questions this awe beyond the trust in numbers (Porter, 1996), by questioning the power and awe of solid, packaged data/knowledge products.

7.3.5 Datamentality and datamateriality

The research suggests that the dominant mode of thinking in development evaluation is centred on knowledge products (Mosse, 2004a: 77; Williams, 2014). These carry representational evaluation results that can be legitimately circulated for decision-makers in development bureaucracies and markets (Wallace et al, 2006; Harper, 2003; Quarles van Ufford, 1988).

These knowledge products are exchanged between development NGOs, evaluators, practitioners, funders and governments, but are to a lesser extent also available to publics, partners, researchers and target communities or participants (Roche, 1999; Groves, 2015). This product view (Blackler, 1995) is foregrounded in TIEK, DIKW and much data/knowledge management. The view is instrumental in governmentality, where products need to be created and exchanged in competitive markets (Lemke, 2002; Hayes et al, 2017), but it is also implicated in the information profiling, knowledge networks and smooth strategies of

developmentality, albeit participation and partnering processes are added to this more recent perspective (Lie, 2015). From a development 2.0 perspective, capacity, capacities and skills and resources (e.g. Heeks, 2008; 2018) to make processes and products are the legitimate elements for making knowledge products to market. This narrowing of data/knowledge to market exchange conforms with Vygotsky's critique of the methods/results function of scientific endeavours, which elide changes in unfolding practices. In contrast to the method/results frame, Vygotsky termed this a developmental process (1978; Holzman, 2006b), which includes the agents and practices that change, not just the end or products. The effect of this narrowing audit 1.0 gaze is that the contingencies, politics and unexpected impacts upon evaluation, NGOs and governing are not in the professional view. The lost data, farmer clamour, doubts about accurate database statistics, the inability to contract a professional evaluator, the out-of-scope housefire impact, the changes in staffing etc, are too local, or too material to travel and be packaged for markets. However, they still do occur, they are still impactful and developmental.

In contrast, audit 2.0 focuses on process views of evaluation, what happens during evaluation interactions, the situated data/knowledge practices and contests (Blackler, 1995), and how agents like individuals or NGOs change and develop (Vygotsky, 1978; Holzman, 2006b). This is the developmental view, not the results view. Such a view of knowledge as processes, is consistent with CHAT and with practice-based approaches. Practices imply material relations such as who sits down where to evaluate, plan, design, analyse, what format are reports in, what information systems shape data, store information or send knowledge, and where does the knowledge land? They also concern relations about who, when and where we think and conceptualize, using ideas, methods and theories. Thus, we have both mental (e.g. design thinking, evaluative thinking, systems thinking, development thinking) and material concerns about evaluation processes. If we connect these ways of thinking, acting and governing, to the demands for data as a foundation for knowledge and data framed by experts as a form of evaluation control, we arrive at the concepts of datamentality and datamateriality. Datamateriality as part of audit 1.0 amplifies impact governing and produces socio-political inequalities.

Datamentality is related to distal evaluation prescriptions and methods, to the governing of impact knowledge, not through expert authority, but the expert deferral of authority to claims of factual objective data. The objective, empirical data is seen to dictate evaluation results and

claims. Evidence-based policy (Pakhurst, 2017; Hammersely, 2005) and Results-based management (e.g. OECD-DAC, 2002) depend on this rhetorical shift from expert-as-authority to data-as-authority. DIKW (Zeleny, 1987; Davenport & Prusak, 1998) and New Public Management (e.g. Townley, 1995; Hood, 1995) both support the view of evaluation as driven and grounded by empirical data. However, in Tuomi's (1999) reversal of DIKW, pointing out the need for extensive knowledge and values which are needed in order to make any data, and Blackler's (1995) view of knowledge (and data) as situated, contestable, and negotiated, the audit 1.0 claim of neutral objectivity, of datamentality, is found wanting. The audit 2.0 devices problematise datamentality, and elevate an alternative, datamateriality.

Datamateriality, aligns with audit 2.0, and is a notion that supports seeing data as constructed over time in specific material activities, through specific design processes, technologies, socio-political tactics, and material spaces, places and channels, including digital technologies. A datamateriality view acknowledges the complex relations which pre-shape empirical data captured, extracted and bifurcated in evaluation processes (Tuomi, 1999; Fricke, 2009; Lambe, 2011). In impact evaluation, the materiality of the evaluation process is often deleted when results are moved into focus for decision-makers, managers or market exchange. This displacement is common in scientific modes of data/knowledge construction (e.g. Law, 2004: 20). In doing scientific lab work, Law (2004: 20) quotes Latour & Woolgar (1986: 51) concerning how the process, activities and material relations of doing data and knowledge work "*melt into the background*" when results become the object of attention. In CHAT, this is understood via mediation and transformation, how data is mediated by rules, technologies and subjects, and transformed into outcomes which feed neighbouring activities (Miettinen et al, 2012, Kelly, 2018). In doing evaluation work, in order to see the practices and power relations, datamateriality must be asserted, so we can see who, how, where and with what tools and concepts evaluations are done, and the impacts of these doings. There issue again is *what happens* (practices), *not what works* (solid results).

CHAT encourages researchers and practitioners to pay attention to tools, rules and norms when making data. Information Systems researchers also focus on how social relations are embedded, entangled or imbricated with technologies and materialities[35]. For evaluation, this

---

[35] How such social and material concerns influence each other is a matter of lively debate in organization studies. See Orlikowski (2007) and Leonardi (2011).

means paying attention to the construction of the impact machine, its tools, places, authorities, and how assumptions, designs and technologies shape data/knowledge in practice (Kelly, 2018). Marketing software, discrete data cells and Android tablets cannot transport evaluation contexts and complexities to London. London is a different situation, a different materiality. The received data/knowledge does new work there. Managers and donors must reflect on their own practices, needs, and how these shape evaluation designs, capture, bifurcations and results.

Both datamentality and datamateriality are propositional concepts from the study. However, they enable us to begin to unfreeze the iceberg's composition, to ask questions about evaluation products, processes, conditions and claims, and to ultimately understand more about the incremental, developmental impacts of data/knowledge chains and networks. In development 2.0 today, impacts are generated by shifts towards discrete data designs, adoption of marketing techniques and digital technologies, importation of blueprint methods, bifurcation of legitimate/illegitimate data, expert packaging of knowledge, and pitching, or brokering final products. Each of these create new capacity inequalities in the sector and make more power sensitive ways of evaluating, clamour, doubt, and beneficiary lifeworlds more illegible, more discardable. These are problems of datamentality, which can be articulated through dialogue about datamateriality, i.e. what is happening in evaluation today.

7.4   Contribution 3: Critical engagement

This section reflects on how Critical Engagement and CHAT featured in the study, to arrive at a minor contribution for future research which seeks to simultaneously critique and engage. The contribution stresses the tensions or vibrations between critique and engagement as constitutive parts of the collaboration process (Figure 7.10). The research attempted to critique NGO impact evaluation in practice, but also to engage with NGOs as partners. This meant navigating the difficult territory between critical research on aid sector power dynamics and offering consulting advice for NGOs. The engaged consulting secured access to hard to reach NGO field work and enabled the study to probe questions about how critiques may be received or rejected by small NGOs.
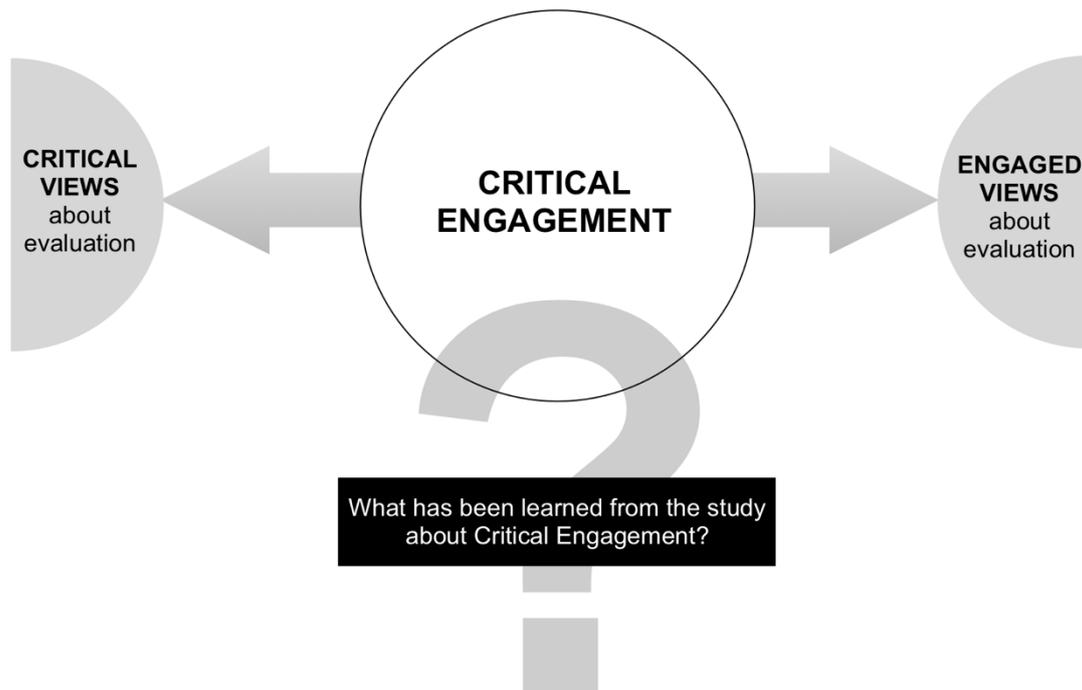
Figure 7.10: What has been learned about critical engagement from the study?

Two questions are important. Firstly, how successfully did CHAT render the power/data/knowledge relations? And secondly, how viable is Critical Engagement? The main audiences for this section include: evaluation practitioners, researchers in development evaluation, ICT4D, KM4D, and CHAT or practice-based researchers.

This critical engagement problem was articulated in the methodological framework and concerned diverse but related points. How do outsiders engage with development settings (Ferguson, 1994)? Do anthropologist activists damage local cultures (Gardner & Lewis, 2015)? Do radical perspectives become moralizing, polemic prescriptions (Avgerou, 2010: 12), de-radicalized (Mosse, 2004a), or decapitated (Leal, 2007)? How can critiques support reform, but avoid being diluted by managerial controls (Fournier and Grey. 201; Alvesson and Willmott, 1996: 18; Anthony, 1998: 279; Burrell, 1993). Commentators advocated for research to step outside the bounds of scholarly exchange (Marcus, 1998: 204, 228), for aid researchers to analyse the aid apparatus (Henkel & Stirrup, 2001: 183), and for a closer focus on expertise (Gibson-Graham, 2006: 192), agency, power, political economy, and granulated practices (Escobar, 2008; 202-203). In describing what has been learned in the study, the following pages conclude that critical engagement requires specific kinds of tactics.

7.4.1 CHAT tensions and responses

How successfully did CHAT render the development evaluation power/data/knowledge relations as part of the critical engagement approach? Four tensions are presented below, to illustrate how CHAT mediated between critique and engagement.

A first tension was the need to zoom in to detailed activities, but also zoom out to wider networks. Other commentators have also argued that practice-based approaches need to zoom in and out of research sites Nicolini (2014), and that practice should focus on micro-level agency and wider issues like political economy (Escobar, 2008).

There was tension in the lens switches, in moves from the field to literature, theory, networks, and back to the field sites. Analysing the NGO settings, from cell data, yes/no questions, evaluator stresses, and humble attitudes, to sector wide NGO service contracting (Banks et al, 2015; Wallace et al, 2006) and audit cultures (Strathern, 2000) required ways of zooming in and out of the empirical site.

CHAT's ability to zoom in and out (Karanasios & Allen, 2013: 300) helped in navigating the tensions. CHAT was used to navigate these tensions by switching lenses from the temporal activity chains to understand micro-level activities (Kelly, 2018) to wider activity systems and networks (Engeström, 1987: 103; Blackler et al, 2000: 281).

Engeström (2009) has argued that the activity systems frame requires continual development. Schools or hospitals for example, are sites of both local and global influences (Hedegaard et al, 1999: 14). Activity systems articulate elements and contradictions, but incremental mediations, such as bifurcating data, packaging knowledge, and transforming NGOs, are diffused across agents, time, activities and relations in longer chains and wider networks. This is why, in contrast to many CHAT based studies, the temporal chains highlight sequential micro-activities, rather than central systems. Zooming out to networks and diverse agents, and into chains of specific activities supports a more fluid view of complex overlapping activities. This fits with Engeström's (2009) discussion of contemporary institutional change, runaway objects, wildfire activities, temporary knotworks, and shifting objects of work in globally connected phenomena today. With data/knowledge representations travelling so rapidly and broadly now, understanding what travelling rationalities (Craig & Porter, 2006), subjugating

knowledges, methods, results and prescriptions do when they land in specific practices and activities is increasingly important. Cost Benefit Analysis (Porter, 1996) and Human rights (Tsing, 2004; Mosse, 2013: 232) were local innovations, which over time became global rationalities. The NGO cases showed how zooming in and out supports an understanding of how and why transformations of processes (e.g. evaluation), agents (e.g. NGOs), and sectors (e.g. development) unfolds.

A second tension concerned CHAT's own vocabulary. Concepts, such as "activity systems", "elements", "subjects", "objects", "contradictions", "expansive learning" and "transformation" were not used explicitly with case participants, because time and opportunities to meet, research and discuss were limited. CHAT terms were replaced in the field by more everyday or development terms such as "stakeholders", "participants", "networks", "ideas", "aims/goals", "technologies", "learning", "reports" or "data". CHAT terms are specialized, and the professionals had not been trained on CHAT, approaches to social practice or critique. English was a second or third language for many of the staff at the second case study, and although the author spoke Thai and there were other English speakers, technical terms from CHAT or evaluation itself risked confusion. From a CHAT perspective, such translations were a mediation (Engeström, 1987; Kontinen, 2007: 140-144) to support dialogue between the author and NGO participants.

A third tension concerned probing or testing NGO ways of working, to learn about their openness to multiple voices and critical perspectives. The engagements with NGOs did not follow CHAT Change Labs (Engeström et al, 1996; Virkkunen & Newnham, 2013) where formal plans are made for participants, such as doctors, patients, and hospital administrators to repeatedly convene. Interactions were more opportunistic to fit in with NGO availability. Bringing NGO stakeholders and participants together is challenging. Donors are often in international settings. Senior managers in urban headquarters. Evaluation experts need to fly in. The case one NGO management, staff, volunteers and farmers were geographically distributed across central India. Case two NGO management, staff, volunteers, villagers and legal/illegal Burmese migrants were in different locations across urban and rural Thailand. In development, there are significant social, economic, cultural, linguistic and power gaps between participants (Bamberger, 2010: 2), meaning multi-stakeholder forums, require extensive planning and resources. In a small-scale research project, formal change labs were thus not possible.

Nevertheless, CHAT did help to navigate the uneven representation in development evaluation, by conceptualizing how a researcher is an involved agent, who can probe and test alternative or critical insights. The researcher probed and tested assumptions and alternatives with NGOs to understand how they might respond. For example, the author proposed qualitative research in case one as a way of learning about partner activity systems and objects of work. In case two, the author encouraged a network view of evaluation, ran evaluation workshops for staff and drafted an evaluation sub-section in a successful funding bid. These were example ways of bringing evaluation knowledge into NGO activity systems, but simultaneously probing how open to critique, inclusion and diverse voices the NGOs were (Gaver et al, 1999; Salancik, 1976). In this view, researchers are part of the network of activities. Their critical offerings are not accounts of truth or moralizing prescriptions, but experiments to learn about legitimate and illegitimate ways of performing evaluation.

A fourth tension was in how CHAT initiated and supported critique, but did not dominate it. CHAT opens critical analysis with use/exchange value contradictions and activity networks (Engeström, 1987; Blackler, 2000; Karanasios, 2014), but these can be supplemented by other perspectives. CHAT is flexible, it does reject other critical views, but can embrace them, especially through diffused activity networks. In this study, once activities and networks were identified, views from critical Development Studies, Information Systems, and governmentality (Foucault, 1979/2000; Dean, 2010) supplemented the CHAT's analysis of contradictions. Through CHAT's ability to locate specific activities, we could learn more about the locations of data/knowledge edits, elevations, submergences. These can be elaborated upon by drawing on audit culture and the audit 2.0 devices, the 6Ps, DIKW assumptions, governmentality or developmentality rationalities. In this respect, CHAT mediates both towards engagement with participants and towards critical perspectives from other theoretical bodies of work. Effectively, CHAT brokers or mediates critical engagement.

In summary, the four tensions of zooming in and out, translating concepts and terms, testing critical probes in the field, and opening up deeper critiques, all illustrate how CHAT as a framework is well is positioned to gel critique and engagement. It is not overly critical or detached from the field. Neither is it overly engaging or easy to dilute, as it has integral critical tools and concepts. It is versatile, and can be used in ethnographic studies (Kontinen, 2007), action research (Orland-Barak & Becher, 2011) or critical realism (Allan et al, 2013; Mukute

& Lotz-Sisitka, 2012) for example. With these reflections in mind, CHAT can also respond positively to claims about its uncritical consultancy mode (Avis, 2007; 2009).

7.4.2 Critical Engagement: tensions and responses

This section reflects on whether critical engagement is a viable endeavour? It was earlier argued that critique or engagement alone are not sufficient. Critiques could be de-radicalized, decapitated, diluted or captured by management (Mosse, 2004a, Leal, 2007; Burrell, 1993). Critiques may fail to trickle down from scholarly exchanges to development practice (Harris, 2016). Moralizing or prescribing changes that recipients can not satisfy, might lead to rejection, produce negative impacts (Avgerou, 2010: 2, Li, 2007), or increase the data/knowledge demands and commoditized capacity gap. Similarly, engagement alone can exacerbate impact silences. This section cannot definitively resolve such dilemmas, but can offer three lessons learned about critical engagement from the study, namely: the acceptance of failure, expected and unexpected responses to critical engagement, and boundary crossing.

The first lesson is that critical engagement is challenging, prone to failure and therefore creates good learning opportunities. Attempts in the field studies failed to divert NGO managers from donor compliance and market-oriented impact commodity construction. The first NGO was committed to market pitches in their evaluation work. The second NGO saw impact evaluation as strategic for them in stabilizing their volatile funding environment, although they began to broker data/knowledge with diverse partners. This acknowledgement of failing, accepts Ferguson's (1994) claim that outside interventions, including critical ones, may not be warranted. It accepts Blackler's (1995) contention that knowledge, and thus critique is negotiated and contested, not universally agreed or transportable. It accepts that neither evaluation experts nor critics have a monopoly on all relevant expertise. And the acknowledgement also accepts that critical trickle down (Harris, 2016) may increase demands and the data/knowledge capacity gap. This acceptance of conflict and failure in trying to both critique and engage, and the researcher's own position in the wider network (e.g. producing a thesis according to academic regimes) suggests that tensions and vibrations are normal. Development contradictions, such as the six proposed by Kontinen (2007: 10-13), surround the researcher, particularly those caught between generating solid scholarly knowledge for

academic markets, representations of impact for development markets, or immobile critical configurations of messy practices.

In contemporary participatory design (Nicholls, 1999; Mama, 2000) or service design (Edmunds, 2014), tests and prototypes are privileged elements of collaborative learning. Tests, probes and prototypes may well be useful for navigating the tensions between critique and engagement too. In the first NGO case, a proposal to generate qualitative interview data was rejected. In the second case, a proposal to talk with pregnant mothers was replaced by a best practice quantitative survey. The probes both failed, but they revealed some of the quantitative and evaluative thinking, assumptions, practices, and power dynamics in place.

In parallel, it is not just influencing development that may fail in presenting critiques, but influencing academia may fail in proposing engagements. Probing palatable engagement with critics who wish attack the management class (Burrell, 1993) may result in solid, pushback, based on articulate scholarly arguments, values and journal demands upon critical researchers. This response or pushback can be termed a failure, but it is also part of dialogue and learning, brokering between disciplines or sectors. Pushback can be revealing of positions, when aid commentators argue that academic papers are irrelevant, too abstract, lack evidence, or promote conspiracies (Green, 2018), or when scholars criticize activist anthropologists (Lewis, 2005), tyrannical participation (Cooke & Kothari, 2001), or damage to pure cultures (Schonhuth, 2002).

Between the extremes, pushbacks and failures, critical engagement elevates attention to power and practice, as in the 6Ps. A researcher or evaluator seeking to both critique and engage, needs to accept these tensions and vibrations, the pushes and pulls, and make contributions specific to time and place. One of the most prominent critics of expertise and interventions was Foucault himself, who also led a busy activist life. Critical engagement suggests that contributions to research and practice may be different in terms of products, processes, practices and power, may fail in their immediate reform intentions, but may succeed in developing new dialogues and diffused, messy impacts.

A second lesson is how critically engaged contributions may be expected in some ways, but also unexpected. Expected responses were evident in the more technical knowledge exchanges. In the NGO cases, sharing technical knowledge of evaluation, such as how to draft more open-ended interview questions to generate qualitative data, or survey questions to generate more

quantitative data, was largely welcomed. This technical exchange formed part of the researcher's tactic for engagement. However, there was a difference in working with the first expert NGO and the second novice NGO.

The expert NGO were selective with both technical and critical advice from the researcher. The technical advice about qualitative interviews was rejected as it was difficult to process for the NGO, and it risked damaging the philanthropy's impact marketing certainty. The novice NGO were open to both technical and critical perspectives. They welcomed evaluation brainstorming, network maps of potential evaluation consumers. This suggests, that as learners, they were more open to critical insights than the expert NGO.

Finally, the use of evaluation data/knowledge at the second NGO, also had unexpected trajectories. As they started to build a database, frame questions, and engage different organizations, they began to share with a local hospital, doctors, a health authority, and a local network of NGOs. They also involved migrant leaders in early Migrant Project evaluation.

These responses suggest that as data/knowledge capacity, skills, goals and attitudes change, agents will create new objects of work with new partners and groups, often in ways unanticipated by researchers or trainers. As part of critical engagement, this means that one cannot presume to fully comprehend the context of data/knowledge opportunities in the environments of partners or NGOs. Nevertheless, the level of expertise may also be important in determining how open an NGO is to critical insights on evaluation. NGOs and funders established in impact markets may be materially bound to existing relations.

A third lesson is the notion of boundary crossing that critical engagement appears to require. The researcher, in performing engagement and critique, was crossing between research and consultancy, academia and development. The idea that organizations and individuals traverse knowledges, sectors, disciplines and values (e.g. Lewis & Mosse, 2006) was evident.

In evaluation, the notion of boundary partners has influenced approaches such as outcome mapping (Earl et al, 2001) and in development settings, boundaries and fluid interactions have been theorized around how technologies are used by communities (de Laet & Mol, 2000). Boundary crossing has also been used with CHAT (Kerosuo & Engeström, 2003; Daniels et al, 2013), and is related to boundary objects (Star, 2010; Star & Griesemer, 1989). Two examples show how boundary crossing helps us understand critical engagement.

At the first NGO case, a qualitative profile had been created and edited to show positive impacts for a female farmer. From a critical point of view, this was not authentic, it was the voice of impact marketing. In terms of boundary crossing, the researcher had to jump from the marketing to the critical, and back, to understand the text, its function, and how it silenced authentic voices. At the second NGO however, the researcher adopted the position of evaluation expert, training the team on evaluation data, knowledge and marketing. Here, the author was not acting as the critic, but as a consultant. In suggesting that the NGO market their impacts more professionally, one staff member pushed back, arguing that they were not used to this work, that they had more humble motives. In terms of boundary crossing, the staff member was providing the multi-voiced perspective, and the researcher the TIEK prescriptions. Ironically, the roles were reversed.

The concept of boundary crossing helps to understand the different roles the researcher (and participants) adopt, and the different knowledges they draw on. In critical engagement, the move from critique to engagement is strategic and can be immediate in some circumstances. This mediation of roles, used to learn about evaluation or other data/knowledge intensive practices, can be further developed in future by drawing on concepts such as boundary crossing, where mediators work between proliferating specialisms, sectors, professions and disciplines. Although much research has been done on boundary objects and boundary crossing, the key for critical engagement is to help critical insights travel in practitioner spaces.

7.4.3 Eight tactics for critical engagement

When these lessons regarding the use of CHAT and the viability of critical engagement are brought together, critical engagement can be seen to involve brokering between diverse partners, but this brokering requires a critical element if it is to pay attention to practices and power. This is not the academic exchange of critiques to find single scholarly truths, to defend or "win" academic arguments (Law, 2009: 150), or publish certainties in top draw journals but is instead the testing or critical insights in the wilds, here, of evaluation-as-practice.

This constitutes the methodological contribution of the study, that critical engagement requires a specific form of brokering, a form which embeds insights and resources for understanding power and practice, such as the 6P sensitivities and impact iceberg (Figure 7.11).
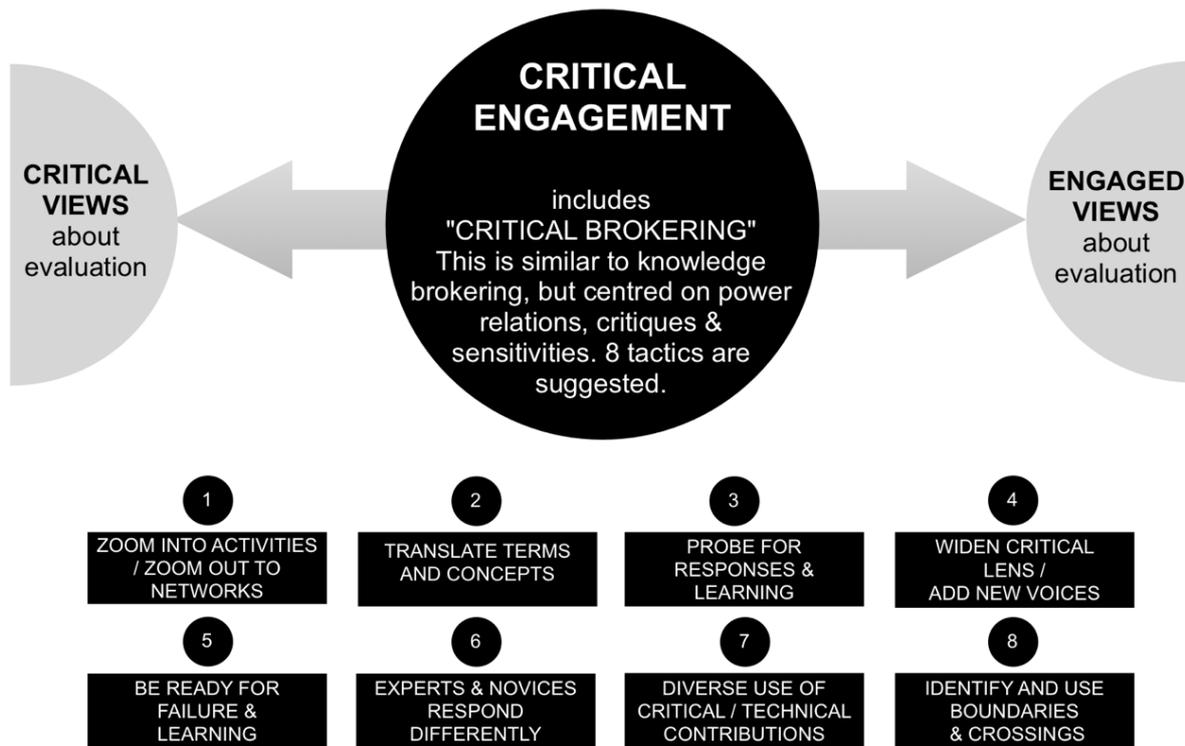
Figure 7.11: Critical engagement requires critical brokering tactics

In the methodological framework, critical engagement was advocated as a way of amplifying and responding to power/data/knowledge silences in mundane impact evaluation work. The study suggests brokering engagement and critique can be supported by 8 tactics:

1) ways of zooming into evaluation micro-activities and out to evaluation networks
2) translating from social theory or evaluation speak into aid speak or everyday terms
3) probing methods, technologies and assumptions via critically informed alternatives
4) strengthening alternatives by adding critical perspectives from texts or participants
5) accepting critiques may fail, but inform learning about evaluation / critical palatability
6) acknowledging that experts and novices may respond differently to critical insights
7) acknowledging that critical or technical contributions can be rejected, adopted, adapted, or lead to unexpected activities as agents increase their expertise
8) Other ideas can support critical brokering, e.g. boundary crossing, boundary objects

7.5 Stepping back to see broad impacts

This chapter has argued for three contributions as responses to impact evaluation data/knowledge narrowing. These three contributions support a view of evaluation-as-practice, and power, which contrasts with evaluation as independent (Picciotto, 2013), expert science (Picciotto, 2012, Duflo & Kremer, 2005), local participation (Chambers, 1994), autocratic, bureaucratic or democratic (McDonald, 1993; Norrish, 2015), or part of developmentality strategies of participation or professional networks (Lie, 2015; Ilcan & Phillips, 2008; Escobar, 1995/2012).

Evaluation-as-practice involves stepping back, reflecting on short-term deliverables, datamentality, results, and impacts, and looking more broadly at the shifting development 2.0 landscape. This brings into focus the transformations and inequalities produced by a sector data/knowledge intensive feeding frenzy around narrow, mobile representations and knowledge products, made for markets and governing practices (Hayes et al, 2017; Kelly, 2018).

8 Conclusion

8.1 Introduction

This study opened with two questions. Has development aid gone from tangible help, food and medicine to a data/knowledge intensive landscape, development 2.0, in which processes such as impact evaluation proliferate? And have increased data/knowledge intensive processes, such as impact evaluation, helped broaden, deepen and speed up our understanding and responses to problems such as poverty and inequality? Claims for the benefits of data/knowledge intensity and better worlds that result are made by many international organisations, technology vendors and researchers (e.g. Quaggiotto, 2007; 2009; 2010; Walsham, 2001; 2012; Heeks, 2008; 2010; Thompson, 2008, Markoff, 2005; Paul, 2010). In contrast, this study has looked at some of the dangers, costs and losses involved in the development 2.0 shift, and thus contributes to making issues of power, practice, and inequality in data/knowledge intensive work more transparent.

The study found data/knowledge to be highly material, manifest in complex temporal activity chains and diffused networks of offices, professional communities, training programs, reports, methods, charts, computers, databases, data cells, prescriptions and diverse processes, such as impact evaluation. Furthermore, the data/knowledge intensive landscape was found not to be neutral or exclusively emancipatory. It was accompanied by governing and audit processes (Li, 2007; Strathern, 2000; Townley, 1995; Escobar, 1995/2012), developmentality claims of equal partnering and participation (Lie, 2015), competitive markets (Hayes et al, 2017), and a commoditisation of capacity (Phillips & Ilcan, 2004), where agents such as small NGOs permanently need to catch up (Sumner, 2006) with large players such as the UN, INGOs,

universities, ICT vendors, or private sector consultancies (Hayes & Westrup, 2012a). Arguably, calls for a Google or Amazon of aid illustrate most clearly the permanent capacity gap and the imaginary of a techno-commercial fixing of old aid (Quaggiotto, 2007; Heeks, 2010; Beer, 2018). This thesis has sought not to dismiss the digital reformist zeal, but to bring balance to it by articulating how power remains entangled within the practices of data/knowledge construction, mediation and consumption. It has done this through interrogating the practices of impact evaluation at two small NGOs.

Where the shift to data/knowledge intensity, volume, and datamentality has supported an imaginary set of benefits (Beer, 2018: 20) this thesis has balanced the fancy towards the materiality and narrowness of data/knowledge that is currently captured and constructed by impact evaluation machines. In two small NGO case studies, impact data/knowledge landscapes were found to include: silences, such as: doubt, clamour, emotion, losses, mess, peripheries, humility, ignorance, diverse traditional and local voices and incidents, novice trials, struggles and organisational changes in efforts to master evaluation mechanics, expert instrumentalities, critical insights, sales pitches, inequalities, socio-political demands for impact products, and last but not least the critical configurations of evaluation itself, NGOs and governing processes. These silences concern individuals, small and large organisations, development 2.0 skills and abilities, attitudes and orientations, and the data/knowledge practices lost by shifting to new skills, orientations, technologies and prescriptions. Thus, the study shows that the landscape of data/knowledge assets and flows in practice, globally, is much wider than the highly articulate, professionally targeted, legible, loud and digitally documented evaluation audibles. Narrow data/knowledge is pre-scribed, expert and oriented to professional groups, policy makers, funding markets and managerial controls. The wider landscape of doubts, transformations and inequalities is comparatively silent, illegible, inarticulate, and messy.

This wider and more material view of impact data/knowledge has implications for evaluation research, practice, policy and dialogue. These implications are discussed in detail at the end of the case chapters and throughout the discussion chapter. The groups who may find value in the implications, and likely have an interest in power/data/knowledge dynamics, include:

1. development sector professionals whose work overlaps with impact or evaluation e.g. evaluators, M&E specialists, evaluation commissioners, trainers, donors, project managers, front-line staff, and aid sector technology designers

2. non-professionals such as activists, communities, leaders, and beneficiaries who wish to collaborate with, develop alternatives to, or contest evaluation processes or claims

3. academics and researchers who focus on Evaluation Studies or Impact Evaluation, particularly those who question or explore how critical perspectives are adopted, adapted, or rejected by professionals in different sectors and organisations

4. academics and researchers in Knowledge Management, KM4D, ICT4D, Information Management, Information Systems, Data Management and Data Analytics

5. researchers using CHAT and other practice-based approaches

In this final chapter, section 8.2 recaps key elements in the study, section 8.3 outlines limitations, section 8.4 suggests future research, and section 8.5 concludes the thesis.

8.2 Review of core argument

Development 2.0 has brought new opportunities for those involved with the development sector (Quaggiotto, 2007; Heeks, 2010; Green, 2015). The shift to data/knowledge intensity also poses challenges too (Thompson, 2008), not altogether dissimilar to the 1960s shift from colonialism to technical development (Kothari 2005a; Sumner, 2006). The new shift (re)produces omissions and elisions, which are related to power/data/knowledge silences considered in this study. The overarching research problem was articulated as:

- Does development impact evaluation include power relations?

Based on this problem, three specific research questions were generated and NGO impact evaluation at small NGOs selected as sites to explore the problem and possible responses. Development 2.0 provided a background to the study, impact evaluation was the specific and exemplar data/knowledge intensive process under interrogation, and small NGOs were the sites of data generation. The three research questions were:

1. How is impact evaluation data/knowledge constructed at small development NGOs, in practice?

2. Are / how are power relations generated during impact evaluation data/knowledge construction at small development NGOs?

3. Can / how can power relations be addressed and responded to in impact evaluation practices at small development NGOs?

These questions can be presented in an equation, showing the overarching problem and the three research questions (Figure 8.1).
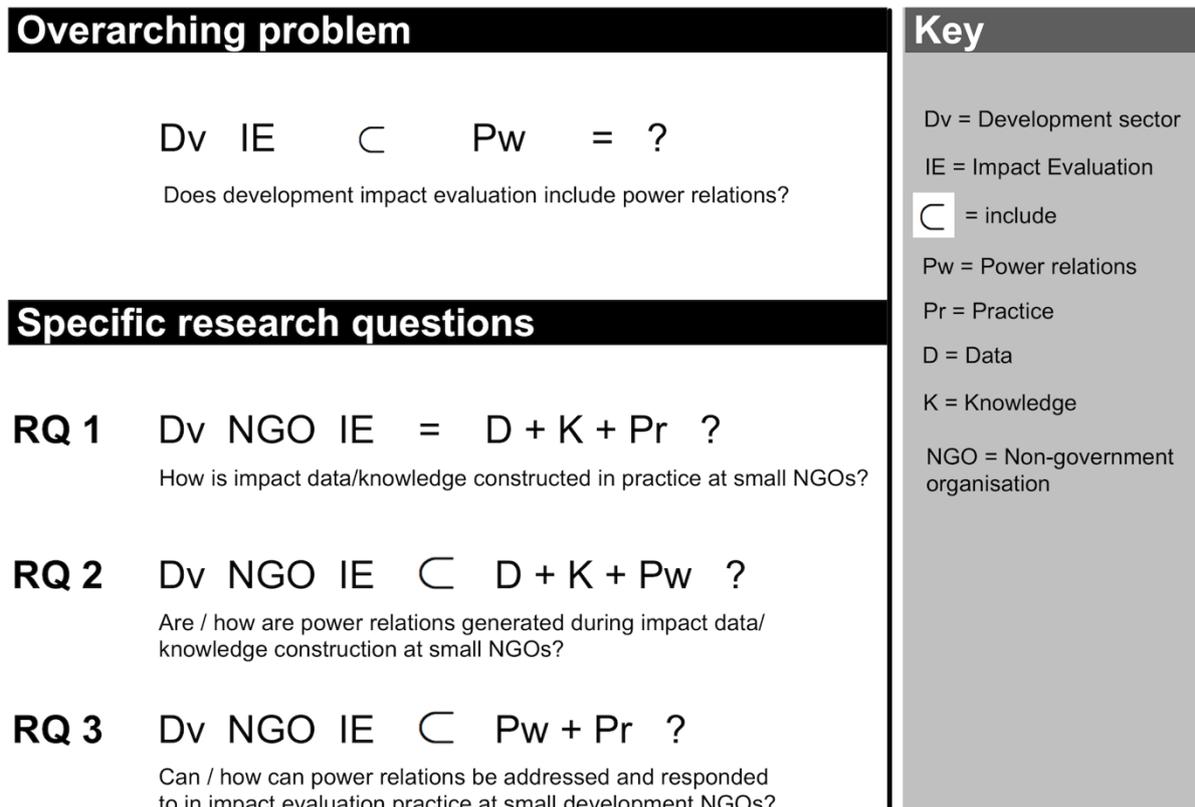
Figure 8.1: Research problem and questions represented visually as an equation

The core argument runs as follows. In the introduction chapter, impact evaluation was outlined as related to data/knowledge intensities which had arisen in the sector historically and accelerated over time. Development 2.0 incorporated accelerating production of data/knowledge in three pertinent areas: development, NGO work, and evaluation itself. These accelerations were illustrated in historical trajectories and contemporary demands.

The overarching problem and research questions concerned the need for a clearer articulation of power and practice in development 2.0 more broadly, and development NGO impact evaluation more specifically. As such, the literature interrogated technical impact evaluation knowledge (TIEK), revealing an extensive and prescriptive discourse. The historical increases

in data/knowledge intensity, sector demands and TIEK prescriptions collectively contribute to the emergence and sustaining of critical configurations, in which NGOs and impact evaluations are situated. However, power and practice remained unclear in TIEK.

A theoretical question arose concerning the unclear grounding of data and knowledge in impact evaluations. This led to a review and critique of the canonical DIKW pyramid, and recognition that this widespread model underpinned TIEK, valorising factual data for pragmatic, managerial decision-making, at the expense of engaging with power and practice. In order to articulate an alternative, another foundation was required. The alternative needed to acknowledge sensitivities to power and practice in data/knowledge work, and support engagements with NGOs making data and knowledge, rather than increase demands and prescriptions upon them. This alternative was labelled the 6Ps.

After outlining the rationale for critique plus engagement, rather than critique versus engagement, an approach using Cultural Historical Activity Theory (CHAT) was proposed (Engeström, 1987; Blackler, 1995; Karanasios & Allen, 2013). CHAT's resources for both engagement and critique were described. Given the limited time, finances and access to NGO sites, the research design focused on opportunistic, collaborative work with two NGOs. This meant research activities would prioritise advice and consultancy to secure access, but also foreground critical suggestions, collaborative learning, stimulating partner responses (Salancik, 1976), and probing (Gaver et al, 1999) evaluation activities in order to learn about practice and power relations. Contributions to partner NGOs' work on evaluations through co-designed evaluation tools, conference calls, workshops, conversations, and spreadsheet/report analysis were selected as ways of learning about evaluation practices and as sites to test critical suggestions and responses in everyday work. Two NGOs were selected from seven potential collaborations, as they illustrated expert and novice approaches to evaluation, and provided access to management participants in positions of authority, not available in other cases. The research spanned 2013 to 2016, with one NGO based in India, and one in Thailand.

In the first case, the Indian NGO and their philanthropic funder demonstrated expertise at navigating evaluation in practice. The philanthropy migrated techniques and tools from their financial sector parent organisation into their development work and partnerships. Evaluation involved a temporal chain of activities, from establishing a need-to-know impact, through assembling an impact machine, to bifurcating and transporting data, packaging and processing knowledge, and pitching convincing stories to investors at funding events. Quantitative

measures and data cells were elevated, along with digital technologies, data management, funding orientations, and organisational innovative strategies, into which impact data flowed. Despite submerging local participation, beneficiary voices, and open qualitative narratives, the case illustrated the commercial politics of evaluation processes and products. Technology and commerce shaped the evaluation chain, and shaped the TIEK or DIKW prescriptions.

In the second case, a similar chain of activities was evident, however this small NGO was a novice evaluator, struggling to learn evaluation in a precarious funding environment. Their evaluation activities incorporated diverse needs-to-know impact, struggles in assembling an impact machine, and no expertise around data bifurcation, data management, impact packaging or pitching. Nevertheless, they demonstrated attempts to share data with local partners, and use data in innovative ways such as to support advocacy work on children's rights or local NGO coalition building. A key finding in this case was the transformation of the NGO towards data/knowledge intensity. They were changing staff roles, skills and attitudes, bringing US interns in, grabbing ad hoc training opportunities, and trying to produce reports and bids which demonstrated evaluation expertise, English competency, and data/knowledge capabilities.

At both NGOs there were CHAT contradictions between impact as a documented, marketable set of convincing representations (Impact-2), and impact as a more local, experiential, contested, uncertain, negotiated, hard to document, yet more intimate aspect of people's changing lives, lifeworlds and experiences (Impact-1). The temporal chains illustrated the micro-activities of how such marketable or exchangeable impact products were constructed, and in the second case, how the small NGO was incrementally transforming from its old community engaged, project management logics (NGO-1 version), towards a data/knowledge intensive organisation (new NGO-2 version). Contradictions and power relations were generated in impact work at both NGOs, pushing them towards governmental and developmental markets and capacities, and today's development 2.0 landscape.

In the discussion, the contradictions and temporal activity chains informed a wider view of impact evaluation power/data/knowledge in the sector. This involved seeing evaluation not just as scientific truth seeking, business bottom line pragmatics, inputs and outputs, methods and results, empirical data foundations, or local participatory intentions, but through a lens of evaluation-as-practice. This lens articulated a more comprehensive, less market-oriented perspective on what happens in impact evaluation activities. Evaluation-as-practice, building on the novel CHAT temporal chains of submerging and elevating data/knowledge elements,

makes impact silences audible for researchers, evaluators, or practitioners, leading to three contributions. Firstly, the impact iceberg which contributes to the body of evaluation literature by delineating many legible/illegible splits evident in the study, what can be said or not, what is audible or silent in contemporary impact evaluation. Secondly, four audit 2.0 concepts or devices which respond to the elision or marginalisation of power and practice in audit cultures, TIEK and the DIKW legacy. These devices were: data/knowledge chains and networks; the 6P sensitivities; the impact spectrum; and datamentality/datamateriality. The third contribution was methodological and highlighted eight tactics for engaging with evaluation agents and testing critical insights with them. Together, the contributions respond to the overarching problem of data/knowledge inequalities and the marginalisation of power and practice in TIEK, and much of the development 2.0 shift. Figure 8.2 visualises the problem and contributions.
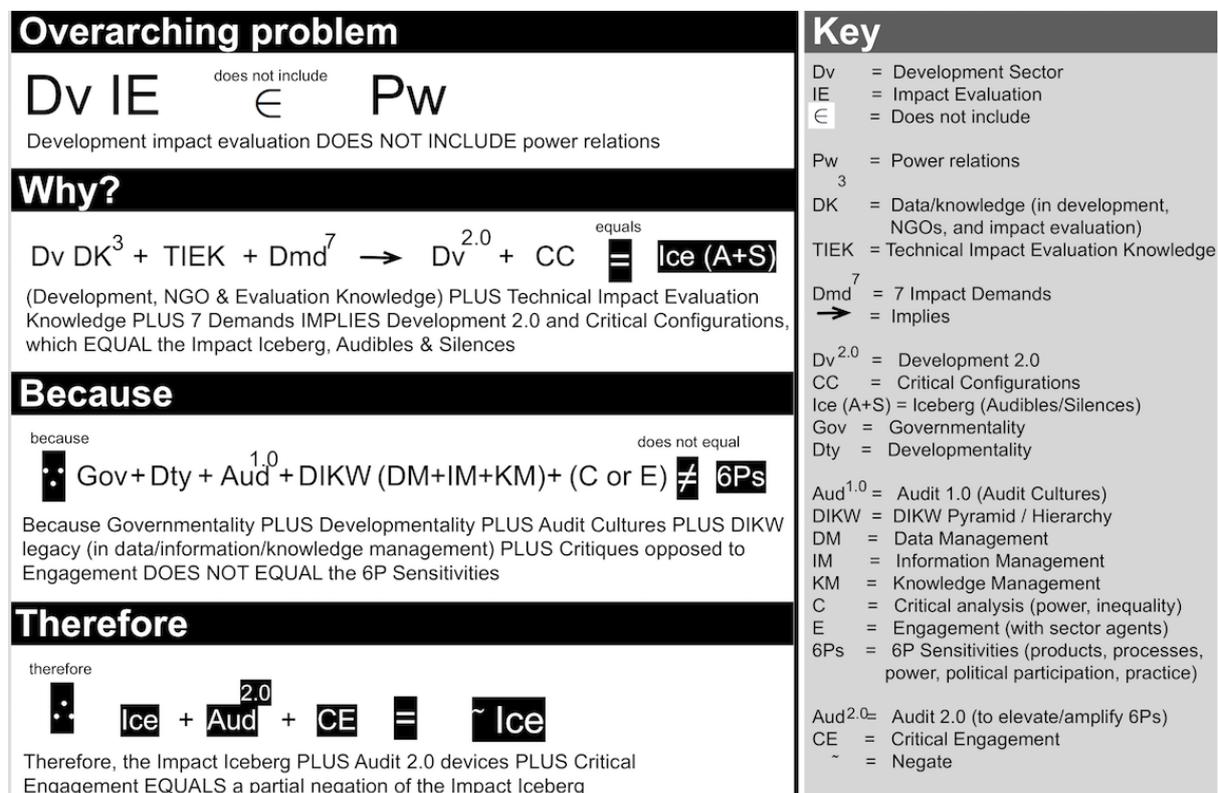
Figure 8.2: Thesis problem-response represented as an equation

In summary, the thesis contributions flesh out an evaluation-as-practice perspective. They throw into sharp relief the agencies diffused across activity chains and networks, which perform governing, developmental, and audit mechanics, and sustain the critical configurations

illustrated in the impact iceberg. Such configurations in practice position evaluation agents, and small NGOs, in seemingly permanent development 2.0 capacity markets.

8.3   Limitations

The study encountered four limitations, which encourage caution and point to future research. Firstly, and most importantly, the contributions of the evaluation-as-practice lens, impact iceberg, audit 2.0 devices and critical engagement tactics, are all limited. They are provocations for dialogue, attempts to bring critiques out of academic zones (Harris, 2016; Law, 2008) and into the wilds of practice. They are thus in need of testing and development. The contributions are not naïve, nor magic bullets to free evaluation from its structural susceptibility to capture by autocratic or bureaucratic institutions (Macdonald, 1993; Norris, 2015; Wallace et al, 2006), neoliberal markets (Picciotto, 2015: 152) or faux participation (Leal, 2007; Cook & Kothari, 2001). They are not blueprints, but stimulations (Salancik, 1979) or probes (Gaver et al, 1999) for dialogue between stakeholders occupying different spaces in evaluation networks. They fail to account for all alternative, local or traditional voices (Anderson et al, 2012; Shivji, 2006).

However, the contributions are inspired by metaphors which critique organisational apparatus to favour those who are marginalised by such apparatus. Other examples are the Ladder of Participation (Arnstein, 1969) which describes degrees of inclusion, or the Aid Chain (Wallace et al, 2006) which deconstructs aid sector dynamics. Such metaphors are limited, yet they make critical reflections accessible. The aim of evaluation-as-practice, the impact iceberg, the audit 2.0 devices, and critical engagement is similar, to amplify data/knowledge inequalities and create spaces to respond to such impact silencing[36].

Secondly, the empirical NGO cases and data generated are limited. Both cases focused on small NGOs, doing evaluation in-house, without external consultants. External consultants are

---

[36] By incorporating wide networks, power relations and silences as ingredients within the contributions, it is hoped that their criticality will be durable and accessible. Risks for such laudable aims are multiple: e.g. dialogue being only for the powerful and radical goals will be diluted by management (Burrell, 1993), the decapitation or tyrannies of participation (Leal, 2007; Cook & Kothari, 2001), or smooth developmentality (Lie, 2015 Li, 2007). Further work is required to test if and how the contributions can live, adapt or endure in the wild.

commonly deployed in aid evaluations (Wallace et al, 2006; Picciotto, 2015). Neither case included field data from target beneficiaries, farmers or villages. Although this submerges subaltern voices, and elevates critical scholarly voices, this was in line with the study aims to focus on the power and practices in development evaluation apparatus. Furthermore, the cases draw on Indian and Thai contexts, with one financially stable NGO funded by a long-term philanthropic partner and one NGO competing in a precarious funding environment, respectively. Further studies with diverse funders, methods, (e.g. ethnography, surveys), agents (beneficiaries, ICT vendors, host governments), and contexts (geographic, types of NGO) can test or add value to this work.

Thirdly, generalising from two cases to a broad sector stretches external validity. For example, in one co-authored conference workshop from the study (Kelly & Secker, 2015), a well-known scholar commented that impact assessments are different to monitoring and evaluation (M&E), and therefore the findings are limited to impact assessments only. Whilst this is true, the research demonstrated how disciplinary definitions and prescriptions do not constrain the way that back-office data/knowledge is moved, merges, shifts, flows or is re-purposed in practice. In this way, impact data can become public relations infographics, monitoring data can become success claims, and both can feed into organisational innovations, as in case one. This suggests the value of the study, as a story about evaluation, power and small NGOs, their contexts, and the impact machines they build, as phenomena wider and more entangled than TIEK boundaries, or experimental sampling validity logics. How many more rich cases would make the findings statistically significant (Crabtree et al, 2013)?

Fourthly, CHAT is not the only theoretical or methodological domain that scaffolded the study. Although CHAT's conceptual flexibility was valuable, the mixing of theory and methods may be unacceptable to some purists. The study used CHAT to respond to theoretical questions about data/knowledge omissions and methodological questions about how to critique and engage with research partners. Responses featured CHAT activity systems, contradictions and temporal chains, and connected the cases to wider debates about power/data/knowledge relations. Expansive learning (Engeström, 1987) and the CHAT hierarchy of operations, actions and activities (Leontyev, 1978; Allen et al, 2011; Karanasios, 2014: 6) were not key themes, however another school of thought was significant, i.e., governmentality (Foucault, 1979/2000).

Why this mix using two schools of thought? Firstly, expansive learning suggests formal and bounded researcher interactions and collaborative change processes. Neither was possible with either NGO. The first NGO was already expert and rejected concerns about power. The second NGO worked in ad hoc ways, learning to perform evaluation. Their journey towards evaluation mastery was messy and problematic. These observations became key case results. Secondly, Leontyev's hierarchy of activity was not significant in either case, primarily because of the focus on and salience of globally diffused and distributed evaluation chains and networks, not the unconscious or psychological levels of evaluation activity. Thirdly, as Foucauldian governmentality is well established in Critical Development Studies, it was difficult to avoid its insights. Although, one conference paper arising from the study looked at the CHAT/governmentality interplay (Kelly & Cowen, 2014), and one journal paper focused on governmentality specifically (Hayes et al, 2017), a discussion of the CHAT/governmentality tensions and synergies requires an in-depth study of its own.

8.4   Further research

Aspects of the data/knowledge critiques in this study inform future work and audiences. Audiences will likely be those mentioned in section 8.1 who share interests in the practices of and power dynamics endemic to mundane data/knowledge intensive work, particularly evaluation. There are six areas for future research implicated by the study.

Firstly, further work with small NGOs and groups with limited resources in the development sector is important to understand if and how they develop expertise in data/knowledge intensity. These could be citizen groups, advocacy groups or service NGOs who have expertise or potential to carve out data/knowledge niches for their own sustainability. A key question would be: Do NGOs intensify market orientation as they gain expertise? [37]

Secondly, another focus would be on the impact iceberg, audit 2.0 devices, and how critical engagement can be tested and extended. How can organisations be encouraged to adopt or adapt critical insights, and how can academic research be best oriented towards socially

---

[37] In case two, it was unclear whether HTSG would develop in future to focus on a) instrumental market exchanges of impact products for financial resources, b) innovative knowledge sharing arrangements, or c) both.

responsive results and incentives beyond journal markets for example (Harris, 2016: 180). Can boundary crossing help academics and practitioners co-design outcomes that mitigate power/data/knowledge inequalities in aid and journal markets? What other concepts, audit 2.0 tools, wider infrastructures (Beer, 2016) or artefact biographies (Williams et al, 2012) can help? Can humility be resuscitated, how? Do we need critical infographics? How do academic practitioners and aid practitioners critically engage? How can the impact iceberg, audit 2.0 contributions or critical brokering be matured and made more appealing for use in the wild?

Thirdly, is the CHAT analysis of temporally diffused power dynamics in funder-NGO partnerships unique, or does it resonate with broader development 2.0 relations? For example, does ICT2.0 deliver progress that is increasingly responsive to poor people's demands, as Heeks (2008: 33) argues? And are data-intensive technologies on balance both relevant and beneficial to developing nations (Walsham & Sahay, 2006: 7)? Answering such questions implicates development 2.0 relations between kinds of organisations (e.g. investors, governments, IT vendors, consultancies), and kinds of processes (e.g. planning, e-development, data analysis, and technology-centric innovations such as blockchain for development) (Hernandez, 2017). More research on development 2.0 is required to mature these debates. It will need to draw on broad forms of analysis (Brigham & Hayes, 2013: 127), illuminate the diversity of development 2.0 organisations and processes, and elevate power/data/knowledge relations. CHAT is well suited to such challenges.

Fourthly, the opportunity for further CHAT use in the development sector or other sectors concerning power/data/knowledge intensities, exchanges, circulations, markets, chains or networks. Specific areas would include further application of the temporal activity chains concept and the data/knowledge audit 2.0 networks; how expansive learning can be ad hoc, partial, or unguided; how Leontyev's activity hierarchy (1978) could be extended or adapted to analyse more globally diffused relations; and how other critical schools of thought, specifically governmentality from a Foucauldian perspective, can complement CHAT to bring critiques and local agency/learning into productive dialogues. Similarly, research on how activity systems, chains and networks are manifest in the politics of knowledge (Lewis & Gardner, 2015), the politics of evaluation (Guijt, 2015), and the battlefields of development (Long & Long, 1992) would be informative.

Fifthly, how can approaches to social practice and power be made pragmatic, palatable, and useful for adoption and adaptation by development agents? If sector demands, governing or

market logics, TIEK or the widespread DIKW legacy continue to normalise and silence power and practice, then how can more sensitive approaches be re-embedded into evaluation and development 2.0 more broadly? From development, Gardner & Lewis (2015: 180–181) implore anthropologists and aid agencies to collectively engage and critique. Guijt (2015: 207) asks practitioners to acknowledge the messy politics of their work or find "*space and time for more appropriate protocols and methods.*" Nicolini (2012: 240–241), in reviewing practice theories, concludes that testing practice-based toolkits in fieldwork is required. Engeström, Virkkunen, Helle, Pihlaja, & Poikela (1996) and others (e.g., Virkkunen & Newnham, 2013: 24–25) facilitate CHAT change laboratories, bringing stakeholders together to explore transformations. However, in global development, stakeholders are geographically dispersed across data/knowledge supply chains. Participatory forms of ICT4D, knowledge management, and evaluation emphasise local communities and local knowledges as a response. But being sensitive to data, knowledge, and power means acknowledging wider activity chains across development 2.0. As such, the study has no clear answers but suggests multisite responses distributed across aid chains may better amplify the problem of power inequalities than single-site interventions. Approaches to social practice can help articulate these multisite issues, but more critical engagement, critical brokering, pragmatic tools and translatable concepts will be required.

Sixthly, there are several concepts or movements in practitioner-oriented research in the development sector that share commitments with critical engagement. Each of these movements encourages local participation, and in that sense at least, can elevate multi-voiced alternatives to top-down governing, managerialism and audit cultures. These include:

- Evaluation forms that trace networks e.g. Outcome Mapping (Earl et al, 2001)
- Knowledge brokering (Shucksmith, 2016; Lewis & Mosse, 2006)
- Knowledge for Development (K4D) which is an emerging field comprising conferences (e.g. The World Bank, 2017), communities (e.g. at NING http://www.km4dev.org/ and D-Groups https://dgroups.org/groups/pelican), institutional research platforms (e.g. Institute of Development Studies, 2019), journals (e.g. KM4D, 2019), and agendas attached to the UN SDGs (Brandner and Cummings, 2017)
- Development 2.0 (Thompson, 2008; Heeks, 2010; Quaggiotto, 2009); Information and Communications Technologies for Development (ICT4D) (Heeks, 2008; Heeks, 2018)

- New approaches in development e.g. Problem Driven Iterative Adaptation (Andrews et al, 2013); Doing Development Differently (DDD) (Andrews et al, "DDD Manifesto", 2014; Prieto-Martin et al, 2017); Thinking and Working Politically (TWP) (Leftwich, 2011; Booth, 2015); Positive Deviance (Pascale et al, 2010), Complexity Theory and Systems Thinking (e.g. Honig & Gulrajani, 2018; Ramalingam, 2013)
- Design for Development (DfD) (Donaldson, 2009), Human Centred Design (e.g. IDEO Toolkit, 2009), Service Design for Development (Edmonds & Cook, 2014)

These trends foreground local participation, stakeholder engagement, multi-voiced facilitation, complexity, and/or digital participation in new data/knowledge intensive processes. However, these trends do not explicitly interrogate diffused power and politics, or conceptualise governing practices, spread across aid chains, markets and networks. They demonstrate a shift towards people centric development, but do not conceptualise the impacts, inequalities, or messy ontological politics (Law, 2004) which result from critical data/knowledge configurations in aid apparatus. Without attention to such problems, brokering work around these trends might unfortunately exacerbate and widen the data/knowledge capacity gap, producing new market opportunities, new developmentalities (Lie, 2015; Phillips & Ilcan, 2004; 2010; Hayes et al, 2017) or datamentalities. These elevated opportunities for new experts risk submerging inarticulate voices, doubt, clamour, political contestations, or iceberg silences (Areas E & F).

As a caveat, the three contributions and audit devices are not naive. Previous decades have witnessed similar attempts to draw attention to power and politics[38]. With each attempt, governmentality and aid markets (Ferguson, 1994; Li, 2007; Bächtold, 2014; Hayes et al, 2017), managerialism (Dar & Cook, 2008), and developmentality (Ilcan and Phillips, 2010; Lie, 2015) critiques apply. This is because the movements seek to empower the local through expertise, but gain less traction on contesting the silent power/data/knowledge relations embedded in aid apparatus ways of working, evaluating, managing and circulating data/knowledge, from a distance. New movements seek to rebalance control and do things differently; however, they can be susceptible to the silent, mundane shaping of scientific and participatory approaches by distal governing practices (Hayes & Westrup, 2012b; Green,

---

[38] Previous efforts are many e.g. Community Development (Leal, 2007), Participatory Rural Appraisal (Chambers, 1994); more recently Localisation (Bonacker et al, 2016) and Doing Development Differently (Crewe, 2014).

2009), often because they lack a wide network ontology of activities, power, data and knowledge. Approaches to social and organisational practice (e.g. Karanasios, 2014; Nicolini, 2013; Miettinen et al, 2009, Reckwitz, 2002; Schatzki, 1996; Marcus, 1995; Harper, 2003) have much to offer these trends, in widening this network of practices and governing agencies, but the tools of practice theory need to be brokered and translated to gain traction. Approaches to social practice can widen the "*data gaze*" (Beer, 2018) and knowledge gaze, elevate datamateriality, explore networks of influence and offer conceptual resources. Critical brokering is required to build momentum between scholars, aid practitioners and communities, to jointly contest the "*increasingly global material forces*" identified by Mohan (2001).[39]

Finally, important work using lessons and contributions from the study has already been initiated by the author, in collaboration with various organisations and colleagues.[40]

8.5   The impact of evaluation

This study has explored the impact of evaluation, in contrast to TIEK literature which focuses on ever better ways of evaluating. Many such normative prescriptions are founded on widespread assumptions and omissions in data/knowledge work more broadly, particularly ones that have accelerated since the 1990s (Lambe, 2011). This study has contested some of these foundations, particularly two described below.

Firstly, the claim that development has become a data/knowledge intensive industry, although true in one respect, often elides the observation that such inert, intangible assets as theory, ideas, policies, evidence, data cells, and knowledge all have physical, tangible foundations. These are manifested in distributed practices, activities, technologies, people, databases,

---

[39] There are at times heated debates between aid practitioner and academic voices (e.g. Green, 2018; Banks et al, 2015) defending institutional norms and practices. Such disputes can damage the legitimacy of collaborations.

[40] Ongoing related work includes: a review of CHAT and power relations in Information Systems (Simeonova et al, 2018a; 2018b); further collaboration with HTSG Thailand; testing critical brokering in aid governance as part of a knowledge platform (e.g. Kelly et al, 2018); exploring evaluation power dynamics through interviews with professional evaluators conducted between 2013-2017; understanding exploitation in developing nations in global pharmaceutical trial networks, with a UK based international law firm; and a co-authored paper using Scott's metis/schemata opposition to understand competing forms of expertise in Tasmanian public health services.

reports, training manuals, email systems, data cells, and broad social-economic infrastructures such as universities, industry regulators, financial markets and governing mechanisms (Sassen, 2002; Beer, 2018). The social and material basis of datamentality is tangible and unequal, not inert, nor apolitical.

Secondly, the notion that today in development we have more data/knowledge than ever, wider coverage and sub-specialisms, and better value from our development 2.0 capabilities is also challenged. We cannot capture everything. As the impact iceberg illustrates, much is rendered silent or marginal, lost or illegible. When we capture one thing, we elide another. Instead of widening, we are at risk of narrowing vast data/knowledge streams towards centres of unequal and ineffective governing, markets and auditing. Acknowledging that data/knowledge is tangible and diffused, that cells can be myopic, is important in understanding evaluation-as-practice and power, and in seeing the impacts of our evaluations. If we wish to see the power relations generated during scientific, participatory and technological evaluation processes, which flow within development 2.0 data/knowledge intensive capacity markets, we must acknowledge the organisational and political positioning of evaluation concepts, methods, contests, practices, pitches, and agents, within these activities and networks.

Stakeholders contest, imagine and desire different data products. One group may admonish others for their data desires, e.g.:

> *"Donors increasingly want to see more impact for their money, practitioners are searching for ways to make their projects more effective, and politicians want more financial accountability behind aid budgets." (Alkire et al, 2017, in The Guardian)*

Alkire et al, predominantly economists from European and US universities, desire macro-level data to solve poverty. Yet such desires compete and have impacts. Innovation for one group, creates a lack of capacity for another. Humble engagement for one group, represents inarticulate clamour for another. Efficiency in one practice, is a lack of participation in another. The desire for digital data involves its speed, accessibility, revealing nature, panoramic coverage, prophetic promise, and smart intelligence (Beer; 2018: 20; O'Neil, 2016; Boisot, 2004: 10). But acknowledging the limited, more tangible, materiality and narrowness of much impact data/knowledge, is part of facilitating more comprehensive discussions about who owns cells, representations and experiences, where they circulate, how certain they are, what

methods we use to know them, and the wider politics of knowledge (Gardner & Lewis, 2015), evaluation (Roche & Kelly, 2012; Picciotto, 2012), and methods themselves (Law, 2004).

The argument, from a critical engagement perspective, is that we require a broader understanding of power and practice in development evaluation (Mosse, 2008: 123; Eyben, 2009), particularly in how power is generated within mundane scientific, technological, business-oriented, or participatory data/knowledge practices. Power is not necessarily loud, direct, aggressive or coercive (Blackler, 2011; Foucault, 1979/2000). It sits there, in the data and knowledge and our passing it around.

Research and critiques have roles to play here, but only if they step out of the box, to test critical insights in the wilds of evaluation practices. The trickle-down notion of academic insight influencing practice from a journal article, in a linear fashion, is not justified anymore (Harris, 2016; Parkhurst, 2-17; Mosse, 2004b). What is required are more mobile and accessible accounts of power and practice that can be used in tests, provocations, reflection, adoption or adaptation activities in fields and offices. Critical engagement and critical brokering have roles to play in understanding and encouraging this kind of commitment. If critiques live only in professorial debates (Law, 2008: 150) or journals, they are like aid impacts, at risk of commodified myopia, trapped in their own institutional regimes of truth and power (Foucault, 1980; Introna, 2003). Such critiques cannot get out, cannot encounter other local knowledges (Willcocks, 2004: 244). The liveliest projects are when critics break with their conservative practices in the academy and get out and engage with other power brokers, institutions and power/knowledge regimes (Marcus, 2011: 204-6).

To conclude, the development 2.0 capacity gap is perhaps the most significant inequality encountered in the study. It is a market opportunity worth sustaining for many impact network agents. Gaps will undoubtedly widen in future, around technical evaluation, development sub-disciplines, datamentality, and new digital innovations. Despite the promise and practice of development 2.0, and impact evaluation itself, both are subject to shaping pressures from audit cultures, governmentality, developmentality, and commercial pragmatics. The danger is that the sector moves further towards a future of datamentality, evaluative thinking, systems thinking, performance metrics, risk mitigation, impact indicators, big data, e-development systems, data analytics, poverty verticals, and digital expertise without acknowledging the impact silences, wide activity networks, the 6P sensitivities, or the concerns (not all clamour)

of targeted communities. Critical engagement is required to make impact evaluation more sensitive and politically transparent in the development 2.0 landscape.

There are many areas of work which lack mobile and accessible critical tools and ideas with which to respond to power/data/knowledge omissions, assumptions, inequalities, inefficiencies and silences. This thesis has responded to one area in particular, that of development impact evaluation. More critically engaged tools and infrastructures are needed, and the study's three contributions support the case for mobilising critiques, making them more accessible or even desirable.

Despite the limitations of this study, the need for further research and for testing the contributions in further evaluations, data/knowledge brokering and critical engagements, the study has interrogated the impact of evaluation and the narrowing of data/knowledge in the digital age.

The contributions and lessons learned from the cases support more ways of taking theory out of labs and libraries and into the wilds of evaluation practice, more acceptance of humility, politics and "*clamour*" as other ways of knowing, more spaces for dialogue, critical reflection, multi-voiced exchanges, and more anxiety around the silences wrapped up in digitally circulated cells, imaginary plans, knowledge products, sales pitches, and solid certainties. These all relate to the impacts of data/knowledge intensive evaluation, and the power relations such impact evaluations generate.

9 References

3ie. (2012). *Impact evaluation practice: a guide for grantees*. International Initiative for Impact Evaluation. New Delhi, India.

3ie. (2012). *Impact evaluation glossary*. International Initiative for Impact Evaluation: New Delhi, India.

3ie. (2016) *About.* International Initiative for Impact Evaluation. Retrieved from http://www.3ieimpact.org/en/about

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, *16*, 3–9.

ActionAid. (2012). *Critical Stories of Change: Let's do it together for our village. Pu pang hsaung yet doh year a twet a si a sin*. The Fellowship programme in Myanmar. Retrieved from http://actionaid.org/sites/files/actionaid/critical_story_for_change.pdf

Acuna, A. (2012). *DIKW Model. User comment in the Knowledge Management for Development (KM4Dev) community wiki page discussion record*. Retrieved from http://wiki.km4dev.org/DIKW_model

Adriaans, P. and van Benthem, J. (2008a). Introduction: Information Is What Information Does. In P. Adriaans & J. van Benthem (Eds), *Handbook of the Philosophy of Science*. Amsterdam: Elsevier.

Adriaans, P. and van Benthem, J. (Eds) (2008b). Philosophy of Information. In *Handbook of the Philosophy of Science*. Amsterdam: Elsevier.

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly 25*(1), 107-136.

Alkire, S., Bédécarrats, F., Deaton, A., Giraud, G., Guérin, I., Harriss-White, B., Heckman, J.,… & Stiglitz, J. (2018, August 16). Buzzwords and tortuous impact studies won't fix a broken aid system. *The Guardian*. Retrieved from https://www.theguardian.com/global-development/2018/jul/16/buzzwords-crazes-broken-aid-system-poverty

Allen, D. K., Brown, A., Karanasios, S., & Norman, A. (2013). How should technology-mediated organizational change be explained? A comparison of the contributions of critical realism and activity theory. *MIS Quarterly, 37*(3), 835-854. DOI: 10.25300/MISQ/2013/37.3.08

Allen, D. K., Karanasios, S., & Norman, A. (2014). Information sharing and interoperability: the case of major incident management. *European Journal of Information Systems*. *23* (4), 418–432. https://doi.org/10.1057/ejis.2013.8

Allen, D., Karanasios, S., & Slavova, M. (2011). Working with activity theory: Context, technology, and information behavior. *Journal of the American Society for Information Science and Technology*, *62*(4), 776–788. https://doi.org/10.1002/asi.21441

Alvesson, M., & Willmott, H. (2012). *Making sense of management: A critical introduction*. London: Sage.

Anderson, A. A. (2006). *The community builder's approach to theory of Change: A practical guide to theory development*. Retrieved from https://assets.aspeninstitute.org/content/uploads/files/content/docs/rcc/rcccommbuildersapproach.pdf

Anderson, M. B., Brown, D., & Jean, I. (2012). *Time to listen: hearing people on the receiving end of international aid*. Retrieved from http://live-cdacollaborative.pantheonsite.io/wp-content/uploads/2016/01/Time-to-Listen-Hearing-People-on-the-Receiving-End-of-International-Aid.pdf

Andrews, M. (2014). *The DDD Manifesto*. Retrieved from https://buildingstatecapability.com/the-ddd-manifesto

Andrews, M., Pritchett, L., & Woolcock, M. (2013). Escaping capability traps through problem driven iterative adaptation (PDIA). *World Development*, *51*, 234-244.

Anthony, P. (1998). Management education: Ethics versus morality. In M. Parker (Ed.), *Ethics and Organization* (pp. 269-81). London: Sage.

Argyris, C. (1985). Action science, concepts methods and skills for research and intervention. San Francisco, CA: Jossey-Bass.

Arnstein, S. R. (1969). A Ladder of Citizen Participation. *Journal of the American Institute of Planners*, *35*(4), 216-224. DOI: 10.1080/01944366908977225

Arrow, K.J. (1969). Classificatory notes on the production and transmission of technological knowledge. *The American Economic Review*, *59*(2), 29-35.

Attewell, P. (1991). Big brother and the sweatshop: computer surveillance in the automated office. In C. Dunlop and R. Kling (Eds.), *Computerization and Controversy: Value Conflicts and Social Choices* (insert page numbers). Boston: Academic Press.

Avgerou, C. (2002). *Information systems and global diversity*. Oxford: Oxford University Press.

Avgerou, C. (2010) Discourses on ICT and development. *Information Technologies & International Development*, *6*(3), 1-18.

Avis, J (2007). Engeström's version of activity theory: a conservative praxis? *Journal of Education and Work*, *20*(3), 161-177.

Avis, J. (2009). Transformation or transformism: Engeström's version of activity theory? *Educational Review*, *61*(2), 151-165.

Bächtold, S. (2015). The rise of an anti-politics machinery: peace, civil society and the focus on results in Myanmar. *Third World Quarterly*, 36(10), 1968-1983.

Bamberger, M., Rao, V., & Woolcock, M. (2010). Using mixed methods in monitoring and evaluation: experiences from international development. World Bank Policy Research Working Paper 5245. Washington DC: The World Bank.

Banks, N., & Hulme, D. (2012). *The Role of NGOs and Civil Society in Development and Poverty Research Network*. Retrieved from http://papers.ssrn.com/abstract=2072157

Banks, N., Hulme, D., & Edwards, M. (2015). NGOs, states, and donors revisited: Still too close for comfort? *World Development*, *66*, 707–718. https://doi.org/10.1016/j.worlddev.2014.09.028

Barr, A., Fafchamps, M., & Owens, T. (2005). The governance of non-governmental organizations in Uganda. *World Development*, *33*(4), 657-679.

Barr, J., Rinnert, D., Lloyd, R., Dunne, D., & Hentinnen, A. (2016). The Value of Evaluation: Tools for Budgeting and Valuing Evaluations. Discussion Paper. ITAD.

Barrow, C. J. (2000). *Social impact assessment: an introduction*. Oxford: Oxford University Press.

BBC News (2015, Sept 8). Migrant crisis: UK aid budget will help fund refugees response – Osborne. *British Broadcasting Corporation.* Retrieved from http://www.bbc.co.uk/news/uk-34167271

Beaulieu, D. (2013). Diversity and tension in knowledge production and dissemination: a closer look at the activities of 10 Ghanaian civil society organisations. *Knowledge Management for Development Journal*, *9*(2), 37-52.

Bebbington, A. J. (2005). Donor-NGO relations and representations of livelihood in nongovernmental aid chains. *World Development*, 33(6), 937-950.

Bebbington, A. J., Hickey, S., & Mitlin, D. (2008). Introduction: can NGOs make a difference? The challenge of development alternatives. In A. J. Bebbington, S. Hickey, & D. Mitlin, (Eds.), *Can NGOs Make a Difference? The Challenge of Development Alternatives* (pp.3-37). London: Zed Books.

Bebbington, A. J. (1997). New states, new NGOs? Crises and transitions among rural development NGOs in the Andean Region. *World Development*, *25*(11), 1755-1765.

Bebbington, A. J. (2004). NGOs and uneven development: geographies of development intervention. *Progress in Human Geography*, *28*(6), 725-745.

Beer, D. (2018). The Data Gaze: Capitalism, Power and Perception. London: Sage.

Beeson, M., & Islam, I. (2005). Neo-liberalism and East Asia: resisting the Washington consensus. *The Journal of Development Studies*, *41*(2), 197-219.

Bernardi, R., & De Chiara, F. (2011). ICTs and monitoring of MDGs: A case study of Kenya HIV/AIDS monitoring and evaluation in a donor multi-agency context. *Information Technology for Development*, *17*, 24–41. https://doi.org/10.1080/ 02681102.2010.511699

Better Evaluation (2016). *What is an approach?* Retrieved from http://betterevaluation.org/approaches

Better Evaluation (2017) *Approaches page*. https://www.betterevaluation.org/en/approaches .

Bhatt, G. (2000). Organizing knowledge in the knowledge development cycle. *Journal of knowledge management*, *4*(1), 15-26.

Bhatt, G. (2001). Knowledge management in organizations: Examining the interactions between technologies, techniques, and people. *Journal of knowledge management*, *5*(1), 68-75.

Blackler, F. (1993). Knowledge and the theory of organizations: Organizations as activity systems and the reframing of management. *Journal of management studies*, *30*(6), 863-884.

Blackler, F. (1995). Knowledge, knowledge work and organizations: An overview and interpretation. *Organization Studies*, *16*(6), 1021–1046. https://doi.org/10.1177/017084069501600605

Blackler, F. (2009). Cultural-Historical Activity Theory and Organization Studies. In A. Sannino, H. Daniels & K.D. Gutiérrez (Eds.), *Learning and Expanding with Activity Theory* (pp.19-39). Cambridge: Cambridge University Press.

Blackler, F. (2011). Power, politics, and intervention theory: Lessons from organization studies. *Theory and Psychology*, 21, 724–734. https://doi.org/10.1177/0959354311418146

Blackler, F., & Regan, S. (2009). Intentionality, agency, change: Practice theory and management. *Management Learning*, *40*, 161–176. https://doi.org/10.1177/1350507608101227

Blackler, F., Crump, N., & McDonald, S. (2000). Organizing processes in complex activity networks. *Organization*, *7*(2), 277-300.

Boell, S. K., & Cecez-Kecmanovic, D. (2015). What is 'Information' Beyond a Definition? *Thirty Sixth International Conference on Information Systems*, Fort Worth, US. Retrieved from https://www.researchgate.net/publication/285581995_What_is_%27Information%27_Beyond _a_Definition

Boisot, M. (2005). Exploring the information space: a strategic perspective on information systems. In D. Rooney, G. Hearn, & A. Ninan (Eds.), *Handbook on the Knowledge Economy* (pp.179-90). Cheltenham, UK: Edward Elgar Publishing.

Boisot, M. H. (1998). Knowledge assets: Securing competitive advantage in the information economy. Oxford: Oxford University Press.

Boland Jr, R. J. (1987). The in-formation of information systems. In R. J. Boland Jr, & R. A. Hirscheim, (Eds.), *Critical Issues in Information Systems Research* (pp.363-394). New York: John Wiley & Sons.

Boland, Jr, R J. & Tenkasi R.V. (1995). Perspective Making and Perspective Taking in Communities of Knowing. *Organisation Science*, *6*(4), 350-372.

Bonacker, T., von Heusinger, J., & Zimmer, K. (Eds.). (2016). *Localization in Development Aid: How Global Institutions Enter Local Lifeworlds*. London: Routledge.

Booth, D. (2015). *Thinking and Working Politically*. GSDRC Professional Development Reading Pack No. 13. Governance and Social Development Resource Centre. Retrieved from https://assets.publishing.service.gov.uk/media/57a0899aed915d622c0002df/twp_rp.pdf

Brandi, U., & Elkjaer, B. (2011). Organizational learning viewed from a social learning perspective. In M. Easterby-Smith & M. Lyles (Eds.), *Handbook of organizational learning and knowledge management* (pp. 21-41). John Wiley & Sons.

Brandner, A., & Cummings, S. (2017). *Agenda Knowledge for Development: Strengthening Agenda 2030 and the Sustainable Development Goals*. Knowledge for Development Partnership, Vienna, Austria. Retrieved from http://www.km-a.net/kma/wp-content/uploads/Agenda-Knowledge-for-Development-and-Statements.pdf

Brigham, M., & Hayes, N. (2013). Hybridity, consulting and e-development in the making: inscribing new practices of impact assessment and value management. *Information Technology for Development*, *19*, 112–132. https://doi.org/10.1080/ 02681102.2012.690171

Brown, V. A. (2010). Multiple knowledges, multiple languages: are the limits of my language the limits of my world? *Knowledge Management for Development Journal*, *6*(2), 120-131.

Bugnion, F. (2009). *From Solferino to the birth of contemporary international humanitarian law*. International Committee of the Red Cross (ICRC). Retrieved from https://www.icrc.org/eng/resources/documents/article/other/solferino-article-bugnion-240409.htm

Burger, R., & Owens, T. (2010). Promoting transparency in the NGO sector: examining the availability and reliability of self-reported data. *World Development*, 38(9), 1263-1277.

Burrell, G. (1993). Eco and the Bunnymen. In J. Hassard & M. Parker (Eds), *Postmodernism and Organizations* (pp.71-82). London: Sage.

Bustelo M., Espinosa, J., Faúndez, A., & Weinstein, M. (2015). *Guide to including a gender+ perspective in VOPEs: innovating to improve institutional capacities*. Latin America and Caribbean Evaluation Network (RELAC), European Evaluation Society and International Organization for Cooperation in Evaluation (IOCE) Working Group. Retrieved from http://www.ioce.net/en/PDFs/GenderPlus%20Perspective%20Guide.pdf

Cameron, J. (2006) A Participatory Approach to Evaluating the Impact of NGOs on Development in Nepal. *Development in Practice*, *16*(1), 91-96

Campbell, D.T. (1974). Qualitative knowing in action research. Kurt Lewin Award address, Society for the Psychological Study of Social Issues, presented at the meeting of the American Psychological Association, New Orleans, LA, September.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago, IL: Rand-McNally.

Canary, H. E., & McPhee, R. D. (2009). The mediation of policy knowledge: An interpretive analysis of intersecting activity systems. *Management Communication Quarterly*, *23*, 147–187. https://doi.org/10.1177/0893318909341409

Cardoso, F. H., & Faletto, E. (1979). Dependency and development in Latin America (Dependencia y desarrollo en América Latina). Berkeley: University of California Press.

CARE (2016). *CARE's History*. Retrieved from http://www.care-international.org/who-we-are-1/cares-history

Cashmore, M., Bond, A., & Cobb, D. (2010). Evaluating the Effectiveness of Impact Assessment Instruments: Theorising the Nature and Implications of Their Political Constitution. *Environmental Impact Assessment Review*, *30*(6), 371–379.

Castells, M. (1997). The Power of Identity: The information age: Economy, society and culture (Vol. 2). Malden, MA and Oxford: Blackwell Publishers.

Castells, M. (2011). The rise of the network society: The information age: Economy, society, and culture (Vol. 1). Chichester: John Wiley & Sons.

Center for Global Development (2006). *When will we ever learn: Improving lives through impact evaluation*. Report of the evaluation gap working group, May 2006. Washington D.C., USA.

Chambers R. (1993). *Challenging the Professions*. London: Intermediate Technology Publications.

Chambers, R. (1994). The origins and practice of participatory rural appraisal. *World Development*, *22*(7), 953-969.

Chambers R. (1997). *Whose Reality Counts? Putting the First Last*. London: IT Publications.

Choo, C. (1998). *The knowing organization*. New York: Oxford University Press.

Ciborra, C. (2006). Imbrication of Representations: Risk and Digital Technologies. *Journal of Management Studies*, *43*(6), 1339-1356.

Clegg, S. R. (1989). *Frameworks of Power*. London: Sage.

Clegg, S. R., Courpasson, D., & Phillips, N. (2006). *Power and Organizations*. Newbury Park, CA: Pine Forge Press.

Cleveland, W. S. (2001). Data science: an action plan for expanding the technical areas of the field of statistics. *International Statistical Review*, 69(1), 21-26.

Clot, Y. (2009). Clinic of activity: the dialogue as instrument. In A. Sannino, H. Daniels & K. Gutierrez (Eds.), *Learning and Expanding with Activity Theory* (pp. 286-303). Cambridge: Cambridge University Press.

Clot, Y., & Kostulski, K. (2011). Intervening for transforming: The horizon of action in the Clinic of Activity. *Theory & Psychology*, *21*(5), 681-696.

Coffman, J. (2002). A conversation with Michael Quinn Patton. *The Evaluation Exchange Newsletter*, 8(1). Retrieved from http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/family-support/a-conversation-with-michael-quinn-patton

Collier, J. (1945). United States Indian Administration as a laboratory of ethnic relations. *Research*, *12*, 265-303.

Contu, A., & Girei, E. (2014). NGOs management and the value of 'partnerships' for equality in international development: What's in a name? *Human Relations*, *67*(2), 205-232.

Contu, A., & Willmott, H. (2003). Re-embedding situatedness: The importance of power relations in learning theory. *Organization science*, *14*(3), 283-296.

Cooke, B. (2003). A new continuity with colonial administration: participation in development management. *Third World Quarterly*, *24*(1): 47-61

Cooke, B., & Kothari, U. (2001). The Case for Participation as Tyranny. In B. Cooke & U. Kothari (Eds), *Participation: The New Tyranny* (pp1-15). London: Zed Books.

Cooke, B., & Kothari, U. (Eds.). (2001). *Participation: The new tyranny?* London: Zed Books.

Cornwall, A. (2005). Critical stories of change: Love of the heart, Tales from Raizas Vivas Brazil. Johannesburg: ActionAid International.

Crabtree, A., Hemmings, T., Rodden, T., Cheverst, K., Clarke, K., Dewsbury, G., & Rouncefield, M. (2003). Designing with care: Adapting cultural probes to inform design in sensitive settings. In *Proceedings of the 2004 Australasian Conference on Computer-Human Interaction (OZCHI2004),* November 2004, Brisbane, (pp.4-13).

Crabtree, A., Tolmie, P., & Rouncefield, M. (2013). "How Many Bloody Examples Do You Want?" Fieldwork and Generalisation. In *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work*, 21-25 September 2013, Paphos, Cyprus (pp. 1-20). London: Springer.

Craig, D., & Porter, D. (2006). Development beyond neoliberalism? Governance, poverty reduction, and political economy. London: Routledge.

Dalkir, K. (2011). *Knowledge management in theory and practice*. Cambridge, MA: MIT Press.

Daniels, H., Edwards, A., Engeström, Y., Gallagher, T., & Ludvigsen, S. R. (Eds.). (2013). *Activity theory in practice: Promoting learning across boundaries and agencies*. London: Routledge.

Danowitz, A. K., Nassef, Y., & Goodman, S. E. (1995). Cyberspace across the Sahara: computing in North Africa. *Communications of the ACM*, *38*(12), 23-29.

Dar, S., & Cooke, B. (2008). The New Development Management: Critiquing the Dual Modernization. London: Zed Books.

Davenport, T. H., & Prusak, L. (1998). *Working Knowledge: How Organizations Manage What They Know*. Boston MA. Harvard Business Press.

Davidson, J. (2014). Anecdata. Genuine Evaluation. Retrieved from http://genuineevaluation.com/anecdata

Davies, R. (1997). Donor information demands and NGO institutional development. *Journal of International Development*, *9*(4), 613-620.

Davies, R., & Dart, J. (2005). *The 'most significant change'(MSC) technique. A guide to its use*. Retrieved from http://betterevaluation.org/resources/guides/most_significant_change.

De Haan, A. (2009). How the aid industry works: an introduction to international development. Sterling VA: Kumarian Press.

De Laet, M., & Mol, A. (2000). The Zimbabwe bush pump: Mechanics of a fluid technology. *Social Studies of Science*, *30*(2), 225-263.

Dean, M. (2010). Governmentality: Power and rule in modern society. London: Sage Publications.

Deepak, S., Kumar, J., Sivarama, C., Giriyappa, R. & Ramasamyn, P. (2016). Life stories of persons with intellectual disabilities in Mandya District, India. *Knowledge Management for Development Journal*, *11*(1), 8-24.

Demos (2003). *Inside out: rethinking inclusive communities*. DEMOS Report, Barrow Cadbury Trust, Birmingham. Retrieved from http://www.demos.co.uk/files/insideout.pdf

Deniston, O. L. (1972). *Program planning for disease control programs*. U. S. Department of Health, Education and Welfare. Washington D.C.: Public Health Service.

Deniston, O. L. (1974). *Evaluation of disease control programs*. Ann Arbor: School of Public Health, University of Michigan.

DFID (2002). *Development Effectiveness Review draft*. London: Department for International Development.

DFID (2003). *Department for International Development Departmental Report 2003*. Presented to Parliament by the Secretary of State for International Development. Crown Copyright, UK.

DFID (2011). *DFID's approach to value for money (VfM)*. Retrieved from https://www.gov.uk/government/publications/dfids-approach-to-value-for-money-vfm

DFID (2012). Broadening the range of Designs and Methods for Impact Evaluations. Report of a study commissioned by the Department for International Development. Working Paper 38. London: Department for International Development.

Donaldson, K. (2009). The future of design for development: three questions. *Information Technologies & International Development*, *5*(4), 97.

Drabek, A. G. (1987). Development alternatives: the challenge for NGOs – an overview of the issues. *World Development*, 15 (Supplement 1), ix-xv.

Dreyfus, H. L., & Dreyfus, S. E. (1986). Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. Oxford: Basil Blackwell.

Duflo, E., & Kremer, M. (2005). Use of Randomization in the Evaluation of Development Effectiveness. In G. Pitman, O. Feinstein, & G. Ingram, (Eds.), *Evaluating Development Effectiveness*, (pp.205-31). World Bank Series on Evaluation and Development, Volume 7. New Brunswick NJ and London: Transaction Publishers.

Duncombe, R. (2009). *Impact Assessment of Mobile Phones on Development: Concepts, Methods and Lessons for Practice*. Development Informatics Working Paper No. 39. Manchester: IDPM, University of Manchester.

Duval, A. M., Gendron, Y., & Roux-Dufort, C. (2015). Exhibiting nongovernmental organizations: Reifying the performance discourse through framing power. *Critical Perspectives on Accounting*, *29*, 31-53.

Earl, M. J. (1994). Knowledge as strategy: reflections on Skandia International and Shorko Films. In C. Ciborra & T. Jelassi (Eds.), *Strategic Information Systems* (pp.53-69). New York: John Wiley & Sons.

Earl, S., Carden, F., & Smuttylo, T. (2001). *Outcome Mapping. Building Learning and Reflection into Development Programs.* Ottawa, Canada: International Development Research Centre (IDRC).

Easterby-Smith, M., Antonacopoulou, E., Graca, M., & Ferdinand, J. (2007). Organizational learning and dynamics capabilities. In H. Scarbrough (2008). *The Evolution of Business Knowledge* (pp.71-88). Oxford: Oxford University Press.

Easterby-Smith, M., & Lyles, M. A. (Eds.) (2011). *Handbook of organizational learning and knowledge management*. John Wiley & Sons.

Easterly, W., & Easterly, W. R. (2006). The white man's burden: why the West's efforts to aid the rest have done so much ill and so little good. Penguin.

Edmonds, R., & Cook, M. R. (2014). Starting a conversation: the need for and application of Service Design in International Development. *ServDes 2014 Conference Proceedings*, Lancaster University, UK http://www.servdes.org/wp/wp-content/uploads/2014/06/Edmonds-R-Cook-M-R.pdf

Edwards, M., & Hulme, D. (Eds.) (1995). Beyond the Magic Bullet: NGO Performance and Accountability in the Post-Cold War World. London: Earthscan.

Edwards, M., & Hulme, D. (1996). Too close for comfort? The impact of official aid on nongovernmental organizations. *World Development*, *24*(6): 961-973.

Edwards, W. (1979). Multiattribute utility measurement: Evaluating desegregation plans in a highly political context. In R. Perloff (Ed.), *Evaluation Interventions*, (pp. 13-54). Beverly Hills, CA: Pros & Cons.

Elbers, W., & Arts, B. (2011). Keeping body and soul together: southern NGOs' strategic responses to donor constraints. *International Review of Administrative Sciences*, *77*(4): 713-732.

Elkonin, D. B. (1972). Toward the problem of stages in the mental development of the child. *Soviet Psychology*, *10*, 225-251.

Ellis, V. (2011). Reenergising professional creativity from a CHAT perspective: Seeing knowledge and history in practice. *Mind, Culture, and Activity*, *18*(2), 181-193.

Els, C., & Cartenson, N. (2015). *Funding of Local and National Humanitarian Actors*. Retrieved from http://www.local2global.info/wp-content/uploads/l2gp_local_funding_final_250515.pdf

Engeström, Y. (1987). Learning by expanding: An activity-theoretical approach to developmental research. Helsinki, Finland: Orienta-Konsultit.

Engeström, Y. (1991). Activity theory and individual and social transformation. *Multidisciplinary Newsletter for Activity Theory*, 7/8, 14-15.

Engeström, Y. (1999). Learning by expanding: Ten years after. Autobiographical Note. Introduction to the German edition. Lernen durch Expansion: *An activity-theoretical approach*

to developmental research. Helsinki, Finland: Orienta-Konsultit. Retrieved from http://lchc.ucsd.edu/MCA/Paper/Engestrom/expanding/intro.htm

Engeström, Y. (2000). From individual action to collective activity and back: Developmental work research as an interventionist methodology. In P. Luff, J. Hindmarsh, & C. Heath (Eds.), *Workplace Studies: Recovering Work Practice and Informing System Design* (pp. 150-166). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511628122.008

Engeström, Y. (2000b). Activity theory and the social construction of knowledge: A story of four umpires. *Organization*, *7*, 301-310.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work, 14*(1), 133-156.

Engeström, Y. (2006). From well-bounded ethnographies to Intervening in Mycorrhizae Activities. *Organization Studies, 27,* 1783–1793. https://doi.org/10.1177/0170840606071898

Engeström, Y. (2008). *From teams to knots: Activity-theoretical studies of collaboration and learning at work*. Cambridge: Cambridge University Press. 10.1017/CBO9780511619847

Engeström, Y. (2009). The future of activity theory: A rough draft. In A. Sannino, H. Daniels, & K. D. Gutiérrez (Eds.), *Learning and Expanding with Activity Theory* (pp. 303–328). Cambridge: Cambridge University Press. https://doi.org/10.1017/ CBO9780511809989.020

Engeström, Y. (2011). From design experiments to formative interventions. *Theory and Psychology, 21*(5), 598–628.

Engeström, Y. & Mazzocco, D. W. (1995). Disturbance management and masking in a television production team: An activity-theoretical study in organizational communication. Paper presented at 45th Annual Conference of the International Communication Association, 25-29 May 1995, Albuquerque, NM.

Engeström, Y., Virkkunen, J., Helle, M., Pihlaja, J., & Poikela, R. (1996). The change laboratory as a tool for transforming work. *Lifelong Learning in Europe, 1*(2), 10-17.

Escobar, A. (1995/2012). Encountering Development: The Making and Unmaking of the Third World. Oxford: Princeton University Press.

Escobar, A. (2008). Afterword. In S. Dar, & B. Cooke (Eds.), *The New Development Management: Critiquing the Dual Modernization* (pp. 198–203). London: Zed Books.

Escobar, A. (2018). Designs for the pluriverse: radical interdependence, autonomy, and the making of worlds. Durham, NC: Duke University Press.

Estrella, M. (2000). *Learning from Change*. London: ITDG.

Eyben, R. (2009). Hovering on the threshold: Challenges and opportunities for critical and reflexive ethnographic research in support of international aid practice. In S. Hagberg & C. Widmark (Eds.), *Ethnographic Practice and Public Aid: Methods and Meanings in Development Cooperation*, (pp. 73-100). Uppsala, University of Uppsala Library.

Eyben, R., Guijt, I., Roche, C., & Shutt, C. (Eds.) (2015). The politics of evidence and results in international development: Playing the game to change the rules?. Rugby: Practical Action Publishing https://doi.org/10.3362/9781780448855

Fals-Borda, O. & Rahman, M. (1991). Action and knowledge: Breaking the monopoly with participatory action research. New York: Apex Press.

Feldman, M. S., & March, J. G. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly, 26,* 171–186. https://doi.org/10.2307/2392467

Ferguson, J. (1994). The Anti-politics Machine: Development, Depoliticization and Bureaucratic Power in Lesotho (New edition). London: University of Minnesota Press.

Floridi, L. (2002). What Is the Philosophy of Information?. *Metaphilosophy*, 33(1–2): 123–145. doi:10.1111/1467-9973.00221

Floridi, L. (2011). *The Philosophy of Information.* Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199232383.001.0001

Floridi, L. (2015). The fourth revolution: How the infosphere is reshaping human reality. Oxford, OUP.

Foucault, M. (1979/2000). Governmentality. In J. D. Faubion (Ed.), *Power: Essential Works of Foucault 1954– 1984. Volume 3* (insert page numbers). London: Penguin.

Foucault, M. (1977/1979). *Discipline and punish: The birth of the prison*. New York: Vintage.

Foucault, M. (1980). Power/Knowledge: Selected Interviews and Other Writings, 1972-1977. New York, NY: Pantheon.

Foucault, M. (1983). On the genealogy of ethics: An overview of work in progress. In H. Dreyfus & P. Rabinow (Eds.), *Michel Foucault: Beyond Structuralism and Hermeneutics* (pp.229–264). Chicago, IL: University of Chicago Press.

Foucault, M. (1991). Governmentality. In G. Burchell, C. Gordon, & P. Miller (Eds.), *The Foucault Effect: Studies in Governmentality* (pp.87–104). University of Chicago Press.

Foucault, M. (2001). *Power. Essential Works of Foucault, Volume 3, 1954-1984*. Edited by James B. Faubion. Penguin Books, London.

Fournier, V., & Grey, C. (2000). At the critical moment: Conditions and prospects for critical management studies. *Human relations, 53*(1), 7-32.

Fox, J. (2007). Accountability Politics: Power and Voice in Rural Mexico. Oxford: Oxford University Press.

Fricke, M. (2009). The knowledge pyramid: A critique of the DIKW hierarchy. *Journal of Information Science, 35*(2), 131–142. https://doi.org/10.1177/0165551508094050

Fyvie, C., & Ager, A. (1999). NGOs and innovation: organizational characteristics and constraints in development assistance work in The Gambia. *World Development, 27*(8): 1383-1395.

Gabay, C. (2012). The millennium development goals and ambitious developmental engineering. *Third World Quarterly, 33*(7), 1249-1265.

Garbarino, S., & Holland, J. (2009). *Quantitative and qualitative methods in impact evaluation and measuring results*. Issues Paper commissioned by the UK Department for International Development (DFID), Governance and Social Development Resource Centre (GSDRC). Birmingham, UK.

Gardner, K., & Lewis, D. (2015). Anthropology and development: Challenges for the twenty-first century. London: Pluto Press https://doi.org/10.2307/j.ctt183p13j

Gaventa, J. (2003). *Power after Lukes: a review of the literature*. Brighton: Institute of Development Studies.

Gaventa, J., & McGee, R. (2013). The impact of transparency and accountability initiatives. *Development Policy Review, 31*(s1), s3-s28.

Gaver, W.H., Dunne, A., & Pacenti, E. (1999). Cultural probes. *Interactions, 6*(1), 21-29.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. London: Sage.

Gibson-Graham, J. K. (2006). *A Postcapitalist Politics*. Minneapolis: University of Minnesota Press.

Giddens, A. (1984). The constitution of society: Outline of the theory of structuration. Berkeley: University of California Press.

Gill, L. (1997). Power lines: the political context of nongovernmental organisation (NGO) activity in El Alto, Bolivia. *Journal of Latin American Anthropology, 2*(2): 144-169.

Ginty, R. M. (2012). Routine Peace: Technocracy and Peacebuilding. *Cooperation and Conflict, 47*(3), 287–308. doi:10.1177/0010836712444825.

Green, D. (2015, August 28). Development 2.0, the Gift of Doubt and the Mapping of Difference: Welcome to the Future. [Blog post]. Retrieved from https://oxfamblogs.org/fp2p/development-2-0-the-gift-of-doubt-and-the-mapping-of-difference-welcome-to-the-future/

Green, D. (2018, 7 November). Evil Donors and 'The Literature': Is there a problem with the way academics write about aid? [Blog post]. Retrieved from https://oxfamblogs.org/fp2p/evil-donors-and-the-literature-is-there-a-problem-with-the-way-academics-write-about-aid

Green, M. (2009). Doing developing and writing culture. *Anthropological Theory, 9*(4), 395–417.

Groves, L. (2015). *Beneficiary Feedback in Evaluation*. Retrieved from https://www.gov.uk/dfid-research-outputs/beneficiary-feedback-in-evaluation

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

Guijt, I. (2015). Playing the rules of the game and other strategies. In R. Eyben, I. Guijt, C. Roche, & C. Shutt (Eds.), *The Politics of Evidence and Results in International Development: Playing the Game to Change the Rules* (pp.193–209). Rugby: Practical Action Publishing https://doi.org/10.3362/9781780448855

Guijt, I., & Shah, M. (1998). Waking up to power, conflict and process. In I. Guijt & M. Shah (Eds), *The Myth of Community: Gender Issues in Participatory Development* (pp.1-23). London: Intermediate Technology Publications.

Gurstein, M. (2000). Community informatics: Enabling community uses of information and communications technology. In M. Gurstein (Ed.), *Community informatics: Enabling Communities with Information and Communications Technologies*, (pp. 1-30). London: IDEA Group Publishing.

Hailey, J. (2000). Indicators of identity: NGOs and the strategic imperative of assessing core values. *Development in Practice, 10*(3-4): 402-407.

Hammersley, M. (2005). Is the evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice. *Evidence & Policy: a journal of research, debate and practice, 1*(1), 85-100.

Hammond, K. R., & Adelman, L. (1976). Science, Values, and Human Judgement. *Science, 194*, 389-396.

Hammond, R. J. (1960). *Benefit Cost Analysis and Water-Pollution Control*. Palo Alto: Food Research Institute, Stanford University.

Hannerz, U. (2003). Being there... and there... and there! Reflections on multi-site ethnography. *Ethnography, 4*(2), 201-216.

Haraway, D. J. (1991). *Simians, Cyborgs, and Women: The Reinvention of Nature*. London: Free Association Books.

Hardy, C., & Clegg, S. (1996). Some dare call it power. In C. Hardy, S. Clegg, & W. Nord (Eds.), *Handbook of Organization Studies* (pp. 622–641). London: Sage.

Harper, R. (2000). The social organization of the IMF's mission work. In M. Strathern (Ed.), *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*. (pp 21-55). London: Routledge.

Harper, R. H. (1998). Inside the IMF: an ethnography of documents, technology and organisational action. London: Routledge.

Harris, R W. (2016) How ICT4D Research Fails the Poor. *Information Technology for Development, 22*(1), 177-192. DOI: 10.1080/02681102.2015.1018115

Harriss, J. (2002). Depoliticizing development: The World Bank and social capital. Delhi: Anthem Press.

Hayes, N. (2011). Information Technology and the Possibilities for Knowledge Sharing. In M. Easterby-Smith & M. A. Lyles (Eds.), *Handbook of Organizational Learning and Knowledge Management* (pp.83-104). Chichester: Wiley.

Hayes, N., Introna, L. D., & Kelly, P. (2017). Institutionalizing Inequality: Calculative practices and regimes of inequality in international development. *Organization Studies, 39*(9), 1203-1226. 0170840617694067.

Hayes, N., Miscione, G., & Westrup, C. (2013). Theorising development and technological change. *Info Systems Journal , 23*, 281-285.

Hayes, N. & Walsham, G. (2001). Participation in groupware-mediated communities of practice: a socio-political analysis of knowledge working. *Information and Organization, 11*(4), 263-288.

Hayes, N., & Westrup, C. (2012). Context and the processes of ICT for development. *Information and Organization, 22*(1), 23–36. https://doi.org/10.1016/j.infoandorg.2011.10.001

Hayes, N. & Westrup, C. (2012b). Power/knowledge and impact assessment: creating new spaces for expertise in international development. *New Technology, Work and Employment*, *27*(1), 9-22.

Hayes, N., & Westrup, C. (2014). Consultants as intermediaries and mediators in the construction of information and communication technologies for development. *Information Technologies and International Development, 10*(2), 19–32.

Hedegaard, M., Chaiklin, S., & Jensen, U. J. (1999). Activity theory and social practice: An introduction. In S Chaiklin, M. Hedegaard, & U. J. Jensen (Eds.), *Activity Theory and Social Practice*, (pp. 12-30). Oxford: Aarhus University Press.

Heeks, R. (2002a). i-Development not e-Development: Special Issue on ICTs and Development. *Journal of International Development, 14*(1), 1-11.

Heeks, R. (2002b). Information and Communication Technologies, Poverty and Development. Working Paper No. 5. Manchester: IDPM, University of Manchester.

Heeks, R. (2008). ICT4D 2.0: The next phase of applying ICT for international development. *Computer*, *41*(6). https://doi.org/ 10.1109/MC.2008.192

Heeks, R. (2010). Development 2.0: The IT-enabled transformation of international development. *Communications of the ACM*, *53*(4), 22–24. https://doi.org/10.1145/1721654.1721665

Heeks, R. (2018). Information and Communication Technology for Development (ICT4D). London: Routledge.

Heeks, R., & Molla, A. (2009). *Impact Assessment of ICT-for-Development Projects: A Compendium of Approaches*. Manchester: Development Informatics Group, Institute for Development Policy and Management, University of Manchester.

Henkel, H., & Stirrat, R. (2001). Participation as spiritual duty: empowerment as secular subjection. In B. Cooke, & U. Kothari, (Eds), *Participation: The New Tyranny* (p.168). London: Zed Press.

Hentschel, J. (1999). Contextuality and data collection methods: A Framework and application to health service utilisation. *Journal of Development Studies, 35*(4), 64-94.

Hernandez, K. (2017). *Blockchain for development—hope or hype?* Rapid response briefing, Issue 17, Institute of Development Studies. Retrieved from https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/12945/RRB17.pdf?sequence=38

Herron, M. C., & Quinn, K. M. (2014). A Careful Look at Modern Case Selection Methods. *Sociological Methods & Research, 45*(3):458-492. 0049124114547053.

Hickey, S. (2012). Turning Governance Thinking Upside-down? Insights from 'the politics of what works'. *Third World Quarterly, 33*(7), 1231-1247.

Hislop, D., Newell, S., Scarbrough, H., & Swan, J. (2000). Networks, knowledge and power: Decision making, politics and the process of innovation. *Technology Analysis and Strategic Management, 12*, 399–411.

Holland, J. (Ed.). (2013). *Who counts?: the power of participatory statistics*. Rugby, UK: Practical Action Publishing.

Holzman, L. (2006a). What kind of theory is activity theory? Introduction. *Theory & Psychology, 16*, 5-11. https://doi.org/ 10.1177/0959354306060105

Holzman, L. (2006b). Activating postmodernism. *Theory & Psychology, 16*(1), 109-123. https://doi.org/10.1177/0959354306060110

Honig, D., & Gulrajani, N. (2018). Making good on donors' desire to Do Development Differently. *Third World Quarterly, 39*(1), 68-84.

Hood, C. (1995). The "New Public Management" in the 1980s: variations on a theme. *Accounting, 20*(2-3), 93–109.

Hout, W. (2012). The anti-politics of development: Donor agencies and the political economy of governance. *Third World Quarterly, 33*(3), 405–422.

HTSG (2009) Healthcare Project Report. [Internal version shared with author]

HTSG (2011) Healthcare Project Annual Report. [Internal version shared with author]

HTSG (2013, December) Healthcare Project 6-month report. [Internal version shared with author]

HTSG (2014) Healthcare Project Annual Report. [Internal version shared with author]

HTSG (2015) Migrant Services Project Grant Application. [Internal version shared with author]

Hughes, J., Jewson, N., & Unwin, L. (2007). Communities of practice: A contested concept in flux. In J. Hughes, N. Jewson, & L. Unwin (Eds.), *Communities of Practice: Critical Perspectives* (pp.1-16). London: Routledge.

Hulme, D. (2000). Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice. *World Development, 28*(1), 79–98.

Hulme, D., & Edwards, M. (1996). Too close for comfort? The impact of official aid on nongovernmental organizations, *World Development, 24*(6), 961–973. https://doi.org/10.1016/0305-750X(96)00019-8

Humanrights.gov (2016). *Fact Sheet: Non-Governmental Organizations (NGOs) in the United States*. Retrieved from http://www.humanrights.gov/dyn/2016/01/fact-sheet-non-governmental-organizations-ngos-in-the-united-states

Hyman, E. L., & Dearden, K. (1998). Comprehensive impact assessment systems for NGO microenterprise development programs. *World Development, 26*(2), 261-276. https://doi.org/10.1016/S0305-750X(97)10022-5

IASC (2006, November 24). *Guidance note on using the cluster approach to strengthening humanitarian response*. Retrieved from https://interagencystandingcommittee.org/node/7059

IDEO (2009). Design Kit: The Human - Centered Design Toolkit. Retrieved from https://www.ideo.com/post/design-kit

IDEO (2014). *Human-Centered Design Toolkit. An innovation guide for social enterprises and NGOs*. Retrieved http://www.ideo.com/images/uploads/hcd_toolkit/IDEO_HCD_ToolKit.pdf

Institute of Development Studies (2016). Knowledge, Evidence and Learning for Development (K4D). Retrieved from https://www.ids.ac.uk/programme-and-centre/knowledge-evidence-and-learning-for-development-k4d

IFPRI (2002). Assessing the Impact of Policy-oriented Social Science Research: Measuring Impact. Washington DC: International Food Policy Research Institute.

Ilcan, S., & Phillips, L. (2008). Governing through global networks: Knowledge mobilities and participatory development. *Current Sociology, 56*(5), 711-734. https://doi.org/10.1177/0011392108093832

Ilcan, S. & Phillips, L. (2010). Developmentalities and calculative practices: The millennium development goals. *Antipode, 42*(4), 844-874. https://doi.org/10.1111/j.1467-8330.2010.00778.x

Imagine Foundation UK (2014). Projects page, The Imagine Foundation public website. Website URL withheld to protect anonymity.

Imagine Foundation UK (2013). Annual Progress Report. [Internal version shared with author]

International Federation of Red Cross and Red Crescent Societies (2015, November 30) Rights of Migrants in Action. Guidelines for call for proposals and award of grants for civil society organisations for protecting and promoting the rights of migrant domestic workers and victims of human trafficking (Thailand edition). Thailand: IFRC.

Introna, L. (2003). Disciplining Information Systems: Truth and Its Regimes. *European Journal of Information Systems, 12* (3), 235–240. https://doi.org/10.1057/palgrave.ejis.3000465

Introna, L., & Whittaker, L. (2004). Truth, Journals, and Politics: The Case of the *MIS Quarterly*. in B. Kaplan, D. Truex, D. Wastell, T. Wood-Harper, & J. DeGross (Eds.), *Information Systems Research: Relevant Theory and Informed Practice* (pp.103-120). Boston, MA: Kluwer Academic Publishers.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med, 2*(8), e124.

Irani, L., Vertesi, J., Dourish, P., Philip, K., & Grinter, R. E. (2010) Postcolonial computing: a lens on design and development. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 1311-1320). Atlanta, Georgia: ACM.

Jarzabkowski, P. (2003). Strategic practices: An activity theory perspective on continuity and change. *Journal of Management Studies, 40*(1), 23-55.

Jenkin, E., Wilson, E., Clarke, M., Campain, R., & Murfitt, K. (2016). Spreading the word: using film to share research findings and knowledge about children with disabilities in Vanuatu and Papua New Guinea. *Knowledge Management for Development Journal, 11*(1): 69-84.

Jennex, M. E. (2009). Re-visiting the knowledge pyramid. *42nd Hawaii International Conference on Systems Science (HICSS), Conference* Proceedings, Waikoloa, Big Island, HI, USA. DOI: 10.1109/HICSS.2009.361

Jerven, M. (2014). Data for development assessment paper: Benefits and costs of the data for development targets for the post-2015 development agenda. Working paper. Retrieved from http://www.copenhagenconsensus.com/sites/default/files/data_assessment_-_jerven.pdf

Karanasios, S. (2014). Framing ICT4D research using activity theory: A match between the ICT4D field and theory? *Information Technologies and International Development*, 10(2), 1.

Karanasios, S. (2018). Toward a unified view of technology and activity: The contribution of activity theory to information systems research. *Information Technology & People, 31*(1), 134-155.

Karanasios, S., & Allen, D. (2013) ICT for development in the context of the closure of Chernobyl nuclear power plant: An activity theory perspective. *Information Systems Journal, 23,* 287–306.

Karanasios, S., & Allen, D. (2013). ICT for development in the context of the closure of Chernobyl nuclear power plant: An activity theory perspective. *Information Systems Journal, 23*(4), 287–306.

Keeley, J. & Scoones, I. (1999) Understanding Environmental Policy Processes: A Review. *IDS Working Paper 89*, Brighton: IDS.

Kelly, P. (2016). Impact knowledge supply chain, mining and packaging with Foucault and CHAT: a story of power, knowledge, a philanthropy and a small development NGO. *European Group for Organisation Studies Colloquium (EGOS),* Athens, Greece.

Kelly, P. & Cowen, M. (2014) CHAT and Governmentality: The uniform in the activities and management of migrant worker populations. *European Group for Organization Studies Colloquium (EGOS)*, Rotterdam, Netherlands.

Kelly, P. & Secker, E. (2015, September 3) Quick and Dirty M&E: Co-designing and co-evolving for engaged impact. *Youth | Participation | Impact Conference*, University of Manchester, UK.

Kelly, P. R. (2018). An activity theory study of data, knowledge, and power in the design of an international development NGO impact evaluation. *Information Systems Journal*, 28(3), 465-488. https://doi.org/10.1111/isj.12187

Kelly, P., Luana, D. & Thomas, B. (2018)) Cultivating a participatory approach to governance in Papua New Guinea: our approach, emerging results and learning how to bring small-g governance into big-G Governance arenas. *PNG Update Conference,* Port Moresby, Papua New Guinea.

Kemmis, S. (1993). Foucault, Habermas and evaluation. *Curriculum studies, 1(*1), 35-54.

Kerosuo, H., & Engeström, Y. (2003). Boundary crossing and learning in creation of new work practice. *Journal of Workplace Learning, 15*(7/8), 345-351.

Kerr, R. (2008). International development and the new public management: Projects and logframes as discursive technologies of governance. In S. Dar and B. Cooke (Eds) *The New Development Management: Critiquing the Dual Modernization*. London: Zed Books.

Khandker, S. R., Koolwal, G.B. & Samad, H. A. (2015) *Handbook on Impact Evaluation: Quantitative Methods and Practices.* Washington D.C, The World Bank.

Kirkpatrick, C. & Hulme, D. (2001). *Basic Impact Assessment at Project Level*. CARE International.

Kling, R. & Iacono, S. (1984). The control of information systems developments after implementation. *Communications of the ACM, 27*(12), 1218-1226. https://doi.org/10.1145/2135.358307

KM4D (2012). DIKW Model. *Knowledge Management for Development (KM4Dev) community wiki page discussion record*. Retrieved from http://wiki.km4dev.org/DIKW_model

KM4D (2016, August 24). Call for Abstracts: Open research, open data, and your development organization: Best practices in information and data management for development. *Knowledge Management for Development Journal.*

KM4D (2019) *Home page.* Knowledge Management for Development (KM4D) Journal. http://www.km4djournal.org/index.php/km4dj

Knights, D. (2004). Writing Organization Analysis into Foucault. In S. Linstead (Ed.), Organizations and Postmodern Thought (pp. 14–23). London: Sage.

Knights, D., Murray, F., & Willmott, H. (1993). Networking as knowledge work: A study of strategic inter-organizational development in the financial services industry. *Journal of Management Studies, 30,* 975–995.

Knutsson, B. (2014). Smooth machinery: Global governmentality and civil society HIV/AIDS work in Rwanda. *Globalizations, 11*(6), 793-807.

Koenig, M. & Neveroski, K. (2008). The origins and development of knowledge management. *Journal of Information and Knowledge Management, 7*(4), 243-54.

Kontinen, T. (2007). Learning challenges of NGOs in development: Co-operation of Finnish NGOs in Morogoro, Tanzania. Doctoral dissertation. Helsinki: University of Helsinki Press.

Kontinen, T. (2013). A Gramscian perspective on developmental work research: Contradictions, power and the role of researchers reconsidered. Outlines. *Critical Practice Studies, 14*(2), 106-129.

Korpela, M., Mursu, A., Soriyan, A., Eerola, A., Häkkinen, H., & Toivanen, M. (2004). Information systems research and development by activity analysis and development: Dead horse or the next wave? In B. Kaplan, D. P. Truex, D. Wastell, A. T. Wood-Harper, & J. DeGross (Eds.), *Information Systems Research: Relevant Theory and Informed Practice* (pp. 453–472). Boston: Springer. https://doi.org/10.1007/1-4020-8095-6_25

Korten, D. C. (1990). Getting to the 21st Century: Voluntary Action and the Global Agenda. West Hartford, CT: Kumarian Press.

Kothari, U. (2005a). From colonial administration to development studies: a post-colonial critique of the history of development studies. In U. Kothari, *A Radical History of Development Studies: Individuals, Institutions and Ideologies*, pp. 47-66. London: Zed Books.

Kothari, U. (2005b). Authority and expertise: The professionalisation of international development and the ordering of dissent. *Antipode*, 37(3), 425-446. https://doi.org/10.1111/j.0066-4812.2005.00505.x

Kothari, U. (2014). Trade, Consumption and Development Alliances: The Historical Legacy of the Empire Marketing Board Poster Campaign, *Third World Quarterly, 35*(1), 1-21.

Krauss, A. (2018) Why all randomised controlled trials produce biased results. Annals of Medicine, 50(4), 312-322.

Kristof, N. (2014, March 14). Professors, we need you! *New York Times*. Retrieved from http://www.nytimes.com/2014/02/16/opinion/sunday/kristof-professors-we-need-you.html?_r=0 Accessed 24/1/2019

Kuutti, K. (1996). Activity theory as a potential framework for human-computer interaction research. In B. A. Nardi (Ed.), *Context and Consciousness: Activity Theory and Human-Computer Interaction* (pp. 17–44). Cambridge, MA: The MIT Press.

Kuutti, K. (1999). Activity theory, transformation of work, and information systems design. In Y. Engeström, R. Miettinen, & R. L. Punamki (Eds), *Perspectives on Activity Theory: Learning in Doing: Social, Cognitive and Computational Perspectives* (pp. 361–376). Cambridge: Cambridge University Press. 10.1017/CBO9780511812774.024

Lam, A. (1997). Embedded firms, embedded knowledge: Problems of collaboration and knowledge transfer in global cooperative ventures. *Organization Studies, 18*, 973-996.

Lambe, P. (2010, February 5). From data with love. [Blog post] Retrieved from: http://www.greenchameleon.com/gc/blog_detail/from_data_with_love

Lambe, P. (2011). The unacknowledged parentage of knowledge management. *Journal of Knowledge Management, 15*(2), 175-197.

Lambe, P. (2011). The unacknowledged parentage of knowledge management. *Journal of Knowledge Management, 15*(2), 175– 197. https://doi.org/10.1108/13673271111119646

Latour, B. (1987). Science in action: How to follow scientists and engineers through society. Harvard University Press.

Latour, B. (2005). Reassembling the social: An introduction to actor-network-theory. Oxford University Press.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate Peripheral Participation*. Cambridge University Press.

Law, J. (2004). After method: Mess in social science research. Routledge.

Law, J. (2009). Actor network theory and material semiotics. In B. S. Turner (Ed), *The New Blackwell Companion to Social Theory*, (pp. 141-158). John Wiley and Sons.

Leadbeater, C., & Cottam, H. (2007). *The user generated state: Public services 2.0*. Retrieved from http://charlesleadbeater.net/2006/10/public-services-2-0-demos/

Leal, A. P. (2007). Participation: the ascendancy of a buzzword in the neo-liberal era. *Development in practice, 17*(4-5), 539-548. https://doi.org/10.1080/09614520701469518

Lefebvre, E. & Lefebvre, L.A. (1996). Information and telecommunication technologies: The Impact of Their Adoption on Small and Medium-sized Enterprises. International Development Research Centre. Ottawa, IDRC.

Leftwich, A. (2011, March). Thinking and working politically: What does it mean, why is it important and how do you do it. In *Background Papers of Research and Policy Workshop* (pp. 3-13). Frankfurt, Germany: Developmental Leadership Program.

Leitch, S. (2012). *Global evaluation framework agreement.* Tender, Department for International Development). Retrieved from http://com.mercell.com/permalink/32642486.aspx

Lemke (2002). Foucault, governmentality, and critique. *Rethinking Marxism, 14*(3), 49-64, DOI: 10.1080/089356902101242288

Lennie, J., & Tacchi, J. (2013). Evaluating communication for development: A framework for social change. Routledge.

Leonard-Barton, D. (1995). Wellsprings of knowledge: Building and sustaining the sources of innovation. Harvard Business School Press.

Leonardi, P. M. (2011). When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies. *MIS Quarterly*, 35(1), 147-167.

Leontyev, A. N. (1978). *Activity, consciousness, and personality*. Englewood Cliffs: Prentice-Hall. https://www.marxists.org/archive/leontev/works/1978/index.htm

Leontyev, A. N. (2009). *The development of mind*. (A. Blunden, Trans.). Marxists Internet Archive. Retrieved from http://marxists.anu.edu.au/archive/leontev/works/development-mind.pdf

Lewin, K. (1943). Forces behind food habit and methods of change. In The problem of changing food habit. *Bulletin of the National Research Council*, 108, 35-65.

Lewin, K. (1946) Action research and minority problems. *Journal of Social Issues, 2*(4), 34-36.

Lewin, K. (1947). Frontiers in group dynamics I. Concept, method and reality in Social Science; social equilibria. *Human Relations, 1*, 5-41

Lewis D (2007). The management of non-governmental development organizations (2nd edition). Routledge.

Lewis, D. (2005) Anthropology and development: the uneasy relationship. In J. G. Carrier (Ed.) *A Handbook of Economic Anthropology*, (pp. 472-486). Cheltenham, UK: Edward Elgar.

Lewis, D. (2005). *Anthropology and development: the uneasy relationship*. London: LSE Research Online. Retrieved from http://eprints.lse.ac.uk/archive/00000253

Lewis, D., & Mosse, D. (Eds.). (2006). Development brokers and translators: The ethnography of aid and agencies. Kumarian Press.

Lewis, D. & Kanji, N. (2009). *Non-governmental organizations and development*. Abingdon, UK: Routledge.

Lewis, D. & Madon, S. (2004). Information systems and nongovernmental development organizations: Advocacy, organizational learning, and accountability. *The Information Society, 20,* 117–126.

Li, T. M. (2007). The will to improve: governmentality, development, and the practice of politics. Durham, NC: Duke University Press.

Lie, J. H. S. (2011). Developmentality. An ethnography of the new aid architecture and the formation of the World Bank Uganda Partnership. PhD dissertation. University of Bergen.

Lie, J. H. S. (2015a). Developmentality: an ethnography of the World Bank-Uganda partnership. Berghahn Books.

Lie, J. H. S. (2015b). Developmentality: indirect governance in the World Bank–Uganda partnership. *Third World Quarterly, 36,* 723–740.

Long, C. (2001). Participation of the poor in development initiatives: Taking their rightful place. London: Earthscan.

Long, N. (2004). Development sociology. Actor perspectives. London: Routledge.

Long, N., & Long, A. (Eds.). (1992). Battlefields of knowledge: the interlocking of theory and practice in social research and development. London: Routledge.

Lyon, D. (1993). An electronic panopticon? A sociological critique of surveillance theory. *The Sociological Review, 41*(4), 653-678.

Lyotard, J. F. (1984). The postmodern condition: A report on knowledge (Volume 10). University of Minnesota Press.

MacDonald, B. (1993). A political classification of evaluation studies in education. In M. Hammersley (Ed.), *Social research: Philosophy, Politics and Practice*. London: Sage.

Machlup, F. (1962). *The production and distribution of knowledge in the United States*, Princeton, NJ: University Press, Princeton.

Madon, S., & Sahay, S. (2002). An information-based model of NGO mediation for the empowerment of slum dwellers in Bangalore. *The Information Society, 18*(1), 13-19. https://doi.org/10.1080/01972240252818199

Mama, A. (2000). Strengthening civil society: participatory action research in a militarised state. *Development in Practice, 10*(1), 59-70. https://doi.org/10.3362/9780855987015.009

Mama, A. (2000). Strengthening civil society: participatory action research in a militarised state. *Development in Practice, 10*(1), 59-70.

MandEOnline (2017) *Home page*. Retrieved from https://mande.co.uk/

Marcus, G. E. (1995). Ethnography in/of the world system: the emergence of multi-sited ethnography. *Annual review of Anthropology*, 95-117.

Marcus, G. E. (1998). *Ethnography through thick and thin*. Princeton University Press.

Markoff, J. (2005) Taking the pulse of technology at Davos. *The New York Times*. https://www.nytimes.com/2005/01/31/technology/taking-the-pulse-of-technology-at-davos.html

Markus, M. L. (1983). Power, politics, and MIS implementation. *Communications of the ACM*, *26*, 430-444. https://doi.org/10.1145/358141.358148

Marsden, D. & Oakley, P. (1990). *Evaluating social development projects*. Oxford, UK: Oxfam GB.

Marx, K. (1845/1969). *Thesis on Feuerbach*. Progress Publishers. Retrieved from https://www.marxists.org/archive/marx/works/1845/theses/theses.htm

Mathison, S. (2004). *Encyclopedia of evaluation*. Sage publications.

McGregor, A., Challies, E., Overton, J. & Sentes, L. (2013). Developmentalities and donor–NGO relations: Contesting foreign aid policies in New Zealand/Aotearoa. *Antipode, 45,* 1232–1253

Merriam, S. B. (2001). Qualitative research and case study application in education. San Francisco: Jossey-Bass.

Miettinen, R., Paavola, S. & Pohjola, P. (2012) From habituality to change: contribution of activity theory and pragmatism to practice theories. *Journal for the Theory of Social Behaviour, 42*, 345–360.

Miettinen, R., Paavola, S., & Pohjola, P. (2012). From habituality to change: Contribution of activity theory and pragmatism to practice theories. *Journal for the Theory of Social Behaviour*, 42, 345–360. https://doi.org/10.1111/j.1468- 5914.2012.00495.x

Miettinen, R., Samra-Fredericks, D., & Yanow, D. (2009). Re-turn to practice: An introductory essay. *Organization studies, 30*(12), 1309-1327.

Miller, P. (2004) Governing by Numbers: Why Calculative Practices Matter. In A. Amin and N. Thrift, (Eds.) *The Blackwell Cultural Economy Reader*, (pp. 179-189). Oxford: Blackwell Publishing Ltd.

Miller, P. & Rose, N. (2008). Governing the present: Administering economic, social and personal life. Cambridge: Polity Press.

Miller, P., & Rose, N. (1990). Governing economic life. *Economy and Society, 19,* 1–31.

Millington, M., Tawake, S., Madden, R., Soares, J., Sindu, P., Lukersmith, S., Scarf, C. & Fa'asala, C. J., (2016). Monitoring of community-based rehabilitation as empowerment: experiences from the Asia Pacific region. *Knowledge Management for Development Journal 11*(1): 25-40

Mitchell, T. 1. (2002). *Rule of experts: Egypt, techno-politics, modernity*. Berkeley: University of California Press.

Mohan, G. (2001). Beyond participation: strategies for deeper empowerment. In B. Cooke & U. Kothari. (Eds.), *Participation: The New Tyranny?,* (pp. 153–167). London: Zed Books.

Mohan, G. (2002) The disappointments of civil society: the politics of NGO intervention in northern Ghana. *Political Geography, 21*(1) pp. 125–154.

Mohan, G. & Stokke, K. (2000). Participatory development and empowerment: The dangers of localism, *Third World Quarterly, 21*(2), 247-268, DOI: 10.1080/01436590050004346

Mohr, L. B. (1995). Impact analysis for program evaluation. Sage.

Mosse, D. (2004). Social analysis as product development: anthropologists at work in the World Bank. In A. K. Giri, A. Van Harskamp, & O. Salemink (Eds.), *The Development of Religion, The Religion of Development* (pp. 77–88). Netherlands: Eburon Uitgeverij BV.

Mosse, D. (2004b). Is good policy unimplementable? Reflections on the ethnography of aid policy and practice. *Development and Change, 35*(4), 639-671.

Mosse, D. (2005). Cultivating development: Ethnography of aid policy and practice. London: Pluto Press.

Mosse, D. (2006). Anti-social anthropology? Objectivity, objection, and the ethnography of public policy and professional communities. Jou*rnal of the Royal Anthropological Institute, 12*(4), 935-956.

Mosse, D. (2007). Notes on the ethnography of expertise and professionals in international development. In *Ethnografeast III: Ethnography and the public sphere* (pp. 1–17). Lisbon 20–23 June 2007. Retrieved from http://ceas.iscte.pt/ethnografeast/papers/david_mosse.pdf

Mosse, D. (2008). International policy, development expertise, and anthropology. *Focaal, 2008 (52),* 119.

Mosse, D. (2011). Introduction: The anthropology of expertise and professionals in international development. In D. Mosse, (Ed.), *Adventures in Aidland: The Anthropology of Professionals in International Development.* Berghahn Books.

Mosse, D. (2011). Politics and ethics: ethnographies of expert knowledge and professional identities. In S. Wright & C. Shore, (Eds.), *Policy Worlds: Anthropology and Analysis of Contemporary Power*. Oxford UK: Berghahn, pp. 50-67.

Mosse, D. (2013). The anthropology of international development. A*nnual Review of Anthropology, 42,* 227-246.

Mowles, C., Stacey, R. & Griffin, D. (2008). What contribution can insights from the complexity sciences make to the theory and practice of development management?. *Journal of International Development 20*(6): 804–820.

Moyo, D. (2009). Dead aid: Why aid is not working and how there is a better way for Africa. Macmillan.

Mukute, M., & Lotz-Sisitka, H. (2012). Working with cultural-historical activity theory and critical realism to investigate and expand farmer learning in southern Africa. *Mind, Culture, and Activity, 19*(4), 342-367.

Muller, J. Z. (2018). *The tyranny of metrics*. Princeton University Press.

Najam, A. (1996). Understanding the third sector: revisiting the prince, the merchant and the citizen. *Nonprofit Management and Leadership 7*(2), 203–19.

Narasimhan, M., & Arun, A. (2017, July 11). Shifting gears: randomised control trials and the future of development evaluation. [Blog post]. Retrieved from https://blogs.lse.ac.uk/southasia/2017/07/11/shifting-gears-randomised-control-trials-and-the-future-of-development-evaluation/

Nardi, B. A. (Ed.) (1996). Context and consciousness: activity theory and human-computer interaction. MIT Press.

National Research Council (2004). Adaptive management theories, frameworks and Practices. adaptive management for water resources project planning. Washington D.C.: The National Academies Press.

Nauta, W. (2004) A moral critique of fieldwork. In: O. Salemink & A. van Harskamp. (Eds.) *The Development of Religion, The Religion of Development*, (pp. 89-97). Delft, Netherlands: Eburon.

Nicholls, L. (1999). Birds of a feather? UNDP and ActionAid implementation of sustainable human development, *Development in Practice, 9(4)*, 396-409, DOI: 10.1080/09614529952882

Nicolini, D. (2012) Practice theory, work, and organization: An introduction. Oxford, UK: Oxford University Press.

Nonaka, I. & Konno, N. (1998). The concept of "ba": Building a foundation for knowledge creation. *California Management Review, 40*(3), 40-54.

Nonaka, I. & Takeuchi, H. (1995) The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford University Press.

Norgaard, R. & Deb, D. (2009). Beyond developmentality: Constructing inclusive freedom and sustainability (First edition). London: Routledge.

Norris, N. (2015). Democratic evaluation: The work and ideas of Barry MacDonald. *Evaluation, 21*(2), 135-142. https://doi.org/10.1177/1356389015577510

OECD DAC (1991). *Principles for evaluation of development assistance*. Development Assessment Committee (DAC). Paris, France: Organisation for Economic Co-operation and Development (OECD).

OECD-DAC (2002). *Glossary of key terms in evaluation and results based management.* Development Assessment Committee (DAC), Paris, France: Organisation for Economic Co-operation and Development (OECD).

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Penguin Random House.

Orland-Barak, L., & Becher, A. (2011). Cycles of action through systems of activity: Examining an action research model through the lens of activity theory. *Mind, Culture, and Activity, 18*(2), 115-128.

Orlikowski, W. J. (2007). Sociomaterial practices: Exploring technology at work. *Organization studies, 28*(9), 1435-1448. https://doi.org/10.1177/0170840607081138

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research, 2,* 1–28. https://doi.org/10.1287/isre.2.1.1

Packer, M. J. & Goicoechea, J. (2000) Sociocultural and constructivist theories of learning: Ontology, not just epistemology. *Educational Psychologist, 35*(4): 227–241.

PARC (2004). Impact Assessment: The Search for the End of the Rainbow? Birmingham, UK: PARC.

Parkhurst, J. (2017). The politics of evidence: from evidence-based policy to the good governance of evidence. Taylor & Francis.

Pascale, R., Pascale, R. T., Sternin, J., & Sternin, M. (2010). The power of positive deviance: How unlikely innovators solve the world's toughest problems (Vol. 1). Harvard Business Press.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. California: Sage Publications.

Patton, P. Q. (2014a). Qualitative research and evaluation methods (4th Edition). London: Sage Publications.

Patton, P. Q. (2014b, November 3). Week 46: Rumination #2: Confusing empathy with bias. [Blog post]. Retrieved from http://betterevaluation.org/blog/anecdote_as_epithet

Patton, P. Q. (2014c, December 4). Week 47: Rumination #3: Fools' gold: the widely touted methodological "gold standard" is neither golden nor a standard. [Blog post]. Retrieved from http://betterevaluation.org/blog/fools_gold_widely_touted_methodological_gold_standard

Paul, R. (2010, April 8). OLPC's Negroponte offers to help India realize $35 tablet. *Arstechnica.* Retrieved from https://arstechnica.com/gadgets/2010/08/olpcs-negroponte-offers-to-help-india-realize-35-tablet/

Peacock, Alan T. & Douglas G.M. Dosser (1958). The National Income of Tanganyika 1942-1954, *Colonial Office Research Study No. 26*, London: HMSO.

Perloff, R. O. (Ed.). (1979). *Evaluator interventions: pros and cons.* Sage Publications.

Perrin, B. (2012). Linking monitoring and evaluation to impact evaluation. Impact Evaluation Notes, (2). InterAction and the Rockefeller Foundation.

Perrow, C. (1984). Normal accidents: Living with high risk technologies. New York: Basic Books.

Phillips, J. J., & Phillips, P. P. (2016). H*andbook of training evaluation and measurement methods*. Routledge.

Phillips, L. & Ilcan, S. (2004). Capacity-building: the neoliberal governance of development. *Canadian Journal of Development Studies/Revue canadienne d'études du développement, 25*(3), 393-409.

Picciotto, R. (2002). The Logic of Mainstreaming: A Development Evaluation Perspective. *Operations Evaluation Department*. Washington D.C.: The World Bank.

Picciotto, R. (2012). Experimentalism and development evaluation: Will the bubble burst?. *Evaluation, 18*(2), 213-229.

Picciotto, R. (2013). Evaluation independence in organizations. *Journal of MultiDisciplinary Evaluation, 9*(20), 18-32.

Picciotto, R. (2015). Democratic evaluation for the 21st century. *Evaluation, 21,* 150–166. https://doi.org/10.1177/1356389015577511

Pinch, T. & Bijker, W. E. (1987). The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Knowledge Might Benefit Each Other. In W. Bijker, T. Hughes, & T. Pinch. (Eds.), *The Social Construction of Technological Systems*, (pp. 17-50). Cambridge, Massachusetts: MIT Press.

Pingali, P.L. (2001). *Milestones in impact assessment research in the CGIAR, 1970–1999*. Standing Panel on Impact Assessment, Technical Advisory Committee. Washington D.C.: CGIAR.

Plaskoff, J. (2003). Intersubjectivity and community-building: Learning to learn organizationally. In M. Easterby-Smith and M. A. Lyles (Eds.), *Handbook of Organizational Learning and Knowledge Management*, (pp. 199-223). John Wiley and Sons.

Polanyi, M. (1967). *The tacit dimension*. New York: Anchor.

Ponte, S. & Richey, L. A. (2014). Buying into development? Brand aid forms of cause-related marketing, *Third World Quarterly, 35*(1), 65-87, DOI: 10.1080/01436597.2014.868985

Porter, T. M. (1996). Trust in numbers: The pursuit of objectivity in science and public life. Chichester, UK: Princeton University Press.

Power, G., Maury, M. & Maury, S. (2002). Operationalising bottom-up learning in international NGOs: barriers and alternatives. *Development in Practice 12* (3-4), 272- 284.

Power, M. (1996). Making things auditable. *Accounting, Organizations and Society, 21*(2), 289-315.

Prasad, A (ed) (2003). *Postcolonialism and Organizational Theory*. New York. St Martin's Press.

Press, G. (2013, May 28). A Very Short History of Data Science. *Forbes*. Retrieved from: https://www.forbes.com/sites/gilpress/2013/05/28/a-very-short-history-of-data-science/#23dcb355cfc0

Prieto-Martin, P., Faith, B., Hernandez, K., & Ramalingam, B. (2017). *Doing Digital Development Differently: lessons in adaptive management from technology for governance initiatives in Kenya*. Making All Voices Count Research Report. Brighton: IDS. Retrieved from

https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/13285/MAVC_DDD_RR
_(Pr4)Final_WEB.pdf?sequence=1

Procacci, G. (1991) Social economy and the government of poverty. In M. Foucault, G. Burchell, C. Gordon, P. Miller, *The Foucault effect: Studies in Governmentality,* (pp. 151-168). University of Chicago Press.

Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y., & Brück, T. (2015). *What methods may be used in impact evaluations of humanitarian assistance?,* Working Paper 22. New Delhi: International Initiative for Impact Evaluation (3ie).

Puri, S., Sahay, S., & Lewis, J. (2009). Building participatory HIS networks: A case study from Kerala, India. *Information and Organization, 19*(2), 63–83.

Quaggiotto, G. (2007, May 1). Development 2.0: A new paradigm for the non-profit sector?. [Blog post]. Retrieved from https://web.jinfo.com/go/blog/2496

Quaggiotto, G. (2009, May 6). A development 2.0 manifesto. [Blog post]. http://blogs.worldbank.org/psd/a-development-20-manifesto

Quaggiotto, G. (2010, October 8). Development 2.0: The skills gap. [Blog post]. *http://blogs.worldbank.org/psd/development-20-the-skills-gap*

Quarles van Ufford, P & Giri, A. K. (2003) Development as a shared responsibility: Ethics, aesthetics and a creative shaping of human possibilities. In A. K. Giri & P. Q. van Ufford, (Eds.), *A Moral Critique of Development: In Search of Global Responsibilities*. Routledge.

Quarles van Ufford, P. (1988). The myth of rational development policy: evaluation versus policy making in Dutch Protestant donor agencies. In P. Quarles van Ufford, D. Kruijt & T. Downing (Eds.), *The Hidden Crisis in Development: Development Bureaucracies*. Tokyo and Amsterdam: United Nations and Free University Press.

Ramalingam, B. (2013). Aid on the edge of chaos: rethinking international cooperation in a complex world. Oxford University Press.

Reason, P. & Bradbury, H. (2001). *The Sage handbook of action research*. Participatory Inquiry and practice. London, Sage.

Reckwitz, A. (2002). Toward a theory of social practices: A development in culturalist theorizing. *European Journal of Social Theory, 5*(2), 243–263.

Richey, L. A. & Ponte, S. (2014). New actors and alliances in development, *Third World Quarterly, 35*(1), 1-21, DOI: 10.1080/01436597.2014.868979

Richmond, O. P., & Carey, H. F. (Eds). (2005). *Subcontracting peace: the challenges of the NGO peacebuilding*. Ashgate Publishing.

Roberts, S. M., Jones III, J. P., & Fröhling, O. (2005). NGOs and the globalization of managerialism: A research framework. *World Development, 33*, 1845–1864.

Roche, C. (1999). *Impact assessment for development agencies: Learning to value change*. Oxford: Oxfam. https://doi.org/10.3362/9780855987701

Roche, C., & Kelly, L. (2012). *The evaluation of politics and the politics of evaluation*. Development Leadership Program (DLP) Background Paper 11. Retrieved from http://publications.dlprog.org/The%20Evaluation%20of%20Politics%20and%20the%20Politics%20of%20Evaluation.pdf

Roloff, K. S., Woolley, A. W., & Edmondson, A. C. (2011). The contribution of teams to organizational learning. In M. Easterby-Smith and M. A. Lyles (Eds.), *Handbook of Organizational Learning and Knowledge Management*, (pp. 249-271). John Wiley and Sons.

Rondinelli, D. A. (2013). Development projects as policy experiments: An adaptive approach to development administration. London: Routledge.

Rowlands, J. (2003). Development methods and approaches: Critical reflections. Oxfam GB.

Rowley, J. & Rubin, F. (2006). *Effective consultancies in development and humanitarian programmes*. Oxfam UK. Retrieved from https://policy-practice.oxfam.org.uk/publications/effective-consultancies-in-development-and-humanitarian-programmes-115364

Rowley, J. E. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science, 33*(2), 163–180. https://doi.org/10.1177/0165551506070706

Saito-Jansen, M. & Pasgaard, M. (2014). Paper. Blocked learning in development aid? Reporting success rather than failure in Andhra Pradesh, India. *Knowledge Management for Development Journal 10*(3): 4-20.

Salancik, G. R. (1979). Field stimulations for organizational behavior research. *Administrative Science Quarterly*, 638-649.

Sandbrook, R. (2011). Polanyi and Post-neoliberalism in the Global South: Dilemmas of Re-embedding the Economy. *New Political Economy, 16*(4), 415-443. DOI: 10.1080/13563467.2010.504300

Sandercock, L. (1998). *Making the invisible visible: A multicultural planning history*. Berkeley: University of California Press.

Sassen, S. (2002). Towards a sociology of information technology. *Current Sociology, 50*(3), 365-388.

Schatzki, T. R. (1996). Social practices: A Wittgensteinian approach to human activity and the social. Cambridge: Cambridge University Press.

Schonhuth, M. (2002). Negotiating with knowledge at development interfaces: anthropology and the quest for participation. In P. Sillitoe, A. Bicker & J. Pottier, (Eds.), *Participating in Development: Approaches to Indigenous Knowledge*. London: Routledge.

Schreyögg, G. & Geiger, D. (2007). The significance of distinctiveness: A proposal for rethinking organizational knowledge. *Organization, 14*(1), 77-100.

Schuurman, F. J. (2009). Critical Development Theory: moving out of the twilight zone. *Third World Quarterly, 30*(5), 831-848.

Scott, J. C. (1985). *Weapons of the weak: Everyday forms of peasant resistance*. Connecticut: Yale University Press.

Scott, J. C. (1998). Seeing like a State: How Certain Schemes to Improve the Human Condition have Failed. Connecticut: Yale University Press.

Seddon, J. (2005). Freedom from Command and Control: A better way to make work work. Buckingham, UK.: Vanguard Books.

Sharma, N. (2007). *The origin of data information knowledge wisdom (DIKW) hierarchy*. Retrieved from https://www.researchgate.net/publication/ 292335202_The_Origin_of_Data_Information_Knowledge_Wisdom_DIKW_Hierarchy

Shivji, I. G. (2006). Lawyers in neoliberalism: Authority's professional supplicants or society's amateurish conscience? *CODERESIA Bulletin 3 & 4*: 16–25.

Shivji, I. G. (2007). Silences in NGO discourse: The role and future of NGOs in Africa. Fahamu/Pambazuka.

Shucksmith, M. (2016). *InterAction How can academics and the third sector work together to influence policy and practice*. Dunfermline: Carnegie UK Trust. Retrieved from https://www.issuelab.org/resource/interaction-how-can-academics-and-the-third-sector-work-together-to-influence-policy-and-practice.html

Silitoe, P. (2002). Participant observation to participatory development: making anthropology work. In P. Sillitoe, A. Bicker & J. Pottier, (Eds.), *Participating in Development: Approaches to Indigenous Knowledge*. London: Routledge.

Simeonova, B., Karanasios, S., Galliers, R. D., Kelly, P., & Mishra, J. (2018a). New ways of organising in collaborative knowledge sharing: examining the effects of power. *European Group for Organizational Studies Colloquium (EGOS)*, 5-7 July 2018, Tallinn, Estonia.

Simeonova, B., Karanasios, S., Galliers, R. D., Kelly, P., & Mishra, J. (2018b). Where is Power in Information Systems Research? Towards a Framework. *International Conference on Information Systems (ICIS)*, 13-16 December, 2018, San Francisco. Retrieved from https://aisel.aisnet.org/icis2018/behavior/Presentations/10/

Snell, R. & Hong, J. (2011). Organizational learning in Asia. In M. Easterby-Smith and M. A. Lyles (Eds.), *Handbook of Organizational Learning and Knowledge Management*, (pp. 635-658). John Wiley and Sons.

Snowden, D. (2002). Complex acts of knowing: Paradox and descriptive self-awareness. *Journal of Knowledge Management, 6* (2), 100–111. https://doi.org/10.1108/ 13673270210424639

Solnes Mitenburg, A., Rijkers, E., Maselle, N., Barass, J., van Roosmalen, J., & Bunders, J. G. (2015). Reflections on the dynamics of the coexistence of multiple knowledge cultures in a community-based maternal health project in Tanzania. *Knowledge Management for Development Journal, 9*(2), 162-184.

Somekh, B., & Nissen, M. (2011). Cultural-historical activity theory and action research. *Mind, Culture, and Activity, 18*(2), 93-97.

Spender, J. C. & Scherer, A. G. (2007). The philosophical foundations of knowledge management: Editors' introduction. *Organization, 14*(1), 5-28.

Springer-Heinze, A., Hartwich, F., Henderson, J.S., Horton, D. & Minde, I. (2002). Impact pathway analysis: an approach to strengthening the impact orientation of agricultural research. *International Conference on Impacts of Agricultural Research and Development*, February 2002, Costa Rica.

Srinivas, N. (2009). Against NGOs? A critical perspective on nongovernmental action. *Non-profit and Voluntary Sector Quarterly 38*(4): 614-626.

St. Clair, A. L. (2006). The World Bank as a transnational expertised institution. *Global Governance, 12*(1), 77-95.

Stahl, B. C. (2004). Whose discourse? A comparison of the Foucauldian and Habermasian concepts of discourse in critical IS research. Proceedings of the Tenth American Conference on Information Systems (*AMCIS)* (pp. 4329-4336), New York.

Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values, 35*(5), 601-617.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science, 19*(3), 387-420.

Stern, E (2015). *Impact Evaluation. A Guide for Commissioners and Managers*. Bond Publications. http://betterevaluation.org/resources/overview/impact_evaluation_bond

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., & Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations*. Report of a study commissioned by the

Department for International Development, Working Paper 38, April 2012, UK Department for International Development. https://doi.org/10.22163/fteval.2012.100

Stiglitz, J. E. (1999). The World Bank at the millennium. *The Economic Journal, 109*(459), 577-597.

Stoll, K., Menou, M.J., Camacho, K. & Khelladi, Y. (2002). Learning about ICT's role in development: A framework toward participatory, transparent and continuous process. Ottawa, Canada: IDRC.

Strathern, M. (2000). Audit Cultures: Anthropological Studies in Accountability, Ethics, and the Academy. Abingdon: Routledge.

Suchman, L. A. (1987). Plans and Situated Actions: The Problem of Human-Machine Communication. Cambridge University Press.

Sumner, A. (2006). What is development studies?. *Development in Practice, 16*(6), 644-650.

Swaans, K., Cullen, B., van Rooyen, A., Adekunle, A. Ngwenya, H., Lema, Z. and Nederlof, S. (2013). Dealing with critical challenges in African innovation platforms: lessons for facilitation. *Knowledge Management for Development Journal 9*(3): 116-135.

Taylor, S. & Osland, J. S. (2003). The impact of intercultural communication on global organizational learning. In M. Easterby-Smith and M. A. Lyles (Eds.), *Handbook of Organizational Learning and Knowledge Management*, (pp. 579-604). John Wiley and Sons.

The Indian Express (2010, July 7). First official estimate: An NGO for every 400 people in India. *The Indian Express*. Retrieved from http://www.indianexpress.com/news/first-official-estimate-an-ngo-for-every-400-people-in-india/643302

Swinford, S. (2016, July 30). Britain to 'leverage' £11bn of foreign aid to build new trade deals after Brexit. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/news/2016/07/30/britain-to-leverage-11bn-of-foreign-aid-to-build-new-trade-deals/

The World Bank (2017). *The World Bank and Knowledge Management: Knowledge for Development Global Partnership Conference 2017*, Geneva, Switzerland. Retrieved from

http://www.worldbank.org/en/events/2017/04/10/world-bank-and-knowledge-management-knowledge-for-development-global-partnership-conference-2017

Thompson, M. (2008). ICT and development studies: Towards development 2.0. *Journal of International Development, 20*(6),821–835. https://doi.org/10.1002/jid.1498

Thompson, M. P. (2002). Cultivating meaning: Interpretive fine-tuning of a South African health information system. *Information and Organization, 12*(3), 183–211. https://doi.org/10.1016/S1471-7727(01)00016-1

Thompson, M. P. (2004). ICT, power, and developmental discourse: A critical analysis. *The Electronic Journal of Information Systems in Developing Countries, 20*(1), 1-25.

Thrift, N. (1997). The rise of soft capitalism. *Cultural Values, 1*(1), 29–57.

Townley, B. (1995). Managing by Numbers: Accounting, Personnel Management and the Creation of a Mathesis. *Critical Perspectives on Accounting, 6*(6) 555–575.

Townley, B., Cooper, D. J., & Oakes, L. (2003). Performance measures and the rationalization of organizations. *Organization Studies, 24*(7), 1045-1071.

Trochim, W. M. (2006). The planning - evaluation cycle page. *Research Methods Knowledge Base, 2nd Edition.* Retrieved from http://www.socialresearchmethods.net/kb/pecycle.php

Tsing A. L. (2004) *Friction: An Ethnography of Global Connection*. Princeton University Press.

Tukey, J. W. (1962). The future of data analysis. *Ann. Math. Statist.* 33 1–67, 812.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison- Wesley, Reading, MA.

Tuomi, I. (1999). Data is more than knowledge: Implications of the reversed knowledge hierarchy for knowledge management and organizational memory. *Journal of Management Information Systems, 16,* 103–117.

Tvedt, T. (2006). The international aid system and the non-governmental organizations: a new research agenda. *Journal of International Development 18*, 677-690.

UN IEAG (2014, November). *A world that counts: mobilizing the data revolution for sustainable development.* United Nations Secretary-General's Independent Expert Advisory group on Data Revolution for Sustainable Development (IEAG). Retrieved from http://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf

UN OCHA (2016). *What is the cluster approach?* United Nations Office for the Coordination of Humanitarian Affairs. Retrieved from https://www.humanitarianresponse.info/en/about-clusters/what-is-the-cluster-approach

UN Women (2015). *How to manage Gender-Responsive Evaluation: Evaluation Handbook.* Retrieved from http://www.unwomen.org/en/digital-library/publications/2015/4/un-women-evaluation-handbook-how-to-manage-gender-responsive-evaluation

UNDP (2002). *UN Handbook on Monitoring & Evaluating for Results*. Retrieved from http://web.undp.org/evaluation/documents/HandBook/ME-Handbook.pdf

UNDP (2006). *Gender: Evaluation of Gender Mainstreaming in UNDP*. Retrieved from http://web.undp.org/evaluation/documents/eo_gendermainstreaming.pdf

United Nations (2016a). *We can end poverty: Millennium development goals and beyond 2015.* Retrieved from http://www.un.org/millenniumgoals.

United Nations (2016b). *Sustainable development goals: 17 goals to transform our world*. Retrieved from https://sustainabledevelopment.un.org/?menu=1300.

USAID (2016, March 28). *Ebola: The response.* Retrieved from https://www.usaid.gov/ebola

van den Berg, R. D. (2004). Evaluating the fundamentalism of evaluation. In A. K. Giri, A. van Harskamp, & O. Salemink (Eds.), *The Development of Religion, the Religion of Development*. Netherlands: Eburon Uitgeverij BV.

Van der Ham, A.J., Shields, L. S., and Broerse, J.E.W. (2013). Towards integration of service user knowledge in mental healthcare in low and middle-income countries: insights from transition theory. *Knowledge Management for Development Journal 9*(2): 125-139.

van Kerkhoff, L. & Lebel, L. (2006). Linking knowledge and action for sustainable development. *Annual Review of Environment and Resources, 31*, 445-477.

Van Maanen, J. (2011). *Tales of the field: On writing ethnography*. University of Chicago Press.

Virkkunen, J., & Newnham, D. (2013). *The change laboratory: A tool for collaborative development of work and education*. Rotterdam: Sense Publishers. https://doi.org/10.1007/978-94-6209-326-3

Vogel, E. (2012). *Review of the use of 'Theory of Change' in international development*. Retrieved from https://assets.publishing.service.gov.uk/media/57a08a5ded915d3cfd00071a/DFID_ToC_Review_VogelV7.pdf

Von Krogh, G. and Roos, J. (1995). *Organizational Epistemology*. New York: St. Martin's Press.

Vygotsky, L. S. (1962). *Language and thought*. Cambridge: MIT Press.

Vygotsky, L. S. (1978). *Mind in society*. The development of the Higher Psychological Processes. Cambridge, MA: Harvard University Press.

Wallace, D.P. (2007). Knowledge Management: historical and cross-disciplinary themes. Westport, CT: Libraries Unlimited.

Wallace, M. (2015). *From Principle to Practice: A User's Guide to Do No Harm*. Cambridge, MA: CDA Collaborative Learning Projects.

Wallace, T., Bornstein, L., & Chapman, J. (2006). *The aid chain: Coercion and commitment in development NGOs*. Warwickshire: Intermediate Technology Publications. https://doi.org/10.3362/9781780440019

Wallerstein, I. (1974). The modern world system: Capitalist agriculture and the origins of the European world economy in the Sixteenth Century. New York: Academic Press.

Walsham, G. (2001). Making a world of difference: IT in a global context. Chichester: Wiley.

Walsham, G. (2012). Are we making a better world with ICTs? Reflections on a future agenda for the IS field. *Journal of Information Technology, 27*(2), 87-93.

Walsham, G., & Sahay, S. (2006). Research on information systems in developing countries: Current landscape and future prospects. *Information Technology for Development, 12,* 7–24. https://doi.org/10.1002/itdj.20020

Walsham, G., Robey, D., & Sahay, S. (2007). Foreword: Special issue on information systems in developing countries. *MIS Quarterly, 31*(2), 317–326.

Weinberger, D. (2010, February 2). The problem with the data-information-knowledge-wisdom hierarchy. [Blog post]. Retrieved from https://hbr.org/2010/02/data-is-to-info-as-info-is-not

Wenar, L. (2006). Accountability in international development aid. *Ethics & International Affairs, 20*(1), 1-23.

White, H. (2009). *Theory-based impact evaluation principles and practice*. International Initiative for Impact Evaluation Working paper 3. New Delhi, India: 3ie.

World Health Organization (WHO) (2010). *Indicators for assessing infant and young child feeding practices part 2: measurement*. Geneva, Switzerland: World Health Organization Press.

Whyte, F. W. (1991). Participatory action research. Sage.

Wiig, K. M. (1993). Knowledge management foundations: Thinking about thinking: how people and organizations create, represent, and use knowledge. Arlington: Schema Press.

Wild, L. & Andrews, M. (2016). *Doing Development Differently: Who we are, What we're doing, What we're learning*. Retrieved from https://www.odi.org/sites/odi.org.uk/files/resource-documents/11199.pdf

Willcocks, L. P. & Lioliou, E (2011). Everything is dangerous: Rethinking Michel Foucault and the social study of ICT. In R. D. Galliers & W. Currie. *The Oxford Handbook of Management Information Systems: Critical Perspectives and New Directions,* Oxford University Press.

Willcocks, L.P. (2004). Foucault, power/knowledge and IS: Reconstructing the present. In J. Mingers & L. Willcocks (eds), *Social Theory and Philosophy for Information Systems*, 238–296. Chichester: Wiley.

Williams, D. (2014). Models, metaphors and symbols for information and knowledge systems. *Journal of Entrepreneurship, Management and Innovation, 10*(1), 79–108. Retrieved from https://www.academia.edu/8243955/Models_Metaphors_and_Symbols_for_Information_and_Knowledge_Systems

Williams, R., & Pollock, N. (2012). Moving beyond the single site implementation study: how (and why) we should study the biography of packaged enterprise solutions. *Information Systems Research, 23*(1), 1-22. DOI: 10.1287/isre.1110.0352

Williamson, J. (1990). What Washington means by policy reform. In J. Williamson, (Ed.), *Latin American Adjustment: How Much Has Happened?, (*pp. 7-20). Institute for International Economics.

Wilson III, E. J. & Best, M. L. (2004). Creating necessary knowledge. *Information Technologies & International Development, 2*(2), pp-iii.

Wittgenstein, L. (1968). *Philosophical investigations*. English text, 3rd edition. Oxford: Blackwell.

World Bank (2016a). *Indicators*. Retrieved from https://data.worldbank.org/indicator/?tab=all

World Bank (2016b). *The World Bank Atlas method – detailed methodology*. Retrieved from https://datahelpdesk.worldbank.org/knowledgebase/articles/378832-what-is-the-world-bank-atlas-method

Worthen, H. (2008). Using activity theory to understand how people learn to negotiate the conditions of work. *Mind, Culture, and Activity, 15*(4), 322-338.

Yanguas, P. (2018). Why we lie about aid: Development and the messy politics of change. London: Zed Books.

Yin, R.K. (2003). *Case study research: Design and methods* (3rd edition.). London: Sage.

Yin, R.K. (2009). *Case study research: Design and research* (4th edition.). London: Sage.

Zahedi, M. & Tessier, V. (2018). Designerly activity theory: toward a new ontology for design research. *Proceedings of Design Research Society*, 25-28/6/2018, University of Limerick, Ireland.

Zeleny, M. (1987). Management support systems: Towards integrated knowledge management. *Human Systems Management,* 7(1): 59–70.

Zuboff, S. (1984). In the age of the smart machine: The future of work and power. New York: Basic Books.

Appendices

Appendix 1: Participant Information Form

This document was used to a provide overview information for participating organisations.

School of Computing and Communications

LANCASTER UNIVERSITY

Project Overview
"Impact Assessment and Service Evaluation" in the 3rd Sector

Paul Kelly, a doctoral research student from the HighWire Doctoral Training Centre at Lancaster University is working on this project in order to understand how technologies and organisational practices contribute to impact assessment and evaluation. The researcher will talk to people from a variety of organisations working with Non-Government and/or third sector organisations (eg NGOs, NPOs).

The research aims to understand how organisational practices, technologies and these two mixed together, create and sustain ways of designing and implementing impact assessments and service evaluations. How do the technologies frame impact? What do different people and institutions require from the assessment/evaluation activities in order to further their own organisational aims and success? And how can explicit evaluation knowledge contribute to the stakeholders practices and relationships across the NGOs, NPOs and their associated network of partner organisations?

Likely outputs include academic publications, press releases and dissemination of findings to organisations with a focus on impact assessment, service evaluation, international development and capacity building.

Research activities such as interviews, workshops, observations and co-working may be recorded in written notes, pictures, audio or video. All data and information provided will be confidential, anonymous and stored securely. All names, work related details such as individual names, organisation names, project names etc will be replaced with pseudonyms so as not to identify the organisation or the people involved.

Participation is voluntary and participants may withdraw from the study at any time, by contacting the researcher and stating their wish to withdraw. If participants withdraw within 2 weeks of a research activity, such as an interview or other kind of participation, their data will be destroyed and not used. After this point their data will remain in the study.

If you have any further questions, please do not hesitate to contact me directly at p.kelly@lancaster.co.uk. In addition, for further issues or suggestions please contact the HighWire Doctoral Training Centre Director, Prof Gordon Blair gordon@comp.lancs.ac.uk, or this specific project supervisor Dr Niall Hayes n.hayes@lancaster.ac.uk.

Thank you for your interest and participation.

Mr Paul R Kelly
HighWire Doctoral Training Centre
Lancaster University, UK
p.kelly@lancaster.co.uk

HighWire Doctoral Training Centre, LICA Building, Lancaster University, LA1 4Y.

Appendix 2: Participation Agreement Form

This form documented research collaboration consent with participating organisations.

Research Collaboration

Development Impact Evaluation Research Collaboration

We, researcher and participant organisation, agree to work on a research collaboration to understand impact evaluation in the international development sector. We agree:

1) the research will support the evaluation work of the organisation.

2) participants may contribute, or comment off record, by telling the researcher at any time

3) the researcher will keep in touch with the organisation, share progress, data, analysis, findings and continue to support the organisation's collaboration and interest.

4) the researcher may make notes, audio record meetings to support the later analysis. Only the organisation / researcher will have access to these.

5) Any data will be stored and encrypted on a password protected secure computer, and will be deleted from any portable devices.

6) Any research plans, materials or data may be used for the organisation's work e.g. surveys, projects, reports, funding applications etc. and for the researcher's work e.g. academic conferences, journal papers.

7) all names of people, places, organisations, projects etc will be deleted/anonymized to secure the privacy of all participants

Further collaborations, plans & agreements can be added verbally, in writing or by email if it supports the interests of both the organisation and the researcher.

Researcher

Name: Mr Paul R Kelly   Signature: _____

Organisation/Role: Lancaster University, PhD researcher in impact evaluation

Organisation

Name: M▮▮▮▮E   Signature: ▮▮▮▮

Organisation/Role: The ▮▮▮▮▮▮ation – Project Facilitator

Appendix 3: Case 1 List of Empirical Data and Interactions

Participants from Rural Indian and The Imagine Foundation were:

- Vijay, Imagine Founder & CEO ("V" in Appendix 3 Table 1 below)
- Leonard, Imagine Project Manager ("L")
- Chandan, Rural India Director ("Ch")
- Paul, doctoral researcher ("P")
- Nigel, researcher supervisor ("N")

Appendix 3 Table 1 below details the specific research interactions and shared documents.

Interactions	Participants	Details and focus
Phone call 1 (45 mins)	L, V, N	Initial contact discussions
Group call 1 (60 mins)	L, N, P	Collaboration background & scoping
Group call 2 (90 mins)	VJ, L, N	Previous evaluations, 2013-14 plans
Group call 3 (90 mins)	L, N, P	Evaluation tools, survey, scripts
Group call 4 (120 mins)	L, N, P	Evaluation tools, survey, scripts
Group call 5 (120 mins)	L, Ch, N, P	Review of tool, implementation plans
Group call 6 (90 mins)	L, Ch, N, P	Implementation plans
Phone call 2 (20 mins)	L, P	Leonard's role transfer
NGO Spreadsheet 1	NGO staff	Mustard crop practices
NGO Spreadsheet 2	NGO staff	Mustard crop costs / value
NGO Spreadsheet 3	NGO staff	Animal husbandry practices
NGO Spreadsheet 4	NGO staff	Cotton crop practices

NGO Spreadsheet 5	NGO staff	Deep ploughing practices
NGO Spreadsheet 6	NGO staff	Vegetable crop practices
NGO Spreadsheet 7	NGO staff	Farmer data
NGO Spreadsheet 8	NGO staff	Farmer data
NGO Spreadsheet 9	NGO staff	Farmer data
NGO Spreadsheet 10	NGO staff	Measurements matrix
Doc 1: NGO Report 1	NGO staff	Narrative of 4 farmers
Doc 2: NGO Report 2	NGO staff	Annual progress report, 2013
Doc 3: Survey draft	NGO staff	Survey draft with discrete questions
Doc 4: Indicators	NGO staff	Indicator draft plans
Doc 5: Survey tool	Ch, L, P	Survey questions (co-designed)
Doc 6: Interview script	Ch, L, P	Qualitative interview script (co-designed)
Doc 7: ICT survey	Ch, L, P	ICT feedback survey (co-designed)
Imagine Website	Imagine Staff	Homepage, About, Project pages etc
Researcher notes	P	Notes taken during calls & tool design
Email exchanges	L, Ch, P	50+ messages

Appendix 3 Table 1: Farmer Livelihoods Program (August 2013 to April 2014)

Appendix 4: Case 2 List of Empirical Data and Interactions

List of key participants at HSTSG Thailand who took part in NGO case 2 activities.

1. Khun, director & founder of NGO ("K" in Appendix 4 Tables 1-4 below)
2. Tik, field operations manager ("T")
3. Susan, donor liaison & US intern ("S")
4. Orr, finance officer, and later translator ("O")
5. Ajarn, board member and university professor ("A")
6. Noi, secretary ("N")
7. Mai, fund manager & US intern ("Mai")
8. Mee, Healthcare project coordinator, ("Mee")
9. Pang, Healthcare project manager ("Pg")
10. Sam, Healthcare project administrator ("S")
11. Cherry, evaluation specialist & US intern ("Ch")
12. Mia, Migrant project writer/ coordinator / US intern ("Mia")
13. Chai, Thai/Karen staff member, driver and later translator
14. Paul, researcher ("P")

Appendix 4 Tables 1-4 below detail the specific research interactions and shared documents.

Type of Data	Participants	Details and focus
Emails	O, S, K, P	Initial contact, planning for field trip 1
Document 1	Staff	Healthcare Project Report (2008)
Document 2	Staff	Healthcare Project Report (2009)
Document 3	Staff	Healthcare Project Report (2010)
Document 4	Staff	Healthcare Project Report (2011)
Document 5	K, O, staff	Healthcare Project Report (2012)

Document 6	K, O, staff	Healthcare Project Report (December, 2013)
Document 7	K, S, staff	Healthcare Project Report (2014)
Document 8	S, K	Healthcare Project Report (2015)
Document 9	S, K	Healthcare Project continuation & funding proposal (2014)
Document 10	S, photographer	Village photo-essay
Document 11	K	Presentation on HTSG history, aims, projects
Document 12	Consultant	Voluntary evaluation & recommendations (2012)
Document 13	Staff	Healthcare outcome summaries
Document 14	Staff	Healthcare participant profiles
Document 15	Staff	Healthcare interview records (4 stakeholders & images)
Spreadsheet 1	Mee	Healthcare family home visits – 116 participants (2013)
Spreadsheet 2	Mee	Healthcare family home visits – 332 participants (2014)
Spreadsheet 3	Mee	Healthcare family home visits – 308 participants (2014)
Spreadsheet 4	Mee	Healthcare family home visits – 105 participants (2013)
Field trip 1		
Conversation 1	K, P	Car ride & refreshments (90 mins)
Meeting 1	K, S, P	Introduction to projects and teams (60 mins)
Meeting 2	P, S, K, A, T	Formal presentation & discussion (90 mins)
Meeting 3	K, S, P	Informal lunch (60 mins)

Meeting 4	K, S, P	Collaboration plans (45 mins)
Conversation 2	T, P	Healthcare project fieldwork (45 mins)
Conversation 3	T, P	Healthcare project fieldwork (30 mins)
Meeting 5	S, K, Aj, O, P	Funding & evaluation (120 mins)
Meeting 6	S, O, P	Funding applications, government, NGO activities (30 mins)
Conversation 4	Mee, P	Healthcare project field work (30 mins)
Meeting 7	K, Aj, S, T, P	HTSG partners & using evaluation data/knowledge (90 mins)
Conversation 5	S, P	S's role, plans, work (60 mins)
Conversation 6	K, P	HTSG projects & aims (60 mins)
Group meeting 8	K, S, Aj, P	Further collaboration & visits (60 mins)
Conversation 7	T, P	On volunteers, villages (30 mins)
Document 16	Mai, Mee	Summary of Healthcare project New Villages survey data
Document 17	P	Analysis of Healthcare Project New Villages data

Appendix 4 Table 1: Phase 1 Healthcare Project and field trip 1 (May 2014 - Sept 2014)

Type of data	Participants	Details / focus
Email	S, O, K, P	Planning for field trip 2, document exchanges
Document 18	P	Collaboration proposal
Field trip 2		
Conversation 1	M, P	Greetings, family, daily work (30 mins)
Conversation 2	O, P	Orr's role change, UN trainings (30 mins)
Meeting 1	T, O, Mee, B, Sa, K, Aj, P	Healthcare project update (90 mins)
Workshop 1	T, O, B, Sa, Aj, S, Mai, Mee, K, Pg	On evaluation stakeholder networks (90 mins)
Workshop 2	T, O, B, Sa, Aj, S, Mai, Mee, K, Pg	Presentation & discussion (90 mins)
Workshop 3	T, O, Sa, Mai, Mee, B, S, P	Evaluation interviews/roleplay (120 mins)
Meeting 2	S, K, P	Lunch; ID cards, hospital births (60 mins)
Workshop 4	T, O, Sa, Mai, Mee, B, S, Pg	Planning: stakeholder questions, roleplaying (120 mins)
Workshop 5	T, O, Sa, Mai, Mee, S, Pg	Timeline, questions (Thai/English), interview guide (120 mins)
Workshop 6	T, O, Sa, Mai, Mee, B, S, Pg	Impact for reports, media, partners, government (120 mins)
Meeting 3	T, O, Sa, Mai, Mee, B, S, K, Pg	Wrap-up, Q&A, stakeholder needs (120 mins)
Conversation 3	K, P	House-fire incident, local conference (60 mins)
Meeting 4	O, S, P	Translation & reporting, O's role, S's plans (60 mins)

Document 19	Mee, B	Interview scripts for stakeholder groups: (co-designed)
Document 20	O, P	Evaluation activities & timelines (co-designed)
Document 21	O, B, P	Interview guide re: privacy contact information (co-designed)
Document 22	K	House-fire incident: response, photograph scans
Document 23	P	Workshop presentation & plans
Audio files	P	4 workshop recordings
Photographs	P	43 visit photographs

Appendix 4 Table 2: Phase 2 Healthcare project and field trip 2 (Sept 2014 - May 2015)

Type of data	Participants	Details and focus
Emails	Ch, Mia, P, K	Arrangements, requests, responses (often multi-page)
Group call 1	Ch, Mia, P	Healthcare evaluation; Migrant Project bid; Skype (90 mins)
Document 24	Ch	Topics/questions on daycare, nutrition & dental hygiene
Document 25	P	Feedback on draft survey
Document 26	Ch	Pilot interview analysis (11 respondents)
Audio files	Ch	Recorded samples of day-carer interviews

Appendix 4 Table 3: Phase 3 Healthcare Project Extension (October 2015 - January 2016)

Type of data	Participants	Details and focus
Email exchanges	Mia, Ch, K, P	Requests, responses, file exchange, evaluation plans for Migrant project bid
Group call	Ch, K, Mia, P	Migrant project proposal, evaluation & design; Skype (60 mins)
Document 27	Ch, Mia, K	Draft Migrant project proposal
Document 28	P	Proposal advice on evaluation indicators, activities & services
Spreadsheet 5	Mia	Draft Migrant project log-frame plan
Spreadsheet 6	P	Log-frame re-draft with advice on aims, indicators, outcomes
Document 29	Ch, Mia, K, P	Final Migrant project proposal (co-designed)

Document 30	Mia	Migrant project survey (co-designed) & pilot data (9 respondents)
Document 31	P	Feedback on baseline pilot survey
Spreadsheet 7	Mia	Data sample (75 survey responses), 127 discrete question items
Document 32	P	Data analysis & report for Migrant survey data
Document 33	K, Mia	Invitation to "Child Statelessness" event (11/2016), including evaluation interim results & child attendance & stories

Appendix 4 Table 4: Phase 4 Migrant Project funding bid & activities (October 2015 - December 2016)

Appendix 5: Sample list of TIEK approaches and methods

Below is an illustrative list of 96 evaluation approaches or methods. The term "*approach*" indicates *"an integrated set of options used to do some or all of the tasks involved in evaluation"* (Better Evaluation, 2016). The list below is evidence of the expanding diversity of TIEK, particularly since the 1990s. It demonstrates the breadth of methodological knowledge. The list is illustrative, not exhaustive.

The list does not categorise individual methods as there are many debates about classifications, for example whether qualitative/quantitative labels relate to methods or data. Garbarino & Holland (2009) follow Hentschel's method-data framework (1999) in relating these labels to data, although many "participatory" approaches concern how to engage with respondents. Participatory statistics is a boundary crossing example (see Holland, 2013). Appendix 5 Table 1 below is used to indicate the increasing array of technical approaches, and opportunities for confusion.

The list is a limited version of a more comprehensive list collated in 2016. It is based on diverse publications, primarily from the following sources:

1) Better Evaluation web resource (supported by ODI, ALNAP, RMIT, IDRC, IFAD, Rockefeller Foundation);
2) MandEonline web resource (supported by Oxfam UK, Action Aid, Save the Children, Christian Aid, CAFOD, IDRC);
3) The World Bank, UN, CGIAR, ODI publications;
4) Academic articles from "Evaluation", "World Development", and "Journal of Development Studies";
5) Work by specialists Michael Patton, Bob Picciotto, Chris Roche, Richard Heeks e.g. Heeks and Molla (2009).

List of approaches and methods	
Anecdotes and anecdata	Kernel matching method
Appreciative inquiry	Knowledge attitude practice behaviour (KAPB)

Before and after designs	Learning for action
Benchmarking	Learning-oriented evaluation
Beneficiary assessment	Livelihoods framework
Binary probit regression	Log-frame analysis
Capabilities framework	Measuring distributional program effects
Capacity development results framework	Medical models e.g. case-controls
Case studies	Meta-evaluations
Cluster evaluations	Most significant change
Collaborative evaluation	Multi-attribute utility measurements
Collaborative outcomes reporting	Multivariate regression analysis
Communications for development	Multivariate statistical modelling
Comparative post-test design	Nearest neighbour method
Comparison of means	Network evaluation
Contribution analysis	Outcome expectations
Control groups	Outcome harvesting
Cost benefit analysis (CBA)	Outcome mapping
Counterfactuals analysis	Participatory appraisals
Critical stories of change	Participatory data
Critical system heuristics	Participatory evaluation
Crowd sourced evaluation	Participatory rural appraisal

Cultural institutional framework	Participatory statistics
Delphic surveys	Pipeline method for comparisons
Democratic evaluation	Positive deviance
Developmental evaluation	Project goals analysis
Discontinuity matching	Propensity score matching (PSM)
Do No Harm	Quasi experimental evaluations
Double-difference / Difference in difference	Rainbow framework
Dreams realised, or Visioning	Random comparison group
Economic models (structural & reduced)	Randomised control trials (RCTs)
Empowerment evaluation	Rapid outcome assessment
Enterprise Variable/Relation/Value Chain	Rapid rural appraisal
Equity focused evaluation	Real time evaluation
Evidence principles	Realist evaluation
Ex post facto evaluation studies	Regression frameworks / analysis
Experimental designs	Responsive evaluation
Expert panels	Robust decision analysis
Factor analysis	Shoestring evaluation
Feminist evaluation	Social return on investment
General elimination methodology	Sub-objectives as design
Horizontal evaluation	Theories of change

Information economics	Theory based evaluation
Information needs mapping	Time series design - comparative
Innovation history	Time series design - interrupted
Institutional histories	Tracer Studies
Instrumental variable method	Utilisation focused evaluation
Instrumental variables	Value for money

Appendix 5 Table 1: Sample list of diverse approaches/methods used in development impact evaluation

Appendix 6: Case 1 Quantitative Survey draft questions

Researcher and case participant interactions centred on co-creating evaluation materials, such as this Quantitative Survey draft set of questions.

QUANTITATIVE SURVEY Draft 1.2

1 Quantitative data is best used with accompanying qualitative data (see qualitative survey draft). Quantitative data is best used with qualitative data, to explain the how and why of the quantitative data. For example:

- Beneficiary may have good income, but never see it because of family control issues around finance;
- Beneficiary may have bank account, but is confused and never uses it;
- Beneficiary's children may complete schooling, but not pass any exams;
- Beneficiary may live near a hospital, but can't afford to use it;
- Beneficiary may have large seasonal fluctuation in monthly earnings;
- Beneficiary may live near roads, but never leaves village

Questions

FARMING

1 What kind of farming do you do?
- ☐ Animal Husbandry
- ☐ Agriculture
- ☐ Other _____

2 What kind of farm animals do you own? And how many? _____

3 What kind of crops do you farm? _____

4 What area of land do you
- own _____
- rent _____

5 How do you get your water supply? _____

FINANCES

1 Do you have excess farm produce to sell each month? Normally   Sometimes   Rarely

2 What are your main monthly incomes?

Source 1	Rs _____
Source 2	Rs _____
Source 3	Rs _____
Source 4	Rs _____
Source 5	Rs _____
TOTAL	
Comments	

3 What are your main monthly outgoings?

Cost 1	Rs _____
Cost 2	Rs _____
Cost 3	Rs _____
Cost 4	Rs _____
Cost 5	Rs _____
TOTAL	
Comments	

4 Do you have enough money? Yes / No

5 What would be enough? _____

6 Do you have a bank account? Yes / No

7 Do you keep savings in it? Yes / No

EDUCATION

1 Did you or your partner finish high school? Y/N

2 At what age did your children leave school/college?

Child 1 _____	Child 2 _____
Child 3 _____	Child 4 _____
Child 5 _____	Child 6 _____

Comments

3 Can your family afford all school fees/books/travel/costs? Yes / No

4 Have any of your children dropped out of school? Yes / No   Why?

HEALTH

1 How close is the nearest doctor?

2 How close is the nearest hospital?

3 Is it easy for you to travel to the doctor / hospital?

4 Have you had any serious health issues in your family in the last year? What issues?

5 Did you have problems accessing healthcare e.g. cost, distance, too busy?

UTILITIES

Please check Yes or No if you have the following facilities / utilities.

1 Drinking water on tap in home Y / N
2 Drinking water <500m from home Y / N
3 Toilet facilities in your home Y / N
4 Water well at your home Y / N
5 A Water pump at your home Y / N
6 Electric power to your home Y / N
7 Electric lighting in your home Y / N
8 A television in your home Y / N
9 An internet connection at home Y / N
10 A personal mobile phone Y / N
11 A computer at home Y / N
12 A family motorcycle Y / N
13 A family car Y / N
14 A mosquito net for whole family Y / N

ACCESS

Please check Yes or No if you have easy access to the following facilities.

(Easy access means you can walk there within 1 hour, or use your vehicle or public transport to get there within 1 hour)

1 A bus stop Y / N
2 A large bus station Y / N
3 A train stop Y / N
4 A train station Y / N
5 An approach road into your village Y / N
6 Road from village to nearest m'and Y / N
7 A bank Y / N
8 Your bank Y / N
9 Your children's school Y / N
10 Your work place e.g. farmland Y / N
11 A vet Y / N
12 An NGO office Y / N
13 An agriculture expert Y / N

14 How many hours away is the nearest town? Hours _____ by walking / vehicle / train

15 How many hours away is the nearest city? Hours _____ by walking / vehicle / train

Appendix 7: Case 1 Qualitative interview draft questions

Researcher and case participant interactions centred on co-creating evaluation materials, such as this Qualitative Interview question set.

Qualitative Interview Questions Draft 1.3

1. 1 Qualitative interviews should be used with a sampling strategy. At this stage 25 short 45 minutes to 1 hour interviews is sufficient to inform the evaluation and triangulation. Not all households require a qualitative survey.
2. 2 Survey volunteers should be trained to do interviews. Critically, asking a question and gently encouraging an interviewee response is recommended. Go slowly.
3. 3 Without rich context qualitative data, to support the quantitative data, it may not be clear how programs impact lives, families, villages and livelihoods of development. Qualitative and quantitative data support the triangulation and validity of impact assessment across all stakeholders.
4. 4 Interviewers don't have to ask all questions, just stimulate conversation about the topics. Interviewers can skips questions and / or sections. Time advice is for guidance only. For example, you may skip education if applicant has no children.
5. 6 Interviews should ideally be recorded on the mobile device and downloaded to an office computer. Then they should be stored on a secure server with Name/Date/Location in the filename/Identifier

Time	Topic and questions
5 mins	**Introductions** • Introductions, Name, WG you work with, How long with WG • Explain purpose of short interview to understand how to make programs and support better in future • Explain that all answers are confidential and will be anonymous.
5-10 mins	**Part 1: About your home** • Is your home sufficient for your family needs? • What else do you need? e.g. more room, better location, toilet etc? • What 3 things would you change if you could? And why? • What do you need to achieve this? e.g. support, knowledge etc? • Any further important issues about your home?
5-10 mins	**Part 2: About facilities** • Do you have the facilities you need e.g. water, electricity, car, tv, internet, phone etc? Why? • Do you have good access to local towns, work, school, hospitals, banks? Good roads, public transport? Why? • What 3 things would you change/get if you could? Why? • What do you need to achieve this? e.g. money, knowledge etc? • Any further important issues about utilities/facilities?
5-10 mins	**Part 3: About education** • Do you children enjoy school? Did/will any drop out, fail? • Would improve education e.g. books, meeting teachers more etc? • What 3 things would you change if you could? And why? • What do you need to achieve this? e.g. time, money etc? • Any further important issues about your children's education?
5-10 mins	**Part 4: About healthcare** • Are you satisfied with current access to healthcare? Why/why not? • Can you describe recent healthcare incidents? What happened? • What 3 things would you change if you could? Why? • What would you need to achieve this e.g. closer doctors, money etc? • Any further important issues about your healthcare access?
5-10 mins	**Part 5: About your work** • What farming do you do? Animal husbandry, agriculture? Do you have good access to markets/transport? Which ones? • What 3 things would you change if you could? Why? • What would you need to do this? e.g. contacts, vet advice? • Any further important issues about your work?
5-10 mins	**Part 6: About finance and money** • Do you have a bank account? How do you manage money? Loans? • Do you use banks regularly? e.g. once a month? Have savings? • What top 3 things would you change if you could? Why? • What would you need to do this? More time, know about banks ...? • Any further important issues about your money/finances?
5 mins	**Close interview** • Express gratitude for interviewee's time; • Tell them this will contribute to our understanding of their livelihoods and hopes for the future, and improve programs • Thank you, Contact Information

Appendix 8: Case 2 Post Workshops Evaluation Plan

Researcher and case participant interactions centred on co-creating evaluation materials. Appendix 8 Figures 1 and 2 below are excerpts from the 2015 Evaluation Plan developed after the workshop series. Names of districts, people etc are blacked out for anonymity purposes.
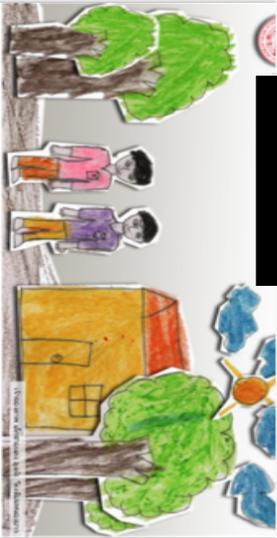
Project Impact Evaluation Interview Guide

Project '... Child Care and Development' modality in disadvantaged sub-districts communities of remote areas in Upper Northern Thailand to Ensuring Holistic Development for very Young Children"

Project areas: M... a sub-district ... rict in M... on Province.

Introductory Message

Thank you for taking part in our project's evaluation research. This project's target areas are Muang ... Sa sub-district, ... Province. Our aim is to collect information on aspects of your life or work that are important for us to learn about. This research is part of evaluating our own work at ... and your help is much appreciated. If you wish to withdraw from the research at any time, please let us know. For further information please contact: Ms. Q... Assistant director and project coordinator at 0... 57.

Confidentiality

The information or stories of this research that you share with us are strictly confidential. We will not use your name, address or phone number to share in any of our reports or future publications. We assure you that your participation is completely anonymous.

Group	Method	District	Time	Translator	When?
• Tambon Authority (TAO)	Interview	• P... 5-10 • Muan... 5-10	30-45 mins /person	-	April – May 2015
• Local Hospital					
• Local Education Authority					
• Expectant mothers	Interview	• Pla... 5-10 • Muan... 5-10	1-2 hour / person	Translator needed	April – May 2015
• Post pregnancy					
• Families with 0-2 year-old-children					
• Parenting Volunteer	Focus group	• Muan... 5-10	2 hours / Group	-	April – May 2015
• Pre – School teacher	Focus group	• Post...10 • Muan... 5-10	2 hours / Group	-	April – May 2015
• Day care teacher	Focus group	• 4-5 ... staff*	2 hours	-	April – May 2015
• ... staff*					

2

Appendix 8 Figure 1: Excerpt from 2015 co-designed evaluation plan

Questions for Interviews / Focus Groups

Local Authorities:
- Tambon Authority (TAO)
- Local Hospital
- Local Education Authority

1. Information of demographics and birth rates in sub-district
2. What is the health plan development for sub-district villagers in terms of taking care of pregnant/post-natal women? Are there any specialists on women's health?
3. Are there other NGOs working on the issue of early childhood care and women's health in this area besides ▇
4. What is the role of Local organizations (TAO, hospitals, and schools) in supporting NGOs or organizations working in this sub-district? What are the results?
6. Do you have any plans in the near future for early childhood care policy?

Parenting Volunteers

1. What is your role in your community?
2. What motivated you to become a Parenting Volunteer?
3. How do you feel when you support your community as a Parenting Volunteer?
4. How do you feel when you provide care and suggestions to expectant mothers and mothers with children ages 0-2 years, and why?
5. What is the most challenging aspect of working as a Parenting Volunteer?
6. Do villagers know that you are a Parenting Volunteer?
7. What will you do after the ▇ project has ended? Will you continue working with early childhood care issues?
8. Do you have any suggestions for better implementation of ▇ project?

Expectant mothers, post-natal mothers, and families with children ages 0-2 years

1. General introduction/conversation starter – How are you?
2. For pregnant mothers: How old are you? When is your baby due?
3. Who is taking you to see the doctor?
4. Why/when do you visit the doctor? / Why have you decided not to visit the doctor?
5. How do you take care of yourself while you are pregnant / after the birth?
6. Has anyone been providing you with information on prenatal care?
7. Do you have a family plan on how to take care of your baby?

8. Have you ever left your village? Why / why not?
10. If you have some issues regarding prenatal care, who do you talk to or how do you handle the situation?
11. Do you know any Parenting Volunteers in the community? What is their role?
12. How many times this year have you received home visits from Parenting Volunteers? What do you think of the home visits?
13. Are there any organizations beside ▇ that provide you with support?
15. What would you do to increase the quality of your family's life?
16. Do you know ▇ or its staff members?
17. Do you have any suggestion for better implementation of ▇ 's project?

Pre-school teachers and Day-care centre teachers

1. How long have you been working as a teacher?
2. How are the kids in general?
3. What do you think of their early childhood development?
4. What knowledge or skills on early childhood care would you like to increase/learn more about?
5. Are there any organizations or government sectors that provide support to your day-care centre/ school? If so, what kind of support is provided?
7. Do you think the support that _____ has been providing is helpful, and if so, why?
8. What type of support do you need the most?

ECCD staff

1. Would you please give me an overall picture of ▇ ?
2. Are you proud to work for the ▇ project?
3. What have you learnt from the project, and how can you apply the knowledge gained from your working experience to your daily life?
4. Do you think the project is benefiting the project areas?
5. What are indicators of the project's sustainability that you have seen?
6. Where are the levels of relevant agencies' participation?
7. Who has a key role in supporting the project?
8. What are the successes of the project?
9. Do you have any suggestion for better implementation of the ▇ project?

Dental Care/Malnutrition-Related Questions (Draft by S▇)

1. General Introduction (How are you, etc.)
2. How many meals do you and your children eat per day?
3. What does an average meal look like in your home? (i.e. what are the main types of food that you eat?)
4. Do you usually cook yourself or do you buy food from elsewhere?
5. Did you/are you breastfeeding your children?
6. In general, how willing are your children to eat fruits?
7. In general, how willing are your children to eat vegetables?
8. Do your children lead an active lifestyle (i.e. do they run around and play with other children, or would they rather sit at home and play video games)?
9. Do you and your children own toothbrushes and toothpaste?
10. How often do you and your children brush your teeth? / Is tooth-brushing a part of your children's daily routine?
11. When was the last time you saw a dentist?
12. Have you ever taken your children to see a dentist/dental care expert?
13. Have you or your children ever reported feeling pain in your teeth/gums/mouth?
14. How highly do you value dental hygiene? I.e. Do you think about it a lot or a little?

Appendix 8 Figure 2: Excerpt from 2015 co-designed evaluation plan