THE UNIVERSITY of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

Accounting for Crisis:
The Power of Ambiguity in the Management of Humanitarian Emergencies

Marian Konstantin Gatzweiler

Doctor of Philosophy

The University of Edinburgh

2017

Declaration

This is to certify that the work contained within has been composed by me and is entirely my own work. No part of this thesis has been submitted for any other degree or professional qualification.

Signed:

Marian Konstantin Gatzweiler

Abstract

A defining feature of humanitarian crises is their unpredictable nature, making them interesting sites to analyse how accounting systems can facilitate engagement with the unexpected. This thesis explores the question of how evaluation systems can be designed and practiced to engage with the complexities of humanitarian crisis settings, in which the potential for disastrous errors is overwhelming. Informed by empirical research on the management practices in a large-scale refugee camp, the study investigates principles and tactics that allow humanitarian evaluation systems to make a resource of the inevitable ambiguity and incompleteness that define their contexts. In doing so, the thesis draws from and further develops the concept of heterarchy, defined as 'governance through difference', and shows how it provides promising insights for accounting research. To explain how evaluation systems can become performable in the dynamic humanitarian environments, the study theorizes four interlinked principles that emerge from the empirical findings. These principles are: (1) in-built tensions between evaluation dimensions; (2) open and participatory design; (3) relational value and incompleteness; and (4) enacting minimalist control through a community of practitioners. In doing so, the study makes three contributions. Firstly, the study contributes to the accounting literature on the enabling role of ambiguity by theorizing *how* evaluation systems can foster approaches and techniques that embrace ambiguity as a resource to engage with complex settings. Secondly, it further develops the notion of heterarchy by explicating how heterarchical tensions can become productive without leading to chaos and by theorizing additional principles that are necessary to sustain heterarchies in an organized fashion. Thirdly, departing from the emerging

literature on humanitarian crises that primarily focuses on how accounting systems can be used to normalize and control disaster settings, the thesis advances understanding of how accounting technologies can serve as anomalizing devices for the adaptive management of crises.

Lay Summary:

This thesis explores how performance evaluation systems can be designed and practiced to enable reflective engagement with unstable and unpredictable settings such as humanitarian crises. While evaluation systems are commonly developed with the aim to align pre-specified objectives with actions and measures for 'correct' performance, the thesis shows that it is impossible to design a system that can cope with dynamic humanitarian environments without producing a multitude of blind spots. If evaluation systems draw attention away from such blind spots, a resulting concern is the increased prospect that these systems actively prevent detection of emerging issues to the point that they become impossible to contain. Therefore, rather than aspiring to an impossible ideal of precise measurement, the thesis takes seriously the incompleteness and ambiguities of evaluation systems in such settings and attempts to develop ways to pro-actively embrace them to avoid oversimplification and confirmation seeking. In other words, the thesis explores tactics and principles that allow humanitarian evaluation systems to reflectively engage with the ambiguity of crisis settings. Drawing from empirical research in a large-scale refugee camp, the study theorizes four such principles. In doing so, the study contributes to knowledge by outlining how evaluation systems can mobilize ambiguity in a strategic manner to detect anomalies and facilitate engagement with the unpredictability of humanitarian crisis settings.

Dedication

Table of Contents

List of Tables:

List of Figures:

1. Introduction

1.1. <u>Introduction</u>

> Today's humanitarian emergencies are beyond anything we have experienced in living memory. Unless we urgently change the way we tackle them, the world will become an increasingly difficult place to live for millions of people who have already lost almost everything.

<div align="right">

Antonio Gueteres (2015), United Nations High

Commissioner for Refugees

</div>

Interest in the role of accounting systems in anarchistic and unstable environments, in which clarity of goals and cause-and-effect relationships cannot be taken for granted, dates back to the beginning of interdisciplinary accounting research (Cooper, Hayes & Wolf, 1981). Since then, critical studies have begun to explore the role of accounting systems in situations of uncertainty beyond the post-hoc rationalization of organizational actions, towards a more proactive outlook (Burchell, Clubb, Hopwood, Hughes & Nahapiet, 1980; Hopwood, 1990, 2005; Mouritsen, 2016). Several studies have drawn attention to the enabling role of incomplete measures (Dambrin & Robson, 2011; Jordan & Messner, 2012; Quattrone, 2015a) and the power of ambiguity in strengthening accounting systems in "not just communication of what is known but transformation of what is knowable" (March, 1987, p.165).

These studies raise questions about the potential of ambiguity to engage with unpredictable and unstable settings, such as humanitarian crises. However, theoretical understanding of the question of *how* designers might develop accounting systems that productively embed ambiguity to foster their evocative capacities remains underdeveloped. In line with these concerns, drawing from empirical research on performance evaluation practices in a large-scale humanitarian crisis, this thesis investigates principles and tactics that allow evaluation systems to make a resource of the inevitable ambiguity and incompleteness that define this context.

Despite recent calls for research and the rising importance of the topic for policy-makers, the role of accounting and evaluation systems in the engagement with extreme situations such as humanitarian crises has so far received relatively little attention in the literature (see Sargiacomo, 2015; Walker, 2014; Taylor, Tharapos & Sideway, 2014). However, in light of the recent refugee crisis as one of the major contemporary socio-political challenges, interest in the role of accounting techniques in the management of such emergencies has increased significantly (Walker, 2016; Vollmer, 2013). Humanitarian crises may seem like unusual settings for accounting scholars given their unpredictable and highly unstable nature. Humanitarian organizations have to operate under harsh conditions, in which the consequences of errors are extreme and not uncommonly result in deaths, the marginalization of vulnerable groups, or the prolonging of conflict and war (see Ramalingam, 2013). When large errors or misconduct occurs, there is commonly widespread media coverage with the ensuing consequence that humanitarian organizations can lose sources of funding, support and public goodwill (Barnett, 2011; Everett and Friessen, 2010). In other words,

humanitarian organizations have to operate in extraordinarily challenging conditions, in which failure to produce desirable results can have wide-ranging consequences.

The primary focus of the emerging accounting research on humanitarian disasters has been on the question of how accounting technologies are mobilized to control and 'normalize' crisis situations. In a recent study, Sargiacomo (2015, p. 68) highlights how government agencies re-asserted control over a humanitarian crisis following a large-scale earthquake by constructing an improvised calculative infrastructure that combines scientific with accounting classifications. Accordingly, by enabling government agencies to act upon the crisis setting from a distance, this extraordinary accounting assemblage shaped the construction and enactment of programs of control (Miller, 1991; Robson, 1992, 1994), permitting a return to more 'normal' forms of governance (see also Sargiacomo, Ianni & Everett, 2014). Other recent research has advanced knowledge in relation to the question of how humanitarian organizations may employ distinct accountability scripts to justify their actions in a post hoc manner and engage with communities in emergency relief operations and development programs (Everett & Friessen, 2010; Taylor, Tharapos & Sideway, 2014; Lai, Leoni & Stacchezzini, 2014; O'Leary, 2016). Combined, these emerging studies have begun to explore concerns in relation to the possible roles of accounting technologies in humanitarian crises by focusing on the question of how accounting can act as a normalizing technology (Jordan, Mitterhofer & Jorgesen, 2016).

While these studies provide important insights, in the extreme environments that commonly define humanitarian practice, the key organizing and sense-making challenges do not necessarily relate to the implementation of predefined programs of control. As emphasized by Weick & Sutcliffe (2015), during crises, control mechanisms are commonly experienced by way of their limitations, incompleteness and failures. Such limitations may take the form of insufficient knowledge of system boundaries (Cooper et al., 1981; Cooper, 1983; Garud, Sanjay & Tuertscher, 2008), ambiguities of choice (March, 1987; Mouritsen, 2016), shifting centers of control (Quattrone & Hopper, 2005; Dechow & Mouritsen, 2005), and uncertainty of funding and budgeting horizons (Barnett, 2011; Baker, 2014). Given that these factors constitute the norm rather than the exception in crisis settings (see Ramalingam, 2013), they provide challenging theoretical and practical problems for humanitarian accounting and performance evaluation systems. In particular, the role of categorizations and reporting mechanisms is critical in the way organizations and decision-makers relate to crises (Hopwood, 1990). That is because accounting categorizations create specific forms of visibility and invisibility (Hoskin & Macve, 1986; Hopwood, 1990; Miller & O'Leary, 2007), which in turn induce blind spots that might miss unexpected events by labelling them too quickly as 'in-family' (Weick & Sutcliffe, 2015). Thus, while the use of accounting categories is unavoidable to make sense of humanitarian crises, they run the danger of locking in expectations that are inappropriate for a response context. In other words, if accounting and performance evaluation systems draw attention away from important elements that are not captured by them, a resulting concern is an increased likelihood that these systems impede the detection of and the engagement with emergent issues (Stark, 2009).

Therefore, instead of focusing on the question of how accounting systems might be implicated in the attempt to normalize disaster situations (Sargiacomo, 2015), the more provocative question is how accounting systems might be designed and practiced to facilitate engagement with anomalies, ambiguity and the unexpected. Focusing on anomalies does not necessarily lead to an abandonment of the concept of control, but instead it favours a more minimalist attitude (Quattrone & Hopper, 2005) that takes seriously the limitations and incompleteness of accounting systems in such contexts (Dambrin & Robson, 2011; Jordan & Messner, 2012; Quattrone, 2015a) and attempts to build on them. In this sense, focusing on engagement with anomalies has the potential to draw attention to possible problems and work against a tendency to overlook emerging issues to the point that they become impossible to contain (Weick & Sutcliffe, 2015).

In line with these concerns, the primary interest of this study does *not* lie on the particular outcomes of particular humanitarian actions, which are widely documented (see Harrell- Bond, 1986; JEEAR, 1996; James, 2008; Ramalingam, 2013), nor does it strictly lie in the humanitarian organizations themselves. Rather the focus of this study is on the specific evaluation techniques and practices that facilitate engagement with the ambiguous and unpredictable nature of humanitarian crises, which remain under-researched and under-theorized. Since crisis settings constitute their everyday operating environment, humanitarians have developed specific performance systems that can offer promising insights for accounting scholarship on how to engage with such unstable and complex settings. It is widely recognized that the capacity to deal with situations of crisis depends to a large extent on the

technologies and routines that were developed before chaos unfolds (see Weick and Sutcliffe, 2007, 2015). In line with these insights, this study explores tactics and principles that enable humanitarians to evaluate performance and plan for actions under conditions of chaos, complexity and uncertainty.

The theoretical interest in processes of evaluation under conditions of uncertainty also provides direct links with recent scholarly interest in the Sociology of Worth (SOW), from which this dissertation draws (see Berthoin-Antal, Hutter & Stark, 2015; Kornberger, Justesen, Mouritsen & Madsen, 2015). Within the sociological discipline, SOW is the specialization that is concerned with theoretical and societal questions of value and pursues several basic but highly significant questions: How are notions of value conceptualized and practiced in modern societies? What kind of mechanisms are needed to ensure that a greater number of societal members are recognized as valuable, including the poor, refugees and other marginalized groups? What is the role of technologies in constituting and measuring value? How can conflicts between different forms of value be mediated and settled (see Karpik, 2010; Stark, 2009; Lamont, 2012)? A variety of research has begun to advance knowledge in relation to several of these issues. Studies have drawn attention to the material features of simple graphs and visualizations in shaping and constituting value (Pollock and D'Adderio, 2012; Pollock and Campagnolo, 2015; Quattrone, 2009) or the socio-technical arrangement of trading rooms in recognizing and constructing notions of value (Beunza and Stark, 2004). Other studies have explored how accounting systems may be implicated in creating compromise between different value criteria (Chanhall, Hall and Smith, 2013) and the political processes of defining and controlling the valuation of contested goods (see Reinecke, 2015). By conceptualizing value as neither an

objective property of an element, nor a subjective preference, but instead as a derivation of valuation practices, this literature has begun to advance knowledge in relation to the question of how multiple notions of value can be theorized. Furthermore, it has shed light on theoretical and societal implications of over-committing to a singular conceptualization of value, for example market value (Kornberger et al., 2015; Muniesa, 2011; Boltanski and Thevenot, 2006; Sandel, 2012). However, less is known in relation to the design of evaluation processes that not only embrace multiple forms of value beyond notions of post-hoc justification, but that also enable reflective decision-making within situations of uncertainty (Lamont, 2012; Stark, 2009).

To advance the theoretical concerns of this dissertation – the design and practice of evaluation systems to facilitate engagement with the complexity and ambiguity in the management of humanitarian crisis – a specific concept from the SOW literature is mobilized, namely heterarchy (Stark, 2009; Lamont, 2012). As a form of governance, heterarchy encourages the organization of heterogeneous and even dissonant principles of evaluation. While dissonance is commonly associated with destructive forces, and organizations commonly aim for the smoothest alignment between their objectives, activities, and evaluation criteria, the concept of heterarchy highlights how dissonance and ambiguity might be considered a resource in uncertain situations. A basic premise is that when the organizational environment is unstable and there is "uncertainty about what might constitute a resource under changed conditions, contending frameworks of value can themselves be a valuable organizational resource" (Stark, 2009, p.6). Drawing from the notion of heterarchy as a mode of organization implies the recognition that incomplete categorizations can foster contradiction, tension and re-association.

Heterarchical systems hereby shift the focus on processes that seek to institutionalize forms of criticism and competition, foster checks and balances, and recursively create spaces of ambiguity and reflection. Furthermore, rather than separating different value spheres, and instead of attempting to forge compromises between them (Chenhall, Hall and Smith, 2013), the notion of heterarchy raises questions regarding how organizations might encourage a more radical and challenging strategy to regularly foster ambiguous spaces that enable reflection and challenge expectations (March, 1981, 2006).

Drawing from the concept of heterarchy thus provides promising insights on a distinct approach towards engaging with heterogeneous forms of evaluation in contexts of uncertainty, which constitutes a key challenge in the management of humanitarian response operations (Barnett, 2011; Rottenburg, 2009). However, as argued by Lamont (2012), a conceptual understanding of the principles and conditions that might foster or hinder the practice of heterarchies is limited. While diverse principles of evaluation are probably a common element in most organizations, the question of how heterarchical principles might be productively stimulated remains largely under-theorized. Furthermore, theoretical understanding is limited in relation to the question of how coordination can be achieved despite the dissonance that forms an important element of heterarchical structures (Holm, 2010; Lamont, 2012). As such, the notion of heterarchy not only offers important insights for the key interests of this dissertation, but it also provides promising potential for novel theoretical developments to be explored through the empirical engagement with performance evaluation systems for the management of humanitarian crises.

To explore the central research interests of this study, a case study approach was selected as the appropriate method to enable a detailed examination of performance evaluation systems and practices within humanitarian crisis settings. As a definition of 'case study research', the study draws from Czarniawska's (2014) specification of the term as the investigation of the occurrence of a phenomenon that is studied in retrospect and limited in terms of time and scope. While it is commonly assumed that case study research pertains to the study of a clearly delineated organization or realm, Czarniawska's (2014) definition allows for a much wider conceptualization of the research phenomenon (see also Flyvbjerg, 2006). In line with this definition, the phenomenon that is investigated can be based on issues, for example a reform initiative in the public sector, or it can be centred on the design and practice of an accounting technology, which constitutes the focus of this dissertation. The methodological choice of conducting case study research is thus suitable for the key research focus of this thesis as it permits the in-depth investigation of the design and practice of humanitarian performance evaluation systems in their particular social context (see also Cooper & Morgan, 2008; Quattrone, 2006).

In line with this approach, three different case studies were carried out to investigate the research phenomenon. The first case study centres attention on analysing the design of the most widely used performance evaluation system that was developed with the single aim of facilitating engagement with humanitarian crisis contexts: *The Sphere Handbook* (hereinafter referred to as Sphere) (Barnett, 2011; Buchanan Smith, 2003; ECBP, 2007). As emphasized by Czarniawska (2014) as well as Flyvbjerg (2006, p.229), a promising approach when selecting a case study is to focus on the identification of a critical case

which has "strategic importance in relation to the general problem". Following this advice, the case study on Sphere, which is taken as a paradigmatic example, investigates the challenges surrounding the design of humanitarian performance evaluation systems. Drawing from the insights developed in the first case, the second and third cases explore how Sphere's system engages disaster managers and shapes evaluation practices in the management of a particular humanitarian crisis setting, namely a large scale refugee camp in the Middle East. Consistent with the selection strategy pursued with the first case study, case study two and three are based on paradigmatic examples of humanitarian crisis management, focusing on the different operational requirements of delivering nutritional supplies and the governance of water supply chains for disaster affected people. While the delivery of nutrition and water constitute two of the most consequential tasks of humanitarian crisis management (see Ramalingam, 2013; Barnett, 2011), the intensity and complexity of the performance evaluation challenges the managers confronted in both of these areas emerged during the field research and was not planned in advance. Due to the immense pressure and urgency that underlies the need to deliver these basic services in a reliable and consistent manner under constantly evolving circumstances, conducting case studies on the management of nutrition and water constituted an appropriate approach for the theoretical interests of this dissertation. In summary, carrying out the three case studies permits the thesis to investigate and explore tactics and principles that might enable humanitarians to evaluate performance under conditions of chaos.

1.2. <u>Research Questions</u>

In line with the dissertation's central research focus on the design and practice of performance evaluation systems to engage with situations of humanitarian crises, two research questions are proposed:

Research Question 1: What role might evaluation systems play in humanitarian crises which are characterized by high degrees of uncertainty, complexity, and ambiguity?

The first research question provides the overall direction of the investigation. It highlights its exploratory nature in relation to the possible roles of the design and practice of evaluation systems in the complex and uncertain environments that define humanitarian crisis settings. Building on the overall direction of the first research question, a second research question is proposed:

Research Question 2: What kind of tactics and principles can be conceptualized to design and practice an evaluation system to engage with and embrace the ambiguity of humanitarian crises?

The second research question further focuses the attention of the investigation on tactics and principles of performance evaluation systems that facilitate engagement and embrace the fast changing and chaotic settings that humanitarians commonly have to operate in. Implied in the question is a key theoretical and practical concern identified in the introduction. Instead of

focusing on the question of how accounting systems might be implicated in the attempt to normalize and fully control disaster situations, the more promising question in humanitarian crisis management is how accounting systems might be designed and practiced to embrace anomalies, ambiguity and the unexpected.

1.3. <u>Contributions</u>

In pursuing these research questions, the dissertation makes three different contributions, which are briefly summarized in this section. The study's first contribution is to the literature on ambiguity as a resource to enhance the evocative power of accounting systems to engage with unstable and uncertain environments (Cooper et al., 1981: March, 1987; Chenhall et al., 2013; Quattrone, 2015a). Even though there has been a renewed interest in the role of ambiguity and incompleteness in the practice of accounting systems in recent years (see Dambrin and Robson, 2011; Jordan and Messner, 2012), the idea to proactively embrace ambiguity has so far remained under-theorized (see Cooper et al., 1981: March, 1987; Busco and Quattrone, 2015). Based on the findings from the case studies, the dissertation develops four principles that advance understanding of *how* performance evaluation systems can embrace ambiguity and build on it to inform managerial judgement. These four principles comprise of (1) in-built tensions between evaluative principles; (2) open and participatory design; (3) relational value and incompleteness; and (4) the enactment of minimalist control through a community of practitioners. By developing these principles, the dissertation shows that Sphere's evaluation system did not become influential in the engagement with the

humanitarian crisis because users simply 'made do' with its incomplete measures (Dambrin & Robson, 2011), or because managers attempted to fix or distance themselves from measures considered imperfect (Jordan & Messner, 2012). Instead, the explicit incompleteness and organized ambiguity that emerged through engagement with Sphere became influential by counteracting a tendency to oversimplify performance requirements and quickly classify observed challenges into familiar categories. In other words, the four principles not only promoted an attitude of scepticism, which reduced the likelihood that disaster managers were blinded by the evaluation system's limitations, but managers used the openly incomplete measures as adaptable templates for their engagement with the humanitarian crisis.

The second contribution concerns the concept of heterarchy, which raises insights on a distinct notion of engaging with divergent principles of evaluation in contexts of uncertainty beyond notions of compromise (Stark, 2009; Lamont, 2012; Chenhall et al., 2013; Berthoin-Antal, Hutter & Stark, 2015). Through the theorization of the principles of in-built tensions, open and participatory design, relational value and incompleteness, and the enactment of minimalist control through a community of practitioners, this study contributes to theoretical understanding about elements that enable and constrain the practice of heterarchies (Stark, 2009; Lamont, 2012). An important implication emerging from the study is that tensions are not enough to sustain heterarchies. Another implication is that not all types of tensions are insightful and productive in the same manner. As shown in the case studies, while the complex social setting of the refugee camp was defined by a whole range of tensions and conflicts, evaluative tensions only became generative in specific phases of the response operation and through the mediation of

Sphere's evaluation system. These elements ensured that the engagement with heterarchical tensions did not lead to chaos, confusion and strife. Furthermore, the study also highlights that while evaluative tensions were crucial to stimulating processes of exploration, the productive nature of heterarchical tensions diminished when the disaster response entered phases in which requirements had to be adapted to contextual challenges. In such situations, different elements became significant, for example the principles of relational value and incompleteness that facilitated the organised recombination of existing resources. Combined, the four principles this study theorizes contribute to scholarly understanding of the practice of heterarchies, beyond a mere focus on evaluative tensions (Stark, 2009; Lamont, 2012).

The third intended contribution of this study is to the emerging accounting literature on humanitarian response operations. As outlined previously, an important focus of the existing literature on accounting in humanitarian crises has been on exploring how notions of long distance control are re-constituted through accounting technologies (Sargiacomo, 2015) or how humanitarians employ distinct accountability scripts to legitimate their actions in a post hoc manner, thereby contributing to their perceived normalization and rationalization (Everett and Friessen, 2010; Taylor et al., 2014; Lai et al., 2014). By problematizing issues of complexity and ambiguity as key challenges in humanitarian response operations, the study theorizes how accounting technologies can be implicated in the adaptive management of humanitarian crises. Instead of employing accounting technologies to attempt to 'normalize' humanitarian crises, the case studies provide rich evidence of how disaster managers drew from heterarchical principles for performance evaluation to embrace the dynamism of humanitarian contexts and facilitate permanent

adaptation to emergent challenges. In doing so, the study advances theoretical and practical understanding in relation to the question of how accounting systems may be designed and practiced to stimulate reflective capacities in the management of humanitarian crises.

1.4.   <u>Structure of the Thesis and Concluding Remarks</u>

The thesis is structured as follows. Following the introduction, the literature review, methodology, empirical findings, discussion and conclusion outline the theoretical positioning of the study, how it was conducted, what it found and how it contributes to existing knowledge. The literature review comprises two separate chapters, one addressing relevant research in accounting in chapter two and the other one discussing pertinent studies from the Sociology of Worth in chapter three. These chapters further explicate the research gaps and outline the compatibilities between the theoretical concerns and gaps in the two literatures. In chapter four, the study's methodology is discussed, outlining how the study was carried out and which choices underlie the research design. These include the study's broad philosophical assumptions and the methods that were used to collect and analyse empirical data. In chapter five, the empirical findings are presented through three case studies, focusing on performance evaluation techniques and principles that facilitate engagement with the complexity of managing humanitarian crisis settings. In the discussion presented in chapter six, the study develops four distinct heterarchical principles that advance theoretical understanding of how performance evaluation systems might make ambiguity a resource. The final chapter provides concluding remarks, summarizes the three key

contributions, discusses practical implications, outlines possibilities for further research and offers reflections on the study's limitations.

2. Literature Review

2.1. <u>Introduction and Overview</u>

Building on recent calls for research (see Walker, 2016; Sargiacomo, 2015), the review shows how, in relation to one of the key issues for accounting systems in humanitarian crises, scholarly understanding remains underdeveloped: how to account for and evaluate actions in humanitarian environments defined by high degrees of uncertainty, complexity and continuous change. Organizations operating within such uncertain conditions provide several theoretical and practical concerns for accounting systems (Burchell et al., 1980; Cooper et al., 1981; March, 1987; Quattrone, 2015a, b). In humanitarian crises, where governance is commonly experienced by way of its limitations, accounting systems play numerous roles beyond the representation of performance (Carruthers, 1995; Tinker, 1991) and the control of populations (Miller and Rose, 1990). Against this background, the study explores how humanitarian performance evaluation systems and practices enact notions of ambiguity (March, 1987; Cooper et al., 1981; Mouritsen, 2016), incompleteness of accounting information (Jordan and Messner, 2012; Dambrin and Robson, 2011) and a multiplicity of distinct modes of evaluation (Chenhall et al., 2013; Stark, 2009; Coslor, 2016).

To conceptualize these issues, the literature review is structured around four key sections. The first section reviews literature on accounting as an intrinsically social and organizational practice (Burchell et al., 1980; Hopwood, 1985, 1987). This section situates the study within a body of interdisciplinary research that highlights the agency of accounting in the constitution of the social and provides context for the following sections (Chapman, Cooper and Miller, 2009). The second section reviews literature that focuses on theorizing the role of accounting as a technology for social control. It demonstrates advancements in scholarly understanding in the way accounting technologies facilitate the 'conduct of conduct' (Miller and Rose, 2008), shape identities of the marginalized (Walker, 2008), enable action at a distance (Robson, 1991, 1992), and construct governable spaces (Miller and Power, 2013). Insights from this literature have been influential in the emerging research on the role of accounting in disasters. Building on the first two sections, the third section demonstrates that the primary focus of the emerging studies on humanitarian emergencies has been on the question of how accounting technologies can be mobilized to re-assert control over specific crisis settings, thereby contributing to their gradual normalization (see Sargiacomo, 2015; Sargiacomo et al., 2014). By explicating some of the limitations of this argument, the review emphasizes that current literature does not sufficiently engage with a key concern for humanitarian accounting systems beyond notions of control and attempts to normalize disaster situations. The review stresses that more engagement is needed with the question of how accounting systems might be designed and practiced to facilitate engagement with anomalies, ambiguities and the unexpected in the management of humanitarian crisis. In line with this question, the fourth section reviews and analyses literature that focuses on the role of ambiguity, doubt and incompleteness as a resource of accounting

technologies for organizations facing situations of uncertainty. These studies open up questions on how designers might develop accounting systems that embrace incompleteness and ambiguity to foster their evocativeness, interpretive flexibility and reflective capacities (Cooper et al., 1981; March, 1987; Chenhall et al., 2013; Quattrone, 2015b; Revellino and Mouritsen, 2015).

2.2. <u>Accounting and the Social</u>

2.2.1. *Situating Accounting in Society*

In the first line of the founding issue of Accounting, Organizations and Society (AOS) Hopwood (1976:1) proclaimed "accounting has played a vital role in the development of modern society". Hopwood's statement was characteristic of the ambition of AOS as a new journal, marking the starting point of a rich tradition of interdisciplinary accounting research over the next forty years. Attention to the role of accounting in the shaping of modern societies and social relations can also be found in some of the texts now commonly considered as classics of social science literature. In *Protestant Ethic and the Spirit of Capitalism* (1956), Weber stresses the role of accounting, and in particular double-entry bookkeeping, in the development of a particular form of (calculative) rationality, which he considers instrumental for the establishment of modern bureaucracies and the rise of capitalism. In this context, Weber further draws attention to the possible dehumanizing effects of an emerging 'iron cage' of rationalized forms of social organizing. Schumpeter (1950) and Sombart (1967) made similarly strong claims in relation to the connection of accounting and social organization by arguing

that double-entry bookkeeping was a primary driver for the rise of capitalism. In another social science classic, *Capital*, Marx (1976) interprets the social significance of accounting practices in a radically distinct manner. For Marx, accounting practices provide a tool that serve to mystify and hide rather than reveal the nature of social relationships underlying the forces of production. Despite their difference in perspective, what is common in these short descriptions of Weber (1956), Schumpeter (1950), Sombart (1967) and Marx's (1976) analyses is that the intrinsically social character of accountancy systems is highlighted and their agency in ordering patterns of power and influence problematized. In other words, these early writings situated accounting as an influential object of social science research (see also Carruthers and Espeland, 1991; Miller and Power, 2013; Quattrone, 2015a).

Building on these traditions, in another early AOS editorial, Hopwood (1978b:190) encouraged further exploration into the social nature of accounting:

> For those interested in the social aspects of accounting the horizons are particularly bright. In the past such interests have almost always had to be confined to the development of particular social accountings. Now, however, the turmoils of practical discourse have illustrated that such social forms of accounting, important though they may be, are only a particular manifestation of the social nature of accounting itself.

Ever since, accounting has assumed an increasingly important position in the organizing of modern societies to the extent that "it is perhaps the most powerful system of representation for social and economic life today in many national settings" (Miller and Power, 2013, p.558). One of the prevalent

manifestations of this argument is that, despite their disputed nature, almost every area of public life today is evaluated or rated by performance measures (see Power, 1997). These include the arts (Coslor, 2016), higher education (Espeland and Sauder, 2007), holidaymaking (Jeacle and Carter, 2011), cities (Kornberger and Carter, 2010) and health care (Llewelyn and Northcott, 2005). Burchell et al. (1980:6) further describe a few further significant areas in which accounting practices are integral to social and economic management:

> Accounting data are now used in the derivation and implementation of policies for economic stabilization, price and wage control, the regulation of particular industrial and commercial sectors and the planning of national economic resources in conditions of war and peace and prosperity and depression.

The breadth of studies reveals a broad consensus amongst researchers for a need to theorize the power and influence of accounting and performance measurement systems in organizations and society (Hopwood, 2005; Burchell et al., 1980). By raising questions as to why and how accounting practices have come to develop into such significant and pervasive governance tools and what roles they serve in organizations and society, this emerging range of studies contributed to the growth of a rich research community dedicated to addressing these issues (see also Hopwood, 1983, Walker, 2016). Their research agenda stimulated a departure in accounting scholarship away from a predominantly functionalist paradigm, which focused on the chronological analysis of technical developments based on notions of rational-choice and efficient resource administration (see Chandler and Daems, 1979; Boyns and Edwards, 1996).

Research that emerged during the subsequent decades drew emphasis on the historically and socially contingent nature of accountancy practices. In Hopwood's (1987:207) words, accounting continuously exhibited a "tendency to become what it was not" as it spread into being applied into ever more areas of social governance, contributing to a profound transformation in the scholarly understanding of accounting:

> In the space of a little more than a decade, there has been a profound transformation in the understanding of accounting. Accounting has come to be regarded as a social and institutional practice, one that is intrinsic to, and constitutive of, social relations, rather than derivative or secondary (Miller, 1994, p.1).

By presenting four genealogies of distinct accounting calculations, Miller and Napier (1993) showed how accounting assumes distinct contents and forms over time, as new techniques are invented and old ones shifted from one context to another and are given new meanings in this process. Miller and Napier (1993:631) hereby not only raised the question of what can be counted as accounting but also emphasized that "there is no 'essence' to accounting and no invariant object to which the name 'accounting' can be attached". This argument had wide-ranging implications as it drew attention to the ad-hoc fashion in which accounting techniques are assembled and assert agency in relation to historical, social and geographical concerns. Rather than attempting to formulate general principles to establish the boundaries of accounting and to define what is inside and outside of these, Miller (1998) proposes to study

accounting as a form of *bricolage*: an activity that depends to a large extent on improvisation and adaptation to issues and resources based on context. In the face of evidence highlighting the malleability of the boundaries of accounting, Miller (1998:605) further suggested: "accounting is most interesting at its margins." In this view, it is at the margins of accounting where it connects and clashes with other areas of knowledge and expertise and where it is formed as new techniques are added and removed.

The establishment of the interconnection between accounting systems and the constitution of social associations are of particular relevance to this study for at least three interconnected reasons. Firstly, it encourages the investigation of accounting beyond the economic domain (see also Hopwood, 2005; Walker, 2008). While interdisciplinary accounting research focuses on the investigation of the roles accounting performs "in both the construction and realisation of the domains of the social and the political" (Hopwood, 1985, p.366), the accounting academy continues to be dominated by a focus on organizations located in the economic sphere (see Walker, 2016). Secondly, it draws attention to the particularities of accounting as a system of knowledge production that interacts strongly with other knowledge systems (see Miller, 1998), promoting the investigation of their connections and epistemological underpinnings (see Hines, 1988; Morgan, 1988; Tinker, 1991; Carruthers, 1995). As stressed by Quattrone (2015b:53), accounting reports and classifications are political statements whose epistemology acts in a subtle and often silent manner:

> It is this unawareness of the power of representations and the perspective that they inevitably take that makes the…users believe that these numbers are facts, incontestable values where the correctness of

the calculation implies its justness, making people forget that facts are always made.

As will be shown below, in light of the commonly incomplete and ambiguous nature of accounting information in situations of crisis, the problem of representation constitutes a challenging task for the design and practice of humanitarian performance evaluation systems. As emphasized by Chenhall et al. (2013: 269), "accounts of performance are critical because it is in discussions over the different metrics, images and words that can be used to represent performance that the actual worth of things is frequently debated and contested." What is valuable, why it is valued and how it is constructed, and represented thus requires an analytical focus on the systems and practices through which distinct notions of value are performed (see Stark, 2009; Kornberger et al., 2015). Thirdly, attention to the interconnection between accounting systems and the constitution of the social encourages the examination of accounting technologies in the enactment of various forms of social control (Burchell et al., 1980; Miller and Rose, 1990; Walker, 2008; Ezzamel, 2012; Quattrone, 2015a). Combined, these areas of interest provide a fruitful starting point for developing an understanding about the possible roles accounting systems might play in the governance of humanitarian crises and refugee camps. The following sections discuss key contributions in these areas and analyses their relevance as well as shortcomings in relation to key concerns for the design and practice of humanitarian performance evaluation systems.

2.3. <u>Accounting, Control and Action at a Distance</u>

2.3.1. <u>*Governance Programmes and Technologies of Control*</u>

A key take away from the previous section is that accounting does not passively reflect reality but its development and practice is inter-related with other projects for social governance, which depend on accountancy techniques to construct reality by classifying, ordering and visualizing social phenomena (Burchell et al., 1985; Miller, 1994; Chapman, Cooper and Miller, 2009). In this context, the notion of social control emerged as an important concept for researchers interested in the social aspects and implications of accounting. As demonstrated in section 2.4, insights from this literature on control have also been influential in the emerging research on the role of accounting in disasters and crises (Sargiacomo, 2015, 2014; Taylor et al., 2014; Everett and Friessen, 2010).

Against this background, in a review article on interdisciplinary accounting research over the last forty years, Walker (2016: 45) highlights the following:

> While it may be accepted that accounting is potentially an instrument of social control the way in which these technologies are mobilised in this respect is often assumed rather than demonstrated. Although contributions to the behavioural and 'critical' literature may refer to social control, the concept is seldom disentangled from generalised notions of managerial strategies of control, organisational control and power relations. Indeed, the term 'control' itself though considered

fundamental to the accounting discipline, is invariably taken as a given and defined atemporally as simply 'being there'.

Even though Walker (2016) points to some shortcomings in the research canon, a variety of important studies have advanced scholarly understanding in relation to the role of accounting as a technology of social control and governance. By taking the historical construction of concepts of standard costing and budgeting as their unit of analysis, Miller and O'Leary (1987) highlight how these developments, instead of merely advancing the technical refinement of accounting, formed part of a wider project of expanding the role of the state in the active management of the person in a variety of ways. Miller and O'Leary (1987:261) link the emergence of bodies of accounting expert knowledge to the construction of the individual as a manageable and efficient entity with the political aim to increase output and prosperity: "quite literally, the person was to be worked upon, to be managed through a series of interventions into an enhanced state of life." Accounting technologies are thus theorized as implicated in the 'conduct of conduct', in particular by "linking its instrumental aspirations to act on the actions of individuals with attempts to exercise political power over an entire population or nation" (Mennicken and Miller, 2012, p.14).

In line with this reasoning, to investigate the question of how accounting and notions of social control may mutually constitute each other, a number of studies explored the interplay between programmes and technologies for governance. These studies include investigations into the enrolment of accounting practices in the development of modern bureaucracies (Miller, 1990); the spread of discounted cash flow accounting as a result of the

problematizing of economic growth as a public policy target (Miller, 1991); and the reform of American industry in which accounting technologies became linked to an agenda to foster a novel type of economic citizenship (Miller and O'Leary, 1994). Adding to these insights, Ezzamel (2012) analysed and described a variety of elements that allow accounting technologies to shape and construct notions of social order. These elements include valuation techniques, procedures that facilitate memorization, ritualization and organization and the cultivation of accountability relationships through reporting (see also Quattrone, 2009; 2015a; Walker, 2016). As shown below in section 2.4.2, the conceptualization of coherent programmes of governance that are linked to technologies of control is highly problematic in theoretical and practical terms within the dynamic contexts of humanitarian crisis settings.

Miller and Power (2013) further theorize the mutually constitutive nature of accounting, control and organizing practices: "If organizing without accounting is increasingly unthinkable today, accounting also makes organizations thinkable and actionable in a particular way" (Miller and Power, 2013, p.558). Miller and Power thereby identify four key roles that make accounting influential as a technology for social control and governance. These include: (1) territorializing, namely the construction of spaces in a way that they become amenable to calculation and thereby conducive to governance (see also Mennicken and Miller, 2012; Mouritsen, 1999); (2) mediation, which highlights the agency of accounting technologies in connecting distinct actors and domains that are previously separate, such as science and the economy (see Miller and O'Leary, 2007), perceptions of control and creativity (see Chenhall et al., 2011), or different conceptualizations of notions of risk (see

Jordan et al., 2013); (3) adjudication, which highlights some of the disciplinary roles of accounting in not only evaluating performance but also in defining failure and success in organizations and society (see also Pollock and D'Adderio, 2012); and (4) subjectivizing, which highlights how accounting technologies contribute to the regulation and control of its subjects. While the four roles identified by Miller and Power (2013) are interrelated to significant extents, they provide important insights into the role of accounting in enacting various forms of control, disciplinary mechanisms and social organization. These studies outline the potentialities of accounting systems as technologies that allow for the spread of highly subtle and non-coercive forms of influence and control.

In line with these insights, Walker (2008) importantly draws attention to the more sinister side of accounting as a technology of social control in a case study on the role of accounting classifications and calculations in the administration of a poverty relief programme in Victorian England and Wales. With a focus on the role of accounting in creating the identities and in re-enforcing stigmatization of the marginalized in society, Walker (2008:453) emphasises the risk that accounting may serve as a technology to embed oppressive structures through which social control is exercised and perpetuated (see also Walker, 2004):

> Accounting processes comprised degradation ceremonies which compounded the stigmatisation of the recipient of relief, accounting classifications served to inscribe existing and create additional spoiled identities of the pauper, and individualized forms of accounting disclosure compromised the management of stigma by the poor.

The concerns raised by these studies in relation to the creation of stigmatized identities or prejudice through the administration of poverty relief programmes and other accounting classification mechanisms provide a highly relevant problematization of both the enabling and constraining potential of accountancy and control systems (see also Annisette, 2009; Greer and Neu, 2009). This argument serves as a powerful reminder of the importance of continuously balancing means and ends (see Quattrone, 2015a) as well as different forms of value and values (see Chenhall et al., 2013), both of which are of particular importance in the management of crises and the administration of relief to refugees.

2.3.2. *Circulating Inscriptions and Control*

To further theorize the interrelationship between accounting technologies and social control, a number of important accounting studies have drawn from insights from Actor-Network Theory (ANT) (Latour, 1987, 1999, 2005; Callon 1986). These studies draw attention to the process through which centres of calculation, for example government agencies or donors in humanitarian crises, attempt to assert control and act from a distance. Furthermore, this research provided important insights into the epistemological implications for accounting as a practice, which is inherently concerned with the problem of representation. In this context, Robson (1992:691) highlights the following:

> If knowledge is oriented towards acting upon a remote setting, then it is produced and sustained not by 'true' correspondence but by its power in securing long-distance control, through the provision and

maintenance of networks for the gathering, transmission and assimilation of inscriptions. Inscriptions translate the elements of the context (Robson, 1992, p.691).

The terms 'action at a distance' and 'inscription' require some further clarification. As Latour (1987) emphasizes, when people communicate over temporal and spatial distances representational practices gain crucial importance as a form of transporting information from one distant place to another. In such situations, inscriptions play the role of translating elements of the distant entity in its absence (Latour, 1987). For example, if a donor organization wishes to control and act upon a humanitarian agency that works in a disaster setting, i.e. a refugee camp, it needs to acquire the ability to accumulate inscriptions, for example performance measures, to bring information back to the headquarters. As emphasized by Miller (1990:318):

> Written reports, books of accounts, pictures, charts all represent a domain and can be deployed in attempts to administer it. As technologies they do not have a neutral function of recording the real, but literally *re*present in such a way as to make it susceptible to evaluation, calculation and intervention.

Conceptualizing accounting as a type of inscription device allows researchers and practitioners to pay attention to the process through which disparate and diverse entities and actions are translated into performance measures, balance sheets or income statements and through this process are transformed into entities that can be made comparable by a centre of calculation (Robson, 1991, 1992). When a material entity is inscribed, the resulting inscription, for example a number or sign, does not unambiguously represent a 'reality out

there.' Instead, it shows that in the process of inscribing there are a range of assumptions and theories of measurement at work, which exist prior to the inscription, and are thus partly constitutive of the reality that is represented (Woolgar, 1988). In the explanation of the rhetorical persuasive powers of inscriptions, Latour (1988:158) highlights the following:

> The corollary of this "holding of" several elements by one, is a general feeling of strength, economy, and aesthetic satisfaction: the one element may "replace", "represent", "stand for" all the others, which are in effect made secondary, deducible, subservient or negligible.

The introduction of the concept of inscription proved to be theoretically insightful in several ways for accounting researchers. Accordingly, accounting creates a powerful medium, namely the inscription, through which actants can assert influence on the behaviour of other actants in remote places (see Robson, 1991; 1992; Briers and Chua, 2007; Gendron, Cooper and Townley, 2007; Justesen and Mouritsen, 2011). In line with these insights, the introduction of inscriptions as a concept was related to the aim to develop alternatives to a Correspondence Theory of Truth (CCT) perspective on accounting figures and the model *adequatio rei et intellectus* (the equation of thought and thing). The focus on inscriptions involves a theoretical shift from 'true' and 'fair' correspondence between reference and referent towards a focus on the transformations and methodological challenges that are necessary for accounting systems to build chains of references that transport information from distant places to a managerial centre (Andon et al., 2007; Latour, 1999).

The theoretical advancements outlined above spurred numerous innovative studies in a variety of crucial areas of accounting scholarship. Through a detailed ethnographic account, Preston et al. (1992) demonstrate the complex processes of alliance building and the constant transformations the budgeting system goes through as different professional groups translate it to their local contexts. Briers and Chua's (2001) study on the 'implementation' of Activity-Based Costing (ABC) is further in line with arguments on the translation of accounting systems. Briers and Chua (2001:240) hereby show how the fate of accounting technologies always lies "in the hands of those who come after the 'inventor' and success/failure is a social accomplishment by many different human and non-human elements". Llewelyn and Northcott (2005) studied standardization processes, e.g. the construction of the 'average hospital', through which costs could be compared and controlled in the UK health care sector. Their study contributes to understanding the complex process through which the construction of centralized benchmarks can affect everyday practices within widely dispersed hospitals through the UK. Gendron et al. (2007) show how auditors construct notions of professional expertise in relation to the programmatic idea of New Public Management. Similar to Latour's (1987) concept of 'centre of calculation' they show how the Office of the Auditor General functioned as a type of laboratory in which expertise was fabricated and legitimized, networks established and adversaries converted.

2.3.3.   *Theorizing Emergent and A-centred Forms of Accounting Control*

However, despite the insights offered by the studies discussed in 2.3.2., the manner in which the concept of action at a distance is commonly used in the

accounting literature is not unproblematic in both theoretical and practical terms. As noted by Dechow and Mouritsen (2005) as well as Quattrone and Hopper (2005), in theoretical terms it risks reinforcing traditional social science dichotomies, including a relatively linear conceptualization of the relationship between controllers and the controlled. Quattrone and Hopper (2005, p. 760) draw attention to the following problem:

> It is more pertinent to trace continual changes in loci of control rather than trying to identify a specific centre that exerts action at a distance based on modernist presumptions of a dichotomy between the controller and linear and uniform time and space.

In their study on the practical use of management control systems in two multinational firms, Quattrone and Hopper (2005) show how notions of centre and periphery and controller and controlled are always emerging constructs that cannot be inferred from the pre-packaged management control system. While management control technologies have certain enabling and constraining features, Quattrone and Hopper (2005) demonstrate that it is not a stable technology and does not produce stable forms of organizing. These theoretical concerns are mirrored in studies by Andon et al. (2007:273) on the "inherent unsettledness of accounting as a knowledge object," as well as Dechow and Mouritsen (2005), who focus on theorizing how the 'implementation' of Enterprise Resource Planning systems transform networks of power in organizations (see also Justesen and Mouritsen, 2011).

The implications of this theoretical critique of action at a distance is of particular importance for the design and practice of humanitarian

performance evaluation systems. As Walker (2008) has demonstrated, accounting systems that seek to reinforce a strict dichotomy between controllers and controlled can have stigmatizing, and possibly even de-humanizing, effects on the marginalized. Furthermore, conceptualizing accounting as an inscription device not only draws attention to the factors that might make it powerful, but it also highlights the inherent incompleteness of accounting classifications, which refer to the world in its absence (Dambrin and Robson, 2011; Jordan and Messner, 2012; Busco and Quattrone, 2015). Accounting systems as control and ordering devices are therefore always partial and destined to produce blind spots that cannot be anticipated by their designers. As will be argued below, recognizing and building on this inherent incompleteness constitutes a vital pre-condition to address and mitigate the potential de-humanizing effects of performance evaluation systems in the management of humanitarian crises and to navigate situations of uncertainty, complexity and dynamic change. In line with these concerns, the review now more specifically turns to the emerging literature on the role of accounting systems in humanitarian crises.

2.4. Accounting for Humanitarian Crises: From Normalizing to Anomalizing Practices

2.4.1. *Accounting Technologies and the Governance of Humanitarian Crises*

Interest in the role of accounting technologies in the management of situations of crisis dates back several decades. In one of the early calls for further investigation into the topic, Hopwood (1987:231) outlined the following:

> Alongside a more nuanced view of the role of crisis, we need to appreciate the ways in which new bodies of knowledge, new specialists associated with their practice, government regulatory attempts, changing theoretical and practical conceptions of organisational governance and order, and even the development of different accounting rhetoric can provide the basis for action and change.

Despite its remarkable foresight, only recently and only against the background of a variety of corporate (Andon and Free, 2012), financial (Quattrone, 2015b) and humanitarian crises (Sargiacomo, 2015), Hopwood's (1987) call received renewed and more significant attention in interdisciplinary accounting scholarship (see also Walker, 2000). In a special issue celebrating the fortieth anniversary of AOS, Walker (2016:22) re-emphasizes the importance of the topic for the future agenda of accounting researchers, while highlighting the limited understanding of the topic in the accounting academy:

A renewed focus on accounting in crisis situations is likely to accompany the sociology of disruption, trauma and calamity. Emergent agendas concerning the functioning of accounting in the management of natural and humanitarian disasters and the control of dislocated populations would likely assume greater prominence in 'catastrophic societies.'

By employing the term 'catastrophic societies' Walker (2016) draws from the work of Elliot and Turner (2012), who envisage a variety of possible 'social futures' that might shape the twenty-first century. In particular, Elliot and Turner (2012) point to the likelihood of significant social transformations that derive from the increasing threats of conflicts as well as nuclear, biological and environmental hazards. In such 'catastrophic societies', Walker (2016) argues, accounting technologies will become heavily implicated in the global management of increasingly scarce resources, including water, food and fuel. While this outlook is insightful to sketch out novel agendas for the investigation of accounting technologies and emerging forms of global governance with promising theoretical as well as social implications, one does not have to look into the future to find some of these trends unfolding.

In line with this emergent agenda, recent research has begun to explore the role of accounting in disasters, focusing mainly on how accounting technologies contribute to control and gradually 'normalize' crisis situations. Sargiacomo (2015) examines the role of accounting in managing the response to the 2009 earthquake in Italy. Drawing insights from previous research on governmentality (Miller and Rose, 1990), Sargiacomo highlights how accounting classifications formed an important component to demarcate the space of government intervention, providing not only visibility to the costs of

earthquake injured victims, but also facilitating the distribution of resources to distinct segments of the population:

> Much like situations of normal government, accounting numbers help to quantify the 'economic' magnitude of the event. Accounting as 'an inherently territorializing activity' (Miller and Mennicken, 2012), involves dividing territories and people from those most affected to those not affected at all, attaching financial and costing measures to those categorized, and developing boundaries for such issues as healthcare free exoneration or income tax exemption (Sargiacomo, 2015, p.70-71).

In line with this framing, Sargiacomo (2015) defines the calculative infrastructure that emerges in response to the Earthquake as the construction of an "exceptional government assemblage" that eventually facilitated the return to more "normal" forms of governance which enables control at a distance (see Robson, 1991, 1992). To re-assert control and build a rudimentary reporting apparatus, Sargiacomo (2015) highlights how government agencies borrowed from existing accounting as well as from scientific classifications and how these classifications were mobilized in strategic games between the local and national government authorities to address conflicts over resource allocation (see also Sargiacomo et al., 2014).

In another recent historical study on the interrelationship between accounting and crisis, Walker (2014) discusses how accounting served as an important facilitative technology to help communities recover from prolonged droughts that were declared as a national emergency in the United States of the 1930s. Walker (2014:2) uses the concept of heroic bureaucracies, defined as "a

government organization which uses public monies to address a fundamental problem in a new and socially progressive way," to show how an experimental government programme for the resettlement of dislocated families was conceptualized and implemented. Accounting tools such as home planning, record keeping and budgeting hereby served as key instruments to assess the progress of participating families and to monitor the performance of the overall programme (Walker, 2014).

Drawing from the work of O'Dwyer and Unerman (2008) and Gray, Bebbington & Collison, (2006), other recent studies focus on the practice of different notions of accountability in humanitarian disasters. In line with these arguments, Everett and Friessen (2010: 468) argue that humanitarians draw from different and often contradictory accountability scripts to justify actions in emergency relief, leading to "an ironic and agonistic play that occurs both on and off stage, this contest pits humanitarians against their donors." As an example of predominantly 'technical' accountability requirements Everett and Friessen (2010) suggest the Sphere standards, a conceptualization this thesis challenges below. By mobilizing the work of Judith Butler, Everett and Friessen (2010) highlight the performative nature of humanitarian accountability.

In a similar fashion, Taylor et al. (2014) focus on the impact of the increasing influence of hierarchical and holistic accountability in organizations with social and environmental missions through a case study on the 2009 bushfire in Australia. The term hierarchical accountability points to a narrow and often short-term orientation of accountability to stakeholders that control resources.

Holistic accountability connotes a broad conceptualization of accountability towards including a multitude of possible outcomes of NGO activities on communities and society (see also O'Dwyer and Unerman, 2008). As stressed by Taylor et al. (2014), while NGOs commonly attempt to address notions from both hierarchical and holistic accountability, the competing demands posed by these distinct forms of accountability may lead to a corruption of the organizational mission and an unproductive drift of NGO strategy to the detriment of beneficiaries (see O'Dwyer and Unerman; 2008; Gray et al 2006; Ebrahim, 2005). The study particularly highlights that the notion of downward accountability is problematic within a stakeholder-oriented accountability framework that is commonly in place in NGOs and other social and charitable organizations (Taylor et al., 2014). Combined, these studies provide insights into the role of accounting practices in defining 'exceptional' governance spaces in disaster contexts and the use of accounting as a facilitative technology for social transformation in disaster-affected communities.

2.4.2. *'Anomalizing Accounts' and the Management of Humanitarian Crises*

As emerges from the studies reviewed above, the primary focus of the emerging research on humanitarian disasters has been on the question of how accounting technologies are mobilized to control and 'normalize' crisis situations (see Sargiacomo, 2015; Sargiacomo et al. 2014) as well as on the accountability scripts that are used to justify activities in an ex post manner (Everett and Friessen, 2010; Taylor et al., 2014). While these studies provide important insights, in the exceptional and even anarchic environments (Cooper *et al.*, 1980) that commonly define humanitarian practice, important

organizing and sense-making challenges do not necessarily relate to putting into action specific programmes of control. As Weick and Sutcliffe (2007) highlight, a significant challenge of engaging with situations of crisis is to manage and engage with the unexpected. For Weick and Sutcliffe (2007:2/3) crises are defined by:

> Vague notions of contingency resources, incomplete knowledge of the system, unexpected changes in staffing, uneven communication, quotas, and shifting command structures. When the unexpected wind swirled into...[the] system, the vagueness, the incompleteness, and the shifting command were the weak points that gave way.

It follows from the quote that in situations of crises information and governance systems are commonly experienced by way of their limitations, with shifting centres of control (Quattrone and Hopper, 2005), ambiguities of goals, capacities and choices (Cooper et al., 1981; March, 1987), as well as insufficient and incomplete data (see also Barnett, 2011; Baker, 2014). As a result, notions of performance and evaluation are inherently fluid and contested. Against this background, Weick and Sutcliffe (2007:33) further highlight the role of categorizations and reporting mechanisms as crucial elements in the way organizations and decision-makers relate to crises:

> If you want to report something, you need to have words and categories at hand to do the reporting. And those very words can limit what you see and report. Whatever labels a group uses will colour what members of the group think they see and report. This means that people may miss the unexpected and label too many events as 'in-family'.

While accounting categories and classifications are vital to assess the scope and magnitude of humanitarian crises, an unreflective use of performance categories and control mechanisms can contribute to exacerbating the situations they are designed to detect or contain (see also Weick, 1988). As highlighted by Ramalingam (2013), no one knows enough to design an information system that is able to capture and cope with the array of interrelated variables that are at play during humanitarian crises without producing a multitude of blind spots. If accounting and evaluation systems draw attention away from such blind spots and foster a predisposition to quickly classify information into familiar categories, there is an increased likelihood that these systems will actively prevent managers from detecting significant trends or incidents that might have far-reaching consequences (see also Rottenburg, 2009). In line with these insights, a key challenge for accounting systems in humanitarian crisis is thus to work against a tendency to simplify incidents by quickly categorizing them into well-known categories and instead to strengthen a predisposition to question and possibly re-conceptualize an observation into something that can be problematized. In other words, in the management of humanitarian crises, accounting systems need to be judged by their ability to challenge taken-for-granted categories and re-appraise expectations (Revellino and Mouritsen, 2015; Quattrone, 2015a, b).

In line with these insights, instead of focusing on the problem of how accounting technologies might be mobilized in an effort to normalize humanitarian crises (Sargiacomo, 2015), a challenging question is how accounting systems might be designed and practiced to facilitate engagement with anomalies, ambiguities and the unexpected. Focusing on engagement with anomalies and ambiguities has the potential to draw attention to possible

problems and reduce the risk of ignoring and overlooking emerging challenges to the point that they become impossible to manage (see Weick and Sutcliffe, 2007). While stimulating questioning is not commonly a virtue associated with accounting systems, for the management of humanitarian crises it is of critical importance to design systems that embrace and even routinize the exercise of questioning and critique (see Cooper et al., 1981; March, 1987; Mouritsen, 2016).

As emphasized by Ramalingam (2013: 138 – 139), despite rising recent interest, theoretical and practical understanding still remains underdeveloped in relation to the question of how administrative and performance evaluation technologies can inform adaptive capacities in fast-changing humanitarian environments:

> In the way aid agencies learn, strategize, organize, and perform, there is a clear manifestation of an obsession with organized simplicity…This gives us a simple choice: do we continue to struggle to answer questions with an intellectual toolkit not designed with such problems in mind, and a policy and operational toolkit that lags even further behind?

To further advance a conceptualization of performance evaluation systems beyond an 'obsession with organized simplicity' in the management of humanitarian crises, the following section discusses accounting studies that examine the role of ambiguity and incompleteness as a design resource for organizations facing uncertain situations (Cooper et al., 1981; March, 198; Quattrone, 2015b; Revellino and Mouritsen, 2015; Mouritsen, 2016).

2.5. Accounting Systems and the Power of Ambiguity

To contextualize the possible roles of accounting and evaluation technologies in relation to contexts of uncertainty and ambiguity, it is informative to briefly recollect an illustration provided by Burchell et al. (1980) on different possible roles of accounting systems in organizational practice. Burchell et al. (1980) specify that their two-by-two matrix constitutes a simplification for illustrative purposes. Nevertheless, the illustration offers an insightful starting point since it provided inspiration for a variety of important subsequent studies.

Figure 1: Uncertainty, Decision-Making and the Roles of Accounting in Practice
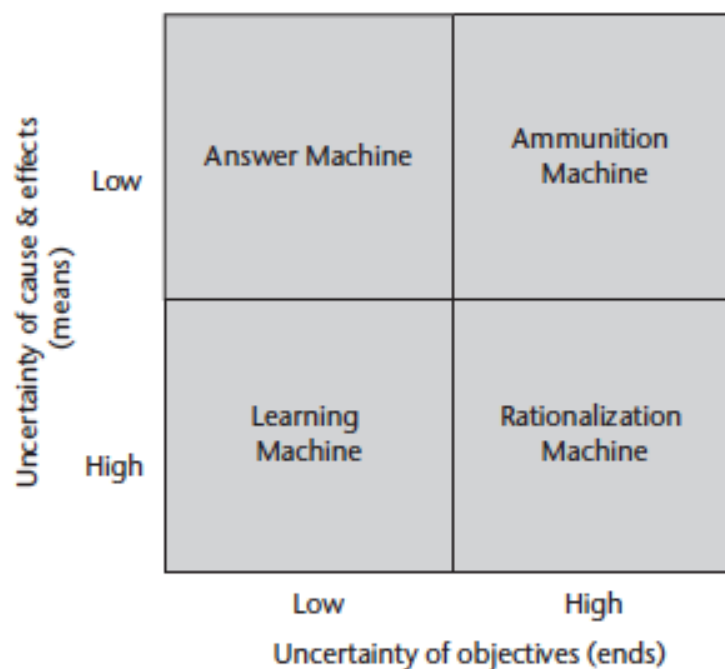
Source: Adapted from Burchell et al. (1980, p.14)

The vertical axis in figure 1 corresponds to the degree of uncertainty of cause-and-effect relationships and the horizontal axis relates to the degree of uncertainty of objectives that are present in a situation. By putting these two dimensions in relation with each other, Burchell et al. (1980) specify four different roles accounting might play. Firstly, if both uncertainty of cause-and-effect and uncertainty of objectives is low, accounting might work as an answer machine. Data can be fed into the machine and it produces straightforward answers, telling users, for example, about inefficient costing in one operational branch in comparison to others. While such situations might be extremely rare, if not impossible, the role of accounting is conceptualized as a computational practice. Secondly, if uncertainty of cause-and-effect is high but uncertainty of objectives is low, accounting might act as a learning machine. For instance, while organizations might have clear costing targets, they may not know how to achieve them. In such situations, accounting can serve as a trial-and-error tool to learn how to reach these targets. Thirdly, if uncertainty of cause-and-effect relationships is low but uncertainty of objectives is high, accounting might serve as an ammunition machine. Here, accounting provides a medium through which different organizational interests can express their diverging objectives. Budgeting is but one example in this context. The role and design of accounting is not neutral in this process. By influencing the accepted techniques and language of such processes, accounting shapes what can be considered as an acceptable solution. Fourthly, when both dimensions are defined by high degrees of uncertainty, accounting might take the role of a rationalization machine according to Burchell et al. (1980). In these situations, accounting commonly serves as a technique to justify decisions after they have been taken. In other words, accounting becomes an ex post legitimating and rationalization tool.

2.5.1. _Designing and Practicing Accounting Systems without Pre-existing Goals_

Building on the work by Burchell et al. (1980), Cooper, Hayes and Wolf (1981) sought to understand possible roles for ambiguity as a design resource of accounting systems. Cooper et al. (1981) accept Burchell et al.'s (1980) point about the role of accounting in the ex-post rationalization and legitimation of organizational decisions. Yet, they adopt a more radical approach. Drawing from the work of March (1971) and Weick (1969), they take as a starting point that "the sequence whereby actions precede goals may well be a more accurate portrayal of organizational functioning than the more traditional goal-action paradigm" (Cooper et al., 1981, p.181). The relationship between accounting information and organizational goals is hereby seen as only loosely coupled: relationships between problems and solutions are often opaque, as are connections between means and ends, between actions in the present and in the past, and between actions in one part of the organization and actions in another part (see March and Olsen, 1976; March, 1987). In this context, Cooper et al. (1981: 187) state the following:

> Organizations require adaptability, flexibility and the ability to create variety in order to survive in changing, varied environments. In contrast to this flexibility, the organization requires a technology for maintaining some consistency and intelligent behaviour. In terms of effectiveness we wish to emphasize that organizations characterized by ambiguous goals, unclear technology and fluid participation need not be pathological.

As follows from the quote, Cooper et al. (1981) recognize the importance of accounting systems as structuring and ordering technologies that allow for the rationalization of actions, the building of historical organizational narratives and as a technology that holds organizations together. Yet, at the same time, they see organizational objectives as the outcome of action rather than their determining factor, which places demands on accounting systems beyond ex-post legitimation or ammunition machines.

Against this background, Cooper et al. (1981:187) ask: "without a pre-existing goal structure, can there be prescription and design?" Cooper et al. (1981) address this problem by exploring how accounting systems might be designed in an organized anarchy. To frame their argument, Cooper et al. (1981) draw from a variety of principles from the organizational behaviour literature (see March, 1971; Weick 1976). Cooper et al. (1981) offer several suggestions for how accounting systems might be designed to play a constructive role in organizations beyond ex-post legitimation. In particular, they highlight the enabling role of deliberately ambiguous accounting systems and information. Ambiguous accounting information, Cooper et al. (1981) argue, can not only take a crucial role in sensitizing managers to reporting on and addressing uncertainties and inconsistencies, but it can also facilitate playfulness and experimentation. By playing with different interpretive models and frames of reference, past organizational experiences may be re-examined through different lenses and thereby provide a basis for new learning experiences. This prescription draws attention to the importance of being sceptical towards what organizations believe to know and thus elevates the importance of doubt to an important characteristic of adaptable organizations. Furthermore, by having minimal confidence in goals, organizations may embed a search for

new value into their practices and systems, making them more prepared for change and possibly more empowered to manage uncertainties. Stimulating attention and sensitivity to uncertainties and unknowns are regarded as key elements to achieving reflective and creative organizations. Furthermore, Cooper et al. (1981) highlight that ambiguity and playfulness can stimulate the use of accounting systems as a shared organizational language that can help in fostering a sense of belonging, the commitment of participants and co-operation between individuals. While Cooper et al. (1981) recognize the challenge for accounting systems to be a source for organizational disciplining at the same time as stimulating playfulness, their article laid out a tentative agenda that emphasizes the constructive and transformative potential of ambiguity in accounting systems.

The contours of this agenda were further sketched out by March (1987), who focuses on the role of ambiguity, accounting information and organizational decision-making. In particular, March's (1987) article encourages enquiry on accounting and information system design beyond traditional theories of choice and decision-making, which emphasize the tight coupling of goals and actions and underplay ambiguities of choice. The following quote describes several possible elements of such an agenda:

> A system of accounts can be judged in terms of its evocativeness, its power to provide not just confirmation of familiar orders, not just communication of what is known but transformation of what is knowable…It may not be entirely ludicrous to imagine a day when professional students of accounting will discuss the aesthetics and evocative power of ambiguity in a proposed accounting procedure with as much fervour as they exhibit in debating its impact on tax liability…It is, perhaps, a strange vision of information engineering to

say that an accounting report should be a form of poetry, using language of numbers, ledgers, and ratios to extend our horizons and expand our comprehension, rather than simplify fill in the unknowns on a decision tree. But it is not an entirely unworthy vision of professions to say that their accounts and reports can be richer in meaning than they are aware or intend, and that they can enrich our senses of purpose and enlarge our interpretations of our lives (March, 1987, p. 165).

To enhance the evocativeness of accounting reports, March (1987) invokes the analogy of theories of language and literature writing (see also Carruthers, 1995). Theories of language, March (1987) argues, engage with questions about the way language might serve as a mediating vehicle to explore and capture meanings that are not fully grasped by the writer. In line with this, poets might deliberately create ambiguous textural accounts about the mysteries of life to invite others to find alternative interpretations in the texts that had not been consciously envisioned by the creator. March (1987), acknowledging the unconventional nature of his proposition at the time, challenges researchers to further theoretically advance implications from these insights for the design and practice of accounting systems.

2.5.2. _The Performativity of Incomplete Measures, Interpretative Reflexivity and the Power of Ambiguity_

Following these initial studies by Cooper et al. (1981) and March (1987), interest in the role of incompleteness and ambiguity re-emerged as an important theoretical concern for accounting scholars in recent years. Dambrin and Robson (2011) advanced this line of reasoning by investigating the

conditions facilitating the performativity of incomplete accounting measures. Dambrin and Robson's (2011) study is situated in the French pharmaceutical industry. A defining element of this industry is that legislation prevents access by pharmaceutical companies to data on doctors' prescriptions that generate sales of drugs. The legislation thus significantly complicates the performance evaluation for bonus payments of the pharmaceutical sales representatives as it leads to an interruption in the chain of circulating references (see Latour, 1999). In this context, Dambrin and Robson (2011: 446) find that methodological opacity in relation to the calculation of performance measures was in fact conducive to their practice:

> Methodological opacity occludes the weak tracing of the relation between sales data and work. Rather than stating that 'because there is faith, problematic and obscure calculations can last', we claim that 'because there is opacity in calculations, people appear to trust in them.'

While methodological opacity formed an important element in the establishment of trust in the imperfect measures, Dambrin and Robson (2011) further identify bricolage, which includes resourcefulness and improvisation, as an important element in the performativity of the broken inscriptions. Managers are described to 'make do' with the imperfect measures by combining them with estimates, observations, self-reporting or standard costs that provide a partial but acceptable proxy for sales performance.

In line with such findings, Jordan and Messner (2012) investigate the problem of incomplete performance indicators within the context of management control systems through a study of the introduction of Lean Six Sigma. Jordan

and Messner (2012) show how the perceived incompleteness of the performance measures did not automatically constitute a problem for operational managers if they were given flexibility in relation to how they could use the indicators to inform their work. If the indicators were treated as means rather than ends, they were likely to be perceived as enabling. Jordan and Messner (2012) point out however that such flexibility is challenging to sustain if senior managers continuously emphasize the importance of attaining the indicators, thus treating them as ends. In such a situation, incompleteness of performance measures was perceived to be problematic by operations managers. As performance measures were now perceived to be coercive rather than enabling, managers now attempted to "repair" them to address their incompleteness. These studies highlight how the opacity of incomplete performance measures can stimulate organizing work and thereby become performable despite their incompleteness (see also Jørgensen and Messner, 2010).

Chenhall et al. (2013) further focus on the question of how the design and operation of performance measurement systems might facilitate compromise between competing principles for evaluation. To develop their argument, Chenhall et al. (2013) propose the concept of 'compromising accounts,' which consists of three elements: (1) concurrent visibility, namely the form through which evaluative principles that are considered important by different organizational groups are made visible; (2) imperfection, which relates to the extent to which the design of the performance measurement system allows for a give-and-take between the different evaluative principles; and (3) the types of responses that emerge in situations when compromise is threatened. Accordingly, if criticisms of the prevailing evaluation criteria are based on

principled arguments rather than personal considerations, a breakdown in compromise or a prolonged conflict might be prevented. Chenhall et al.'s (2013) study provides a contribution to the design and role of accounting formats in the development of compromise to enhance the richness of accounting as advocated by Cooper et al. (1981) and March (1987). Yet, Chenhall et al.'s (2013) concept of compromising accounts is more about co-existence than about innovative or productive re-combination. In Chenhall et al.'s (2013) study, evaluation principles pre-date and remain fixed throughout the process through which the compromising account is developed. Compromise, in the way conceptualized by Chenhall et al., (2013), implies a focus on the virtues of existing principles of evaluation rather than prompting reflection on other future options and possibilities.

Quattrone (2009; 2015a) and Busco and Quattrone (2015) offer further theoretical contributions to this strand of literature by taking the incomplete representational abilities of accounting technologies as a starting point for the conceptualization of valuation practices in conditions of uncertainty. In his historical case study of accounting texts, Quattrone (2009) theorizes accounting as an orthopraxis, i.e. "a type of knowledge which is inextricably linked to the way in which the space between the text and the use is filled with specific forms of practice." Quattrone (2009) argues that accounting becomes performable, a working space and time, by drawing attention to the combination of four factors: (1) the medium through which it is constructed and practiced (the book); (2) its power as a tool for the visual organization of knowledge; (3) its 'orthopraxis' nature; and (4) the procedural flexibility of its method based on the variety of possible ways to segment and recompose accounts. Quattrone (2009) hereby theorizes accounting as an ever-unfolding

technology to interrogate, question and invent new knowledge and thereby inspire doubt about what is valuable (see also Quattrone, 2015a; Busco and Quattrone, 2015). Thus rather than theorizing the emergent nature of accounting primarily in relation to exogenous elements, such as human agency or contingent and contextual factors, these studies focus on dynamic features related to its design, visual features and method.

With the aim to explore how the design of specific accounting inscriptions can be implicated in mediating conceptual ambiguity, Qu and Cooper (2011: 347) further emphasize the importance of the generative nature of interpretative flexibility: "Interpretative flexibility allows different groups to *conceive* their own version of a technique to make it useful in their particular circumstances". According to Qu and Cooper (2011), stimulating interpretative flexibility is thus not an element that should be regarded as undesirable in the construction of inscriptions but instead it should be embraced as an important element to hold together networks of actants with diverging agendas.

Against this background, Quattrone (2015b) argues that accounting could be re-thought as a maieutic[1] machine, prompting users to seek discovery by asking the right questions rather than serving as an 'answer machine' that provides fixed answers or an instrument that serves the purpose of legitimating decisions in an ex post manner (Burchell et al., 1980). Quattrone's (2015b) analysis hereby focuses on theorizing the key features of a maieutic accounting machine, which are specified as a visual performable space, a method of ordering and inventing knowledge, a means to mediate between

---

[1] Maieutic is a Socratic method for philosophical enquiry which uses dialogue as a tool to seek discovery

heterogeneous interests and a motivating ritual. Quattrone (2015b:54) emphasizes the following:

> When calculating values, accounting should embrace an uncomfortable doubt rather than searching for a reassuring certainty. Accounting, rather than giving answers and providing comforting values to be taken as the starting point of decision-making, needs to prompt instead questions on what is not known, on the rationales and rationalities other than the economic ones.

In their recent study on the design and role of accounting systems in a large Italian infrastructure project, Revellino and Mouritsen (2015) combine Quattrone's (2015b) maieutic machine with Miller and O′Leary's (2007) concept of mediating instrument. In their study, Revellino and Mouritsen (2015) focus on the role of the Chronoprogram, a management accounting tool, which links issues surrounding major capital investments with a multiplicity of other dislocated and competing interests on a road map. Revellino and Mouritsen's (2015) specific emphasis lies on the design of the Chronoprogram and how this design is implicated in engaging and mediating between multiple stakeholders and reaches compromise between them. To theorize the mediating instrument as a maieutic machine, Revellino and Mouritsen (2015) mobilize Latour's (2005) dialectic between matters of fact and matters of concern. While matters of fact relate to notions of indisputability, perceived objectivity and closure, matters of concern emphasize openness, questioning and doubt. In Revellino and Mouritsen's (2015) view, generating matters of fact as well as matters of concern is an important element for accounting technologies. The construction of matters of fact allow users to close debates and engage in pragmatic and practical action. At the same time, to be maieutic,

accounting systems should be conceived in a way that inscribes criticism and doubt, in other words matters of concern, into their format and design:

> [Accounting] will be judged by its involvement in inscribing scripts into the design object that defines new and surprising matters of concern, surprising object candidates and their links in the theatre of the world (Revellino and Mouritsen, 2015, p.14).

In Revellino and Mouritsen's (2015) study, the dialectic between matters of fact and matters of concern is achieved by relying on different accounting techniques. While Net Present Value calculations produce matters of fact in capital budgeting, interaction with the Chronoprogram produces matters of concern in relation to the infrastructure project's effects on the natural environment, communities, politics and other unexpected issues. By relying on Latour's (2005) dialectic between matters of fact and matters of concern, Revellino and Mouritsen (2015) thus add a further theorization of the role of ambiguity as a design resource for accounting systems.

Combined, these studies provide useful insights in relation to some of the key theoretical and practical issues facing humanitarian performance evaluation systems. In section 2.4.2, it was argued that instead of focusing on the question how accounting systems might be implicated in the attempt to normalize disaster situations, the more promising question is how accounting systems might be designed and practiced to facilitate engagement with anomalies, ambiguities and the unexpected. While the literature reviewed in this section has begun to further advance knowledge in relation to these concerns, relatively few studies have since followed up on the core agenda set by Cooper

et al. (1981) and March (1987), which as a result remains under-researched and under-theorized. Furthermore, despite recent calls for further investigation (see Walker, 2016) the role of accounting systems in the engagement with extreme environments such as humanitarian crises has so far received little attention in the literature (see Sargiacomo, 2015; Walker, 2014; Taylor et al., 2014). This dissertation follows up on the invitation provided by the above studies. It explores the question how evaluation systems can be designed and practiced to evaluate actions in humanitarian environments defined by high degrees of uncertainty, complexity and continuous change.

2.6. <u>Concluding Remarks</u>

The chapter began by situating the study within interdisciplinary research that has explored the agency of accounting technologies in constituting the social. This offered context and background for the subsequent sections of the chapter (see Burchell et al., 1980; Hopwood, 1985, 1987; Chapman, Cooper and Miller, 2009). The second part of the literature review discussed a rich body of research that has theorized the role of accounting as a technology for social control, which has been influential in the emerging accounting literature on humanitarian crises. In this context, the review discussed several elements regarding the relationship between accounting and social control, including the question of how accounting can shape the 'conduct of conduct' (Miller, 2008), construct identities of the marginalized (Walker, 2008), enable action at a distance (Robson, 1991, 1992), and constitute governable spaces (Miller and Power, 2013). Building on these insights, the third section of the review showed how this literature has been influential in encouraging interest in

exploring possible roles of accounting technologies in humanitarian crisis settings. In particular, this emerging literature focused on how accounting classifications may constitute a crisis intervention space and can be used to re-assert control by government agencies that act from a distance (see Sargiacomo, 2015; Sargiacomo et al. 2014; Everett and Friessen, 2010). Against the background of these studies, the literature review has indicated that scholarly understanding remains underdeveloped in relation to a particular theoretical issue of accounting systems in humanitarian crisis settings. Rather than attempting to enforce pre-defined programmes of control that will quickly break under the complexity and blind spots of crises (see Weick and Sutcliffe, 2007, 2015), a key challenge relates to the question of how accounting systems can embrace the ambiguity of such settings and make a resource of it (March, 1987; Cooper et al, 1981; Mouritsen, 2016). To conceptualize these concerns, the fourth section of the review engaged with literature that explores the power of ambiguity and incompleteness in the design and practice of accounting systems (see Dambrin and Robson, 2011; Jordan and Messner, 2012; Quattrone, 2015a, b). In this context, it was argued that the key point of making ambiguity a resource is not to renounce any form of measurement and calculation. Instead, the research seeks to investigate different forms of engaging with performance evaluation systems that enable a reflective and flexible manner of working with measures to enhance and transform what is knowable (March, 1987). The current study seeks to contribute to this debate by exploring how performance evaluation systems can be implicated in the engagement with the unexpected during the management of humanitarian crises.

To further conceptualize how ambiguity might become a resource for the design and practice of humanitarian performance evaluation systems, in the following chapter the study introduces insights from recent developments in the Sociology of Worth and more particularly the concept of heterarchy as advanced by Stark (2009) and Lamont (2012). It is to these issues that the thesis now turns.

3. The Sociology of Worth: Evaluation, Uncertainty and Governance

3.1. <u>Introduction and Overview</u>

As outlined in the previous chapters, this study investigates the question of how evaluation systems can be designed and practiced to facilitate engagement with anomalies, ambiguity and the unexpected in the management of humanitarian crises. To investigate and conceptualize how ambiguity and incompleteness might be considered a resource for humanitarian performance evaluation systems, this study takes inspiration from the Sociology of Worth (SOW), and more specifically from the concept of heterarchy (Stark, 2009; Lamont, 2012). As a form of governance, heterarchy encourages the organization of diverse and even dissonant principles of evaluation. While dissonance is commonly connected to harmful and negative forces, the notion of heterarchy highlights how dissonance might be considered a resource by regularly cultivating and advancing ambiguous spaces that enable reflection and challenge expectations that commonly emerge in the management of complex projects and situations of crisis (see also Weick and Sutcliffe, 2007; March, 1981, 2006). Drawing from the notion of heterarchy as a mode of organising implies the recognition that incomplete categorisations can foster contradiction, tension and re-association (Stark, 2009; Lamont, 2012).

To explore and discuss these issues, the chapter is structured around three key sections. Firstly, section 3.2 introduces recent discussions surrounding the limitations of many conventional performance evaluation practices and technologies in uncertain and quickly changing environments. In such contexts, organizations are confronted with a challenging type of enquiry that does not centre on identifying solutions to clearly identified problems and therefore cannot be resolved by strictly gathering information on pre-identified evaluation categories (Stark, 2009; March, 1981; Weick and Sutcliffe, 2007). Against this background, the section discusses the possibility of drawing from multiple grammars of value as an important enabling factor for the adaptive capacities of organizations (Stark, 2009; Lamont, 2012; Gehman et al., 2013; Berthoin-Antal et al, 2015). Secondly, section 3.3 reviews developments in SOW research challenging a widespread focus on strictly economic forms of evaluation. In this context, the section analyses important attempts to develop theoretical alternatives to this reductionist view, amongst them the work of Boltanski and Thevenot (2006). While their important contributions are recognized, it is argued that there is a need to move beyond Boltanski and Thevenot's (2006) framework to further address some of the key concerns of this study. In particular, the analysis shows that their framework paints an overly fixed picture of social systems and underemphasizes the explanation of dynamic change and the possibility of adaptation. Thirdly, in line with these insights, section 3.4 discusses recent interest in the concept of heterarchy (Stark, 2009; Lamont, 2012). The concept of heterarchy provides important insights on a distinct notion of engaging with heterogeneous forms of evaluation in contexts of uncertainty and complexity, which constitutes a key challenge in the management of humanitarian response operations

(Rottenburg, 2009; Barnett, 2011). The section is concluded by arguing that despite its promising potential, scholarly understanding in relation to principles and conditions that might foster or hinder the practice of heterarchies is limited (see Lamont, 2012).

3.2. <u>Uncertainty, Risk and Evaluation</u>

In the opening line of *The Sense of Dissonance*, Stark (2009:1) introduces the central orienting principle of his approach to SOW research: "Search is the watchword of the information age." In line with this observation, processes of organizing search and inquiry lie at the heart of his interest. While mechanical engines were central to the industrial age, performance metrics and algorithms facilitate the management and analytics of data in today's information age. The increasing spread of performance measures are one example of how technologies do not only facilitate information management in organizations, but how these technologies themselves assert agency in regards to what is perceived as valuable and worthy in society (Miller and Power, 2013; Kornberger et al., 2015; Quattrone, 2015b). However, despite the increasing influence of such information and evaluation technologies, somewhat paradoxically, these technologies have considerable limitations in addressing one of the crucial problems facing organizations operating in contexts of uncertainty (Berthoin-Antal et. al, 2015). Paraphrasing Stark (2009), a fundamental challenge in situations of uncertainty is organizing a type of search during which you do not know what you are looking for but will recognize it when you find it. This kind of search leads to a more challenging type of enquiry which does not centre on identifying solutions to clearly

identified problems. As a result, it cannot be resolved by strictly gathering information on pre-identified evaluation categories and metrics (see also March, 1981). As emphasized by Holm (2010:334) in a commentary on Stark's (2009) work:

> Whereas the archive could serve as a metaphor for the bureaucratic logic of the past, the search engine is a fitting metaphor for a new age in which worth increasingly is created by combining things that used to be kept apart.

Accordingly, while in relatively stable environments organizations can develop metrics, rules and conventions to make their actions more efficient and gain legitimacy (see DiMaggio and Powell, 1983), in uncertain and quickly changing environments, such as humanitarian crises, adhering to these conventions can rapidly become dated at best, and at worst prevent organizations from recognizing challenges and opportunities. Hence, even though performance evaluation systems might help provide a sense of orderliness to organizational practices, they also risk embedding expectations and biases that make actors blind to the unexpected (see also Garud et al., 2008; Weick and Sutcliffe, 2007).

In this context, it is useful to draw a distinction between the notions of 'risk' and 'uncertainty', which share similarities but are not the same. An insightful distinction between the two notions is attributed to the early work of Knight (1921). Risk, Knight (1921) argues, while also shaped by the fact that the future is unknown, relates to situations in which a probability can be placed regarding the likelihood that an expected event will occur (see also Weick and

Sutcliffe, 2007). In other words, the concept of risk relates to situations where organizations do not know the outcome but can calculate odds. Uncertainty in turn lacks these probabilistic qualities; it involves a multiplicity of unknown factors that make calculation impossible (see also Apandurai, 2011; Stark, 2009; Power, 2004).

In both accounting research and practice, the conventional approach to performance measurement, planning and decision-making in uncertain environments is to frame challenges in terms of pre-specified risk and performance metrics (Power, 2007). However, in the context of humanitarian crises, it is precisely the blind spots that elude calculation that hold significant potential for disastrous errors, which therefore warrant a more serious engagement with the notion of Knightian uncertainty. A quick look at the etymology of the term 'crisis' further helps to illustrate the point. 'Crisis' derives from the Greek word 'krisis', which means decision. In late Middle English, the term became associated with the turning point in a disease in which a difficult decision of great consequence must be made, the implications of which are not foreseeable (Oxford Dictionary, 2016a). Engaging with crises thus entails dealing with uncertainty in relation to information, actions, goals and outcomes. Just as in the type of search identified by Stark (2009), in situations of complex humanitarian disasters, problems are often ill defined, relationships between elements and variables are opaque and value criteria are multiple and competing (Barnett, 2011).

Against this background, Stark (2009:6) questions the common contention that under conditions of uncertainty, stability of organizational value principles and metrics is a key criterion for adaptability and reflexivity:

> Where the organizational environment is turbulent and there is uncertainty about what might constitute a resource under changed conditions, contending frameworks of value can themselves be a valuable organizational resource.

As Stark (2009) emphasises, at an elementary level perplexing situations that encourage reflection are created when there is principled disagreement about what should be recognized as valuable. Instead of merely responding to external demands, organizations might thus pursue a more radical and challenging strategy to regularly foster such confusing and ambiguous spaces that enable reflection and challenge expectations that commonly emerge in the management of complex projects and situations of crisis (see also Callon et. al, 2009). Hence, instead of attempting to impose a sole criterion for evaluation as the key performance objective, organizations might pro-actively generate perplexing situations in which there are multiple principles and legitimate ways to conceptualize notions of value (Stark, 2009). To conceptualize the challenges of such an approach of drawing from diverse notions of value to foster reflective capacities, the following section discusses recent developments in SOW research that focus on moving beyond a reductionist way of theorizing notions of value.

3.3. <u>Embracing a Plurality of Values</u>

Stark's (2009) argument that at an elementary level organizational reflection is created at the confusing point where different principles of evaluation overlap directly situates his research in the emerging field of SOW. Within the sociological discipline, SOW is the specialized field that engages with organizational and societal questions of value and processes of evaluation (see also Lamont, 2012; Helgesson & Muniesa, 2013; Kornberger et al., 2015). In basic terms, SOW research is concerned with addressing several simple but highly consequential questions in relation to some key social problems facing modern societies: What counts as valuable and how is value constituted? How can questions of value be settled when incommensurable frameworks of valuation are at play? Which mechanisms, tools or cultural repertoires are needed to ensure that a larger number of people is considered valuable in modern societies? How are classificatory processes surrounding evaluation practices linked to dominant definitions of citizenship, solidarity with the poor or marginalized, distribution of welfare resources, racism or xenophobia? (see Holm, 2010; Karpik, 2010; Stark, 2009; Lamont, 2012). As Sandel (2012:7) highlights in his analysis of the increasing spread of market-based norms of conceptualizing value, exploring new answers to these questions is timely: "the reach of markets, and market-oriented thinking, into aspects traditionally governed by nonmarket norms is one of the most significant developments of our time." While issues of value are commonly treated as matters of (economic) utility that propose to converge and quantify heterogeneous concerns into a single common measure, SOW research challenges the widespread focus on strictly economic forms of evaluation and seeks to

develop theoretical alternatives to this reductionist view (Stark, 2009; Annisette and Richardson, 2011; Lamont, 2012).

The reductionist perspective on notions of value can also be traced to one of the most influential paradigms in social science research over the last several decades, namely Parsons' Pact (see Stark, 2009; Holm, 2010; Kornberger et al., 2015). Originating from an informal agreement between Harvard's department of economics and sociology, which was led by Talcott Parsons in the 1940s, Parsons' Pact delineated the spheres of influence and expertise between these academic disciplines. In the plans of Parsons, economists would confine themselves to the study of value, whereas sociologists and political scientists would focus their attention on values and the social relations within which economies are performed. While simple in its prescriptions, Parsons' Pact was influential in the development and practice of the modern disciplines of economics on the one hand and sociology, anthropology and political science on the other. By using the concept of worth, SOW research seeks to break with Parsons' Pact. With its simultaneous connotation of economic as well as moral good, the notion of worth merges value and values and thereby challenges the distinction between rational calculation and moral judgement that is implicit in the value-values dichotomy (see also Fourcade and Healy 2007, Shapin 2012). As emphasized by Stark (2009:7):

> The polysemic character of the term – *worth* – signals concern with fundamental problems of value while recognizing that all economies have a moral component. Rather than the static fixture of value and values, it focuses instead on ongoing processes of *valuation*.

One influential attempt to develop an alternative sociological theory of value beyond Parson's Pact is provided by the work of Boltanski and Thevenot (2006) in *On Justification: The Economies of Worth.* By focusing their empirical and analytical attention on everyday situations, Boltanski and Thevenot (2006) demonstrate that social coordination is an ongoing achievement that is shaped by multiple principles of justifying what is valuable. Furthermore, they show that some notions of value are incommensurable and cannot be subjected to comparison by translating them into a common measure (see also Karpik, 2010). In line with this theoretical focus, Boltanski and Thevenot (2006) seek to reconstruct principles of value that meet conditions of generality:

> Our approach to the coordination of human behaviour led us to pay attention to the cognitive ability that allows human beings to establish associations among things that count, to identify beings independently of circumstances, and reach agreement on forms of generality…Among the infinite number of possible associations, we shall be concerned only with those…that can be supported through justifications (Boltanski and Thevenot, 2006, p.32).

Based on the analysis of a variety of classical writings in political philosophy, Boltanski and Thevenot (2006) establish six different orders of worth, each of which defines the good and the just, e.g. the worthy, with different grammars of conceptualizing notions of value. As Annisette and Richardson (2011: 231/232) stress, "an order of worth can be thought of as a hypothetical model of a good society constructed on a singular basis of merit that acts as the sole standard for determining what matters or what is worthy within that hypothesized society". In this context, Boltanski and Thevenot (2006) outline

the following orders of worth, which are based on civic[2], industrial[3], domestic[4], fame[5], inspired[6] and market[7] conceptualizations of the common good. For example, in the domestic world worth might be embodied through notions such as tradition or hierarchy, which contrasts with a civic order of worth, where the higher common principle is civic duty. Each order of worth thus has its principles of constituting the worth of people and things with its own measures of equivalence. For Boltanski and Thevenot (2006), orders of worth are social technologies that offer multiple rationalities on how notions of value might be conceptualized and justified during processes of social coordination. While each order has its own rationality, it also has its own way of legitimating the greater good, and thus offers its own morality and notion of moral justice. Boltanski and Thevenot's (2006) work proved to be insightful for the emerging field of SOW research, in particular in relation to their propositions on the plurality of orders of worth, their amalgamation of value and values and their insights on the relationship between rational calculation and moral judgements (see also Kornberger et al., 2015; Berthoin-Antal et al., 2015).

Despite their important contributions to SOW research, to further address some of the key concerns of this study, there is a need to move beyond Boltanski and Thevenot's (2006) framework. The first section of this chapter raised the argument that under conditions of uncertainty and dynamic change, organizations commonly confront a challenging type of search that does not centre on identifying solutions to clearly identified problems and therefore

---

[2]Exemplified by Rousseau's (1762) *Social Contract*
[3]Exemplified by Saint-Simon's (1819) *Du Systeme Industriel*
[4]Exemplified by Bossuet's (1709) *Politics Drawn from the Very Words of Holy Scripture*
[5]Exemplified by Hobbes' (1651) *Leviathan*
[6]Exemplified by St. Augustine's *City of God*
[7]Exemplified by Smith's (1776) *Wealth of Nations*

cannot be resolved by strictly gathering information on pre-identified evaluation categories. In this context, the section highlighted the importance of concepts such as doubt, adaptability, organized exploration, ambiguity and dissonance. None of these issues form key concerns in Boltanski and Thevenot's (2006) framework. The somewhat fixed nature of Boltanski and Thevenot's (2006) orders of worth provides a relatively static conceptualization regarding the way social relations might be constituted and constructed. Social actors can either settle disputes within one order of worth or attempt to forge a compromise between the orders. As a result, little space is left for questioning, creativity and innovative re-combination of notions of value. Against this background, Stark (2009:15) highlights:

> In Boltanski and Thevenot's framework…conventions (of which orders of worth are a particularly well-elaborated variant) are a way of dealing with uncertainty. They are engines for turning situations into calculative problems…The limitation of this view…is that it does not give adequate attention to the problem that orders of worth cannot eliminate uncertainty. In particular, they cannot eliminate the possibility of uncertainty about which order or convention is operative in a given situation.

To address problems of uncertainty, Boltanski and Thevenot's (2006) relatively static framework therefore has several crucial shortcomings. Firstly, if social actors always draw from the same templates, e.g. fixed orders of worth, why would there be a problem of uncertainty to begin with? In other words, Boltanski and Thevenot's (2006) approach risks theorizing away the key issue it sets out to address. Secondly, in line with this point, the assumption that social actors always orient themselves towards pre-existing orders of worth paints an overly fixed picture of social systems and underemphasizes the

explanation of dynamic change (see also Beckert, 2011). Thirdly, Boltanski and Thevenot's (2006) insights remain at a relatively abstract level. Even though they call for attention to be drawn to the technologies through which notions of worth are disputed, few insights are provided on how such technologies, for example accounting technologies, shape processes of evaluation.

To move beyond these shortcomings, the discussion proceeds by introducing an important concept, namely heterarchy. The concept of heterarchy provides promising perspectives on a distinct notion of engaging with heterogeneous forms of valuation in contexts of uncertainty and complexity (Stark, 2009; Lamont, 2012), which constitutes a key challenge in the management of humanitarian response operations (Rottenburg, 2009; Barnett, 2011).

3.4. <u>Heterarchy: Governance through Difference</u>

In recent years, the concept of heterarchy has been receiving increasing attention in SOW research (Stark, 2009; Lamont, 2012). As highlighted by Lamont (2012:202):

> With growing income inequality and the trend towards a 'winner-takes-all-society', understanding the dynamics that work in favour of, and against, the existence of multiple hierarchies of worth or systems of evaluation (i.e. heterarchies or plurarchies) is more urgent than ever. Indeed, the coexistence of multiple matrices of evaluation is one significant condition for greater social resilience…This grounds the social significance of gaining a better understanding of the processes that sustain heterarchies.

The etymology of heterarchy can provide first insights into some of its implied propositions. Heter-archy derives from two Greek words, namely 'heteros' which translates as 'difference' or 'other' and 'archein', which means 'to govern' or 'to rule'. In colloquial terms, heterarchy can therefore be translated as the 'governance of difference', or even more boldly as 'governance through difference.' The term was first introduced by McCulloch (1945) in his paper *A Heterarchy of Values Determined by the Topology of Nervous Nets*. Informed by his interest in neurology and epistemology, McCulloch's central research focus was on processes of choice-making. By simulating networks of neurons, McCulloch highlights that cognitive processes in the human brain, while orderly, do not work according to hierarchical organizing principles (Crumley, 1995). To theorize an alternative cognitive structure, McCulloch introduced the term heterarchy as a metaphor to conceptualize nervous systems defined by inter-sensitivity.

Since its introduction, the concept of heterarchy has travelled into economic sociology and organization theory. As stressed by Stark (2009), as a form of governance, heterarchy encourages the organization of heterogeneous and even dissonant principles of evaluation and legitimation. While dissonance is commonly associated with destructive forces, and organizations commonly aim for the smoothest possible alignment between their objectives, activities, and evaluation criteria, the concept of heterarchy highlights how dissonance might be considered a resource. A simple illustrative example can help to further clarify this point. As Stark (2009) emphasises, the separation of powers as a crucial design principle of modern government constitutes perhaps one of the oldest and most widely known examples of heterarchical organizing.

Each of the three branches of government, the executive, the legislature and the judiciary, is not only based on a distinct principle of legitimation, but also on the assumption that none of them is superior to the other. Through the interactions, checks and frictions between these branches of government many of the ever-evolving challenges of modern society can be continuously negotiated. As Crumley (1995:3) put it, "heterarchy may be defined as the relations of elements to one another when they are unranked or when they possess the potential for being ranked in a number of different ways." While hierarchies can temporarily emerge within heterarchical systems, importantly no category or ranking is ever complete, all-encompassing or permanent. Ambivalence and even contradiction are thus common features of heterarchies (Crumley, 1995; Stark, 2009).

While the above examples might appear to be relatively abstracted from evaluation practices, the concept of heterarchy provides a fruitful metaphor for SOW research. Rather than strictly separating different value spheres, and instead of focusing on forging compromise between them (Chenhall et al., 2013), heterarchies invite the cultivation of processes of search, enquiry and the recombination of evaluative principles. From this perspective, it is at the point where different evaluative principles overlap that an opportunity for reflexivity is created that might disrupt organizational taken-for-granteds and generate new insights. In Stark's (2009:19/25) terms, heterarchies can be distinguished from more hierarchical forms of organizing along the following lines:

> In contrast to hierarchies, heterarchies are characterized by more crosscutting network structures, reflecting the greater

interdependencies of complex collaboration. They are heterarchical, moreover, because there is no hierarchical ordering of competing evaluative principles… Because resources are not fixed in one system of interpretation but can exist in several, heterarchies make assets of ambiguity.

In line with the quote, this approach thus brings the enabling role of ambiguity into SOW research. Stark (2009) provides a variety of examples from his ethnographic case studies, including on financial arbitrage trading, tech-start-ups and a Hungarian tool factory in the transition from communism to capitalism. The ethnographies provide insights into the way resources and notions of value are re-combined and re-formatted in conditions of economic pressure and uncertainty. As emphasized by Gehman et al., (2013:106), such a distinct form of value governance might be "at once a source of resilience and fragility, depending on what has been knotted together, and what becomes unknotted at any given point in time… values practices emerge not as a kind of terra firma on which an organization's governance might rest, but as one more aspect of organizing that is itself in need of governance." Recent research shows how such a form of value governance might itself be an important factor for organizational resilience (Stark, 2009; Gehman et al., 2013; Berthoin Antal et al., 2015).

One of the evident connections of Stark's framing of the problem with the accounting literature reviewed previously is the emphasis on proactively creating spaces for ambiguity that do not only break familiar routines (March, 1987), but also help detect inconsistencies (Cooper et al., 1981), foster doubt and criticism (Quattrone, 2015b) and integrate knowledge across diverse fields of expertise (Revellino and Mouritsen, 2015). Against this background, the

notion of heterarchy provides an informative approach to engaging with a multiplicity of orders of worth that differs from Boltanski and Thevenot's (2006) focus on how to forge stability and compromise between different orders (see also Chenhall et al., 2013).

Yet, as stressed by Lamont (2012) scholarly understanding remains underdeveloped in relation to the conditions, principles and technologies that might work in favour of, or in contradiction to, the practice of heterarchies of values. While diverse principles of evaluation are probably a common element in most organizations, the question of how heterarchical principles might be productively fostered emerges from Stark's (2009) ethnographies. Did heterarchies simply emerge as an outcome of contextual factors and uncertainties? How is coordination achieved despite the dissonance that defines heterarchical structures? Which elements sustain or obstruct heterarchies? However, raising these questions and inviting further research on them appears to be consistent with Stark's (2009) eclectic approach to social theory. As has been noted in another commentary on his work: "As a theoretician, Stark is a director, not an engineer" (Rona-Tas, 2011, p. 598). Unlike many social theories that aspire to utmost consistency and explanatory power, Stark's (2009) work focuses on several key themes, amongst them dissonance, heterarchy, search, exploration and orders of worth and investigates the interrelationship between them. Hence, while taking insights from Stark's (2009) conceptual and methodological approach, his work provides not only fruitful insights, but, importantly, it also raises many questions to explore for this dissertation; in particular, an underdeveloped theoretical understanding in relation to elements that sustain or inhibit the practice of heterarchies.

In line with this theoretical interest in notions of heterarchy, search, exploration and the interplay between different orders of worth, Stark (2009) also frequently highlights the importance of paying attention to the concept of accounts and its etymological richness. In this context, he emphasizes the following:

> When authority is distributed along lines of lateral accountability, we need to study those who make and keep accounts (and who, most emphatically, are not simply accountants) … etymologically rich, the term simultaneously connotes bookkeeping and narration. Both dimensions entail evaluative judgements, and each imply the other. Accountants prepare story lines according to established formulas, and in the accounts given by a good storyteller we know what counts. In organizations, as in everyday life, we are all bookkeepers and storytellers…It is always within accounts that we 'size up' situations, for not every asset is in a form mobilizable for a given situation (Stark, 2009, p.25).

However, despite the fact that it is frequently mentioned and alluded to, the role of the central valuation and evaluation devices per se, accounting technologies, is largely left unexplored in his research. In other words, while stressing their agential nature, Stark (2009) says less on the way the format, design and practice of specific accounting devices can bring to the forefront how notions of value are conceptualized, visualized, discussed and organized and how they can inform how heterarchies are enacted within organizations and society. By building on these insights, with its key concern with evaluative technologies in the management of humanitarian crises, this study further investigates the theoretical possibilities and implications of heterarchies for accounting systems.

3.5. <u>Concluding Remarks</u>

In relation to the central research concerns of this study, several implications can be summarized at this point. The study investigates principles and tactics that allow humanitarian evaluation systems to make a resource of the inevitable ambiguity and incompleteness that defines their context. The concept of heterarchy, with its focus on 'governing through difference', provides promising insights into a distinct approach for engaging with divergent forms of evaluation in contexts of uncertainty and complexity. Heterarchical systems are a conceptual response emerging from a scepticism towards top-down designs that pre-specify ideal solutions and strive towards a more 'accurate', 'true' or 'fair' manner to account for and represent performance. Instead, heterarchical systems shift the focus on processes that seek to institutionalize forms of criticism and competition, foster checks and balances, and recursively create spaces of ambiguity and reflection. In this sense, SOW research provides a viable theoretical alternative to the reductionism that commonly dominates strictly utilitarian conceptualizations of value (Stark, 2009; Boltanski and Thevenot, 2006; Kornberger et al., 2015). However, as emerges from the above discussion, a conceptual and theoretical understanding in relation to principles and conditions that might foster or hinder the practice of heterarchies of value is limited (Lamont, 2012). When combined, the insights from this discussion provide an approach to investigate the design and practice of humanitarian performance evaluation systems that not only have to confront multiple notions of value but must also enable reflective decision-making in environments defined by dynamic change, complexity and uncertainty.

4. Methodology

4.1. Introduction

The previous chapters have presented the research problem and have situated the study within a specific accounting literature and theoretical approach. In other words, the chapters have specified why the research matters and what the study seeks to investigate. Against this background, this chapter discusses how the study was carried out. It seeks to explain how the various methodological and philosophical choices that inform the investigation are not only coherent but also appropriate for the central questions the research aims to explore.

The chapter is structured as follows. In section 4.2, the philosophical considerations that inform and underlie the study's approach to social science research are outlined. These considerations include ontological assumptions about the nature and existence of entities that form part of the investigation. These ontological considerations are described in section 4.2.1. Building on this discussion, in section 4.2.2 the study's approach to the development and conceptualization of knowledge, its epistemology, is discussed. Both sections justify the rationales for the ontological and epistemological assumptions and discuss how they mutually build on each other. In section 4.3, the chapter continues by discussing the overall research approach and the methods that

were employed to carry out the investigation. It begins in section 4.3.1 by laying out the reasons why qualitative case study research was selected as the appropriate approach. In this context, the section discusses commonly described strengths and weaknesses of this approach in relation to scientific research. Subsequently, section 4.3.2 provides justifications regarding the selection of the three case studies that were carried out as part of this investigation. Furthermore, it offers some general background information on the case studies. Section 4.3.3 describes the different methods that were used to collect empirical data. These included semi-structured interviews, observations and document analysis. It also explains how the selection of these data collection methods is consistent with the study's research design and philosophical approach. Section 4.3.4 then outlines how the collected data was organized and analysed. The chapter finishes with concluding remarks in section 4.3.5.

4.2. Philosophical Considerations and Assumptions

4.2.1. *Ontological Considerations*

To introduce the meaning and significance of ontology for the purpose of scientific inquiry, going back to its etymology can give an insightful starting point. According to the Oxford Dictionary (2016), 'ontology' derives from two ancient Greek words, namely 'ontos', which means being or existence, and 'logos' which translates as discourse or reason. Ontology is concerned with theories and assumptions regarding the existence, nature and becoming of

entities (Woolgar, 1988). While constituting a separate theoretical domain of philosophy (metaphysics), issues concerning ontology are fundamental in relation to the approach towards research and are closely related to the fundamental assumptions underlying any type of scientific inquiry. It is therefore necessary to clarify and justify the ontological assumptions that inform this study's approach (see Burrell & Morgan, 1979), before laying out in detail its epistemology, research design and methodology in subsequent sections.

As emerges from the literature review chapter, this study is firmly situated within a strand of research that can be broadly labelled as 'accounting as a social practice', which emerged as a critique of a largely functionalist and positivist approach that had dominated accounting scholarship for decades (see Willmott, 1983; Chua, 1986; Hopper et al., 1987). As Hopwood (1978: 189, 190) highlights in an early editorial of *Accounting, Organizations and Society*:

> Rather than accepting particular notions of the accounting domain, [researchers] are trying to base their understandings on wider appreciations of the social, economic and institutional context of accounting. Striving for a view *of* accounting, in action, they seek to understand the nature of the social and the organizational interests *in* accounting, the institutional processes through which such interests are articulated and the range of their human impacts.

Such a research programme around accounting and its social and philosophical implications required significant "investment in new conceptual thinking," as Hopwood (1985: 367) described it. At the beginning of this research agenda, Burrell and Morgan's (1979) book on *Sociological Paradigms*

and Organizational Analysis provided influential insights for accounting scholars and other social scientists in relation to alternative ontological positions towards research, which are also informative to describe and situate the positioning of this study. Burrell and Morgan's (1979) book became particularly influential for their creation of a simple two-by-two matrix, which was developed with the aim to categorize and condense all available social science research paradigms along two axes that organize the model. The horizontal axis is comprised of a subjective vs. objective distinction, while the vertical axis includes research approaches that focus on stability vs. change. As evident from figure 2, in addition to the functionalist perspective, three other paradigms emerge, namely radical humanist, interpretive, and radical structuralist. As such, Burrell and Morgan's (1979) work provided an easily accessible overview in relation to different underlying ontological assumptions of social science research that had a significant impact on interdisciplinary accounting scholarship.

Figure 2: Four Paradigms for the Analysis of Social Theory
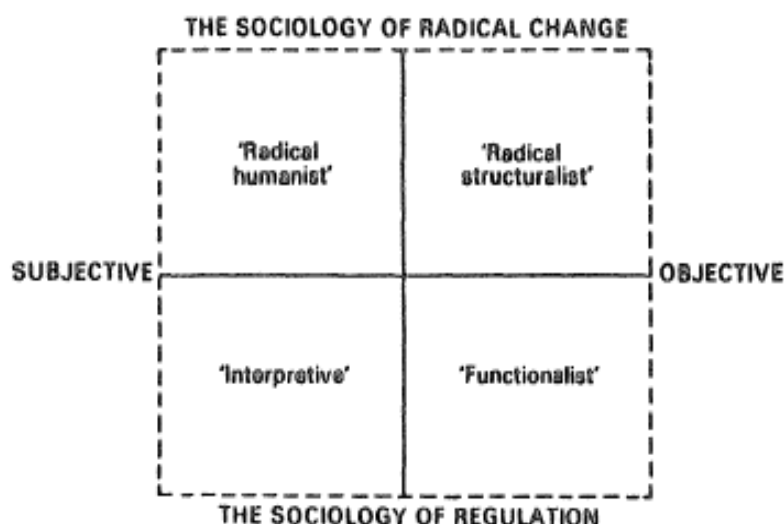
Source: Burrell and Morgan (1979:22)

Yet, Burrell and Morgan's (1979) work also revealed several shortcomings and points of critique that prevent this study from simply 'picking' one of their research paradigms and following their ontological positions. One such critique was that, despite the difference of the paradigms, Burrell and Morgan's (1979) two-by-two matrix essentially limited researchers to two types of explanations for accounting phenomena. Either the phenomenon to be investigated and analysed could be seen as the result of certain presumed macro forces that structure and constitute the accounting phenomenon (as in Marxist perspectives for example), or it can be explained by invoking notions of individual sense-making, as emphasized in interpretivist schools. In both types of approaches, the explanation is external to the accounting phenomenon. In other words, accounting is relegated to the role of a relatively passive entity that is shaped by factors and elements that lie somewhat beyond itself (see Justesen and Mouritsen, 2011).

Thus, while these different sociological and philosophical perspectives constituted a significant enrichment for accounting research, concerns emerged regarding the importance of avoiding what has been described as 'ontological gerrymandering'. Woolgar (1988: 99) offers the following definition of the term: "[Ontological gerrymandering] involves the subtle establishment and manipulation, in the course of the argument, of boundaries between those assumptions and arguments susceptible to deconstruction and those which are not." In other words, Woolgar (1988), sought to draw attention to the way in which the conceptual boundary between phenomena that are considered problematic (and therefore considered worthwhile investigating) and other phenomena that are deemed unproblematic (and can thus be taken

for granted) is manipulated in scientific research and the philosophy of science (see also Woolgar and Pawluch, 1985; Quattrone, 2000).

To work against assumptions that encourage ontological gerrymandering, this dissertation follows methodological advice from Latour (2005). Latour (2005) argues that researchers should treat the 'social' as a flat space whose boundaries are largely defined by empirical observations (see also Callon, 1986; Czarniawska, 2014). This approach, labelled as a 'flat ontology' (Latour, 2005; Czarniawska, 2014), has several advantages. Firstly, it is anti-dualist and rejects artificial dichotomies between subject vs. object, structure vs. agency or technical vs. social that are common in the social sciences, as highlighted in Burrell and Morgan's (1979) framework for example. Such dichotomies can significantly limit what researchers perceive to be relevant by imposing a pre-defined set of concepts on any empirical setting and should therefore be avoided (see also Stark, 2009; Callon, 1986). Secondly, another implication of a 'flat' ontology is that a broader conceptualization of agency is proposed and made possible. As Czarniawska (2014, p.58) highlights: "at the beginning, the only thing that can be distinguished is anything that acts or is acted upon." In other words, any element, including accounting technologies, which makes a difference in the state of affairs of other elements can be considered to have agency and thus warrant attention. Thirdly, the approach discards the assumption of all-powerful macro actors and structures. Instead, by following the advice of keeping the social flat, the question becomes by what route a powerful actor has gained and maintained power in the first place (Czarniawska, 2014).

It is important to clarify that such a 'flat ontology' approach should not be taken too literally. It is evident that for any study the empirical setting and research questions should be carefully selected and it thus follows that certain assumptions about what is interesting and insightful should always guide the process of preparing an investigation. This process of setting up a study is also commonly informed by problems and issues from the established literature, as is the case for this dissertation. Therefore, a flat ontology in an idealized sense is impossible to achieve in pragmatic terms for any study. However, against the background of the shortcomings of ontological gerrymandering, the notion of flat ontology can be considered as a methodological guideline for the research. In other words, it can sensitize the researcher to not over-rely on, or over-specify, the assumptions about the research setting (e.g. structure vs. agency or social vs. technical), treating issues as matters of concern instead of matters of fact. As emphasized by Latour (2005, p.115): "matters of fact may remain silent, they may allow themselves to be simply kicked and thumbed at, but we are not going to run out of data about matters of concern as their *traces* are now found everywhere." Approaching the research through matters of concern is thus in line with an anti-reductionist outlook, which tries to pay careful attention to the multifaceted and complex role of elements and agencies, including accounting technologies. Against this background, the notion of flat ontology informs this study's philosophical underpinnings, its methodological outlook and design, and its approach to data collection and to theory building.

4.2.2. *Epistemological Considerations*

After laying out the underlying ontological considerations of this study, the discussion now turns to the epistemological perspective that informs the research. Epistemology is the domain of philosophy that is concerned with deliberating and advancing different theories of knowledge (Woozley, 1973). The study of epistemology thus seeks to address questions such as: What is knowledge? How can knowledge be acquired? What is truth and method? What are the scope, the possibilities and the limitations of the acquisition of knowledge? While providing an in-depth discussion of each of these questions would be beyond the scope of this section, clarifying the perspective towards the development of knowledge is nevertheless necessary to position this study in relation to the broader body of social scientific research (Bryman, 2008). To do so, this section outlines several concerns that are relevant for the study's epistemological approach.

One concern relates to the link between epistemology as theory of knowledge and ontology as theory of existence. To avoid what Woolgar (1988) called ontological gerrymandering, this dissertation embraced an ontological position of keeping the social flat (Latour, 2005). This position implies that researchers should let empirical observations define the specific boundaries of the setting they are studying. It furthermore implies that notions such as 'power', 'agency' or the 'social' cannot be used to explain research problems but must instead be explained (Czarniawska, 2014; Latour, 2005). In line with these arguments, the notion of 'keeping the social flat' predisposes the research to an epistemological position that can be described as

constructivism. If a flat ontology is considered as the starting point from which a researcher enters an empirical site, a central ensuing question thus becomes how any social space emerged, how it is created, maintained, and constructed in a literal sense (Latour, 2005). Therefore, the notion of a constructivist epistemology is directly linked to the study's ontological positioning.

Against this background, another concern is that the notion of constructivism, which informs this study's epistemological considerations, should not be mistaken for a social constructivist position (see Berger and Luckmann, 1991). In line with the notion of 'keeping the social flat', the constructivist epistemology that is adopted in this study does not assume the existence of any 'social' domain in advance. This point may seem relatively marginal. However, the implications are important. In the constructivist position there is no hidden 'social reality' that lies below the surface and can therefore not be uncovered or deconstructed by the researcher. Instead, a constructivist position implies that any social space "is made to exist by its many ties: attachments are first, actors are second" (Latour, 2005, p.217). By placing the focus on the associations and types of associations that emerge in the empirical site between heterogeneous entities, the task becomes to trace how social worlds are constructed and how specific elements, for example accounting technologies, partake in and shape this construction process.

The crucial take-away emerging from a constructivist epistemology is thus not that 'objects', the 'social' or any other entity are constructed. This is a starting point and is taken as a relatively basic premise (see also Justesen and Mouritsen, 2011). Instead, the argument is that such a constructivist approach

entails the opportunity to bring researchers closer to the entities and associations that are being investigated. As Latour (2004b, p. 231) highlights: "The question was never to get *away* from facts but *closer* to them, not fighting empiricism but, on the contrary, renewing empiricism." In other words, just because a constructivist approach proposes that elements are constructed, it does not mean that they are not real. Instead, it is precisely because something is constructed out of heterogeneous elements that it can exist as a concrete entity. In line with such arguments, a constructivist epistemology is a promising alternative as it encourages researchers to pay attention to the complex and multidimensional associations that shape the emergence and maintenance of social spaces in specific ways (Czarniawska, 2014; Latour, 2005).

Before this methodology chapter moves on to discussing the study's research approach and methods, one final point of clarification should be made in relatioon to the notion of a constructivist epistemology that was outlined above. Similar to the point about adopting a 'flat ontology' approach, the argument of a constructivist epistemology should be regarded with several concerns in mind. While it is taken as a premise that associations between elements that constitute the empirical site should be traced and described in as much detail as possible, it is also implicit that not every association can be followed with the same detail and rigour. This is partly due to limitations in terms of access and time that can be spent in a research site. More importantly however, it relates to the point that the research has a specific focus and interest that is being investigated and pursued. This may lead to situations in which specific choices have to be made about which traces and tracks should be followed and which ones are of lesser interest. It therefore becomes evident

that the researcher always partakes in the construction process and partly stands within the picture that the study seeks to describe and analyse (see Quattrone, 2006). Given that a strict separation between the researcher and the phenomenon under investigation constitutes an epistemological impossibility, it is important to recognize and even embrace this impossibility. Openly acknowledging this epistemological concern is one important step towards an ethical manner of addressing this issue. Another way is to be as respectful as possible of the diversity of matters of concern that speak through the elements from the case settings and in the narrations and analyses that emerge through the investigation in this thesis (Czarniawska, 2004, 2014). Against this background, the notion of a constructivist epistemology offers insightful methodological and conceptual guidelines that inform the research approach, the study's methods, its data collection strategy and analysis. These elements are outlined in the next section.

4.3.   Research Approach and Methods

Having laid out the underlying philosophical assumptions of this research in terms of ontology and epistemology, this section outlines the approach and methodology this study followed. Furthermore, it explains why the chosen approach and methodology is appropriate in relation to the research concerns and the theoretical concerns this study seeks to address.

4.3.1. *Qualitative Case Study Research*

To explore the central research concerns of this study, a case study approach was selected as an appropriate method as it enables the thorough examination of the design and practice of humanitarian performance evaluation systems in their particular social context. Case study research constitutes a long-established and widely used research method cutting across a variety of academic disciplines, including sociology, medicine, law, political sciences, psychology and accounting studies (Flyvbjerg, 2006; Cooper and Morgan, 2008; Quattrone 2006; Muniesa, 2011). However, it is precisely due to this widespread application across disciplines that the exact definition of case study research is often vague and sometimes even misleading (Czarniawska, 2007). It is therefore necessary to clarify what is meant by case study research to position the study's approach in relation to some of the common misunderstandings that arise in relation to this method.

This dissertation narrowly follows Czarniawska's (2014, p.21) specification of a case study as "*the study of the occurrence of a phenomenon* – a chain of events, usually limited in time, usually studied retrospectively." While this definition appears relatively simple and straightforward, it provides an insightful starting point to distinguish assumptions about case study methods from several common misconceptions. One such misconception (see Yin, 1984) is that case study research simply comprises all types of field research. As follows from Czarniawska's (2014) definition above, that case studies focus on the investigation of a specific phenomenon and are therefore inextricably linked to the question: what is this a case of? In line with this question, doing

fieldwork in an organization or a specific research setting is not sufficient to substantiate the claim that the researcher is conducting a case study. Confusing the site with the research phenomenon constitutes a notable and common misconception of case study research (see Czarniawska, 2014; Flyvbjerg, 2006). This distinction is not trivial. As a result, the conceptualization of the research phenomenon can be much wider or much narrower than a single organization. The research phenomenon can be issue based, for example a public policy or a law reform process, or it may comprise an investigation of the design and practice of an accounting technology, as is the case in this study. In other words, the selection of a case study methodology is appropriate for the aims of this dissertation as it allows for the in-depth research of the design and practice of a specific accounting technology in its particular social context.

To enable engagement with the advantages and possible shortcomings of case study research, it is informative to briefly discuss some commonly raised concerns about this methodological approach. The argument that one cannot generalize from case study research is widely considered as a significant criticism of this methodology, questioning the basic validity and justification of its important role in scientific investigations (Czarniawska, 2014). However, the argument does not stand the test of close inspection. A first counterargument is provided by simply looking at the history of scientific discoveries. In a significant breakthrough in physics, Galileo ultimately rejected Aristotle's law of gravity, which had remained unchallenged for nearly two thousand years, through a single, well-selected case study – a conceptual experiment that refuted the claim that heavier objects fall at a quicker speed than lighter ones. This important discovery was not facilitated

by random samples and large quantities of observations, but instead it was made possible by the creative and skilful selection of a critical case (see Flyvbjerg, 2006). This is not to argue that important insights cannot be gained by collecting vast amounts of observations and establishing relationships between them. The key point is that there are many ways of generating important contributions to knowledge, and case study research holds a crucial place in the range of methodological choices that are available to scholars.

Related to this first criticism, another common criticism in relation to case study methods concerns the possible type of knowledge that emerges from the research, which is detailed and context specific instead of broad and context independent (see also Cooper and Morgan, 2008; Quattrone, 2006). The argument goes roughly as follows: if the knowledge produced by case studies is valid only within the specific context of the setting in which the phenomenon is studied, its implications can only be of limited relevance to the broader body of scientific knowledge (Czarniawska, 2014). Given this study's constructivist epistemological approach, it is taken as a basic premise that knowledge is contextually situated. Yet the notion of context-dependent knowledge is not viewed as a shortcoming or weakness in this research. Against this background, Flyvbjerg (2006: 223-224) raises the following important point:

> There does not and probably cannot exist predictive theory in social science. Social science has not succeeded in producing general, context-independent theory and, thus, has in the final instance nothing else to offer than concrete, context-dependent knowledge. And the case study is especially well suited to produce this knowledge… proof is hard to

come by in social science because of the absence of 'hard' theory, whereas learning is certainly possible.

Given that the notion of universal predictive theory is problematic in the (social) sciences, researchers should thus embrace rather than ignore the ensuing consequences. This implies shifting the focus away from placing extensive emphasis on establishing proof of presumed 'facts' about the nature of a phenomenon to in-depth learning about the nuances of a research area (Quattrone, 2006; Czarniawska, 2014; Flyvbjerg, 2001, 2006). This argument is consistent with the shift from focusing on matters of fact to focusing on matters of concern, as proposed by Latour (2005). While a focus on matters of fact closes down debates by reducing phenomena to a few variables that ostensibly determine its nature, approaching research through matters of concern implies an emphasis on the multifaceted being of research phenomena (Latour, 2004a, 2005), such as accounting technologies. In line with these insights, just because knowledge emerging from a specific case study cannot be universally generalized, does not rule out that it can still be a highly valuable contribution to the accumulation of understanding and knowledge of a phenomenon in a given research field. In-depth case study research, with its focus on the investigation of exemplars in a specific context, is thus a particularly appropriate method to produce meaningful contributions to the nuanced understanding of specific research phenomena.

The argument about the appropriateness and validity of specific methods also often takes place as part of a debate between proponents of quantitative and qualitative approaches to research (see Bryman, 2008; Bryman and Bell, 2011). It is therefore informative to briefly position this study's perspective on case

study research in relation to this debate. It is a common misunderstanding to use the term 'case study' as a simple equivalent of qualitative research (Czarniwska, 2014). While most case study research is indeed qualitative, as is this dissertation, there are also several examples of notable case studies using quantitative methods in disciplines ranging from linguistics (Davis & McNeilage, 1990), to psychology (Schacter, Wang, Tulvig, and Freedman 1982). The point is therefore not whether a study is qualitative or quantitative in its approach, but whether the chosen approach enables the study to explore the paradigmatic phenomenon, e.g. the case under investigation, in its richness and depth. The debate over the virtues of quantitative versus qualitative approaches can also be particularly productive and meaningful if it is used to prompt further reflection on the underlying approaches of the method that is used for the investigation, instead of forcibly reducing the diversity of possible research methods. Given this study's exploratory approach on investigating the design and practice of humanitarian performance evaluation systems, a qualitative methodology was chosen as the most appropriate strategy to engage with the research phenomenon (see also Ahrens & Chapman, 2006). Further details about the precise technicalities of the qualitative methods that were employed in this study is given in section 4.3.3. on data collection.

Before moving on to the specification of and justification for the selection of the cases that this study explores in its empirical chapters, one final point of concern in relation to case study methodology should be addressed, namely the danger of verification bias. Verification bias relates to the concern that the research may be designed in a manner that simply leads to the confirmation of pre-existing assumptions the researcher holds about the phenomenon

(Flyvbjerg, 2006). While this concern is relevant for all possible types of research, it has been particularly raised in relation to qualitative case study research (see Diamond, 1996). One of the reasons for this argument is that qualitative case study research requires a range of subjective judgements about the research setting that are criticised as more arbitrary than in studies employing deductive and quantitative methods (Czarniawska, 2014; Flyvbjerg, 2006, 2001). Raising the concern of verification bias is important to sensitize researchers to its possible dangers, its ensuing implications and deficiencies. However, the argument that case study methodology is particularly vulnerable to verification bias in comparison to other methodological approaches does not stand the test of closer inspection. As has been emphasized above, a particular strength of in-depth qualitative case study research is that it enables the scholar to immerse herself in the empirical setting to observe the nuances of how the research phenomenon unfolds in practice. This richness of and closeness to the empirical site is particularly conducive to revisiting researchers' assumptions as they learn about the phenomenon under study. As stressed by Geertz (1995), when conducting in-depth case study research, the empirical setting commonly insists of 'speaking back' which points to the observation that qualitative researchers commonly have to re-visit and change their assumptions and pre-conceptions about the empirical site throughout their study (see also Latour, 2005; Ragin, 1992; Ahrens & Chapman, 2006; Cooper and Morgan, 2008). Furthermore, the issue of subjective judgements and assumptions is one that is a concern for all methodological choices, ranging from the selection of relevant variables in regression analysis to the choices about questions in quantitative surveys. As a result, while all researchers should be sensitive to the issue of possible

verification bias, there is nothing particular about qualitative case study research that would make it a greater concern than in other methods.

4.3.2. *Case Selection and Background*

This section discusses the rationale for the selection of the case studies, which are presented in the empirical chapters. As previously emphasized, the central focus of this study does not lie on the outcomes of particular humanitarian actions, nor does it strictly lie on humanitarian organizations themselves. Instead, the study focuses on tactics and principles that enable humanitarians to evaluate performance and plan for actions under conditions of chaos, complexity and uncertainty. To explore these concerns, three different case studies were selected. The first case study focuses on the analysis of the most widely used performance evaluation system, the Sphere Handbook, which was developed and designed with the sole purpose of engaging with humanitarian crises (Barnett, 2011; Buchanan Smith, 2003). The case was selected because it constitutes a paradigmatic example, a critical case (Flyvbjerg, 2006), of the challenges surrounding the design of humanitarian performance evaluation systems. Sphere is not only commonly recognised as the most critical innovation in the area of humanitarian performance evaluation (Barnett, 2011; ECBC, 2007), but its influence was also confirmed to the researcher during the data collection for this study. Both in the conversations with humanitarian managers and field officers as well as in observations of the crisis management practices in the refugee camps the Sphere Handbook emerged as the central tool for performance evaluation and coordination between humanitarian agencies. As a result, exploring how and

why Sphere's evaluation system is particularly influential in engagement with humanitarian crisis settings became an important focus of the study.

Building on the first case study's analysis on the design of Sphere's evaluation system, the second and third case studies explore how Sphere shapes performance evaluation practices in a specific humanitarian crisis setting, a refugee camp. In line with the selection strategy of the first case, two paradigmatic examples of humanitarian crisis management were chosen, namely the different operational challenges to deliver nutrition and water (Ramalingam, 2013; Walker, 2016). Accordingly, in the second case study the research follows humanitarian crisis managers in their evaluation and organizing practices to build and maintain a functioning system to distribute nutritional supplies for the refugees. The delivery of nutritional supplies constitutes one of the most essential elements of a humanitarian crisis operation, reaching deep into all aspects of life in refugee camps, including survival, health, psychological well-being, cultural norms, logistics, work opportunities, control and politics (see Ramalingam, 2013; Barnett, 2011). As such, exploring the performance evaluation practices surrounding the issue of nutrition allows the research to follow one of the most essential, complex and consequential challenges in humanitarian crisis management.

In line with this approach, the third case study explores the performance evaluation practices in relation to the organization and governance of the refugee camp's water supply chains. Together with the distribution of nutritional supplies, the governance of water is another fundamental operation in the management of humanitarian crises with its own particular

logistical and political challenges. In the desert area where the refugee camp was located, issues surrounding water bear significant potential for conflict between local communities and refugees as well as between humanitarian agencies and the host government. The choice to conduct two separate case studies on the different operations to deliver nutrition and water was not pre-planned but emerged out of the empirical research and engagement with the data. While it is not surprising that issues surrounding nutrition and water are important concerns in any humanitarian response operation (Ramalingam, 2013; Barnett, 2011; Walker, 2016), the centrality and complexity of the evaluation challenges in these two cases only emerged through the engagement with interviewees and the observations that formed part of the field research. Combined, the selection of the different case studies allows the research to thoroughly explore tactics and principles that might enable humanitarians to evaluate performance under conditions of chaos and complexity.

Against the background of the rationale for the case selections, it is important to re-emphasize that the main research focus does not lie on one particular organization or entity. Instead, the study centres its attention on the evaluation and organizing practices (Czarniawska, 2014) in humanitarian crisis management and how these practices relate to the design of Sphere's evaluation system. This approach is in line with the constructivist position that was outlined in the epistemology section. By focusing on the types of associations that emerge between different entities in their organizing and evaluation work, the study can follow the processes that constitute the empirical setting, and thereby place specific emphasis on how the evaluation technology shapes these construction activities (Latour, 2005; Czarniawska,

2014). This approach is suitable given the interconnected nature of elements involved in the camp's organizing and evaluation practices (see Cooper & Morgan, 2008; Latour, 2005; Czarniawska-Joerges, 2007; Barzelay, 2007).

While specific details pertaining to the particularities of each case will be provided in the case studies in the empirical chapter, it is nevertheless informative to provide some broad background on the refugee camp setting and the basic manner through which refugee camps are governed. Located in northern Jordan, Zaatari camp was a key site in which the operation to respond to the Syrian refugee crisis began to unfold. The refugee camp was officially opened in 2012. As roughly four thousand refugees arrived every night, the camp was set up under significant pressures in terms of time schedule and cost (UNHCR, 2015). With this influx of refugees, over time Zaatari not only became the second largest refugee camp in the world, but it is also currently the fourth largest Jordanian city, providing services to around one hundred thousand refugees (UNHCR, 2015). The camp has been in operation for more than four years. Even though this might appear like a long time, in comparison to many other refugee camps around the world it is not. The average operating time of refugee camps from opening to closing lies at around 20 years, while the average stay of refugees in camps is around 12 years (UNHCR, 2013).

The research for this study took place in a period during which the conflict in Syria was in a process of further escalation, with the emergence of terrorist groups such as ISIS beginning to attack border posts and threatening large groups of people within the region. As the threat of terrorism constantly

increased, the willingness of the Jordanian government to accept additional refugees from Syria decreased significantly. The deteriorating security situation was additionally affected by the initiation of bombing campaigns by Syrian, Russian, American and other forces, pressing millions of more people to flee their homes. With its geographical proximity to the Syrian border, managers and refugees in the camp faced significant pressures from these developments in the conflict, including personal security concerns and psychological stress of working and living in this environment. All of these elements provide some further insights into the difficulties, sacrifices and uncertainties of managing a humanitarian response operation. While the number of displaced people was already on a level that was unprecedented since the end of World War II, the management of the international refugee crisis had not yet become a high priority for international policy makers and the media.

At the same time as the policy environment shifted, the Jordanian government remained eager to uphold the temporary nature of the camp. As no end of the Syrian conflict was in sight, refugees became dependent on camps like Zaatari for an indeterminate amount of time. These unfolding events required significant further engagements with multiple areas of management in Zaatari around which conflicts were looming, including issues surrounding human trafficking, educational facilities and the further development of infrastructure. In light of these constantly evolving issues, any solution to the camp's problems was bound to be temporary, including the issues discussed in the case studies. However, these constant changes and unexpected challenges make the analysis and theorization of specific performance

evaluation techniques all the more important, rendering research sites such as Zaatari promising sites for discovery.

In terms of governance, like most other refugee camps around the world, Zaatari is governed through an intricate inter-organizational set-up. The host government exercises its sovereign authority over the territory of the camp, decides how many refugees can enter its borders and is in charge of security. The UN refugee agency, UNHCR, is in charge of administration and the coordination of the refugee camp. While UNHCR is the so-called lead agency, key services are provided by a range of national and international non-governmental agencies (NGOs) and by selected UN organizations, including UNICEF and the World Food Progamme (WFP). The camp is thus governed in the absence of clear-cut command structures. Information collection and decision-making practices are distributed amongst heterogeneous entities and organizations. Due to these interconnections, the most interesting decisions and interactions that shape the future of the camp take place *between* organizations and not necessarily within an organization (see also Stark, 2009; Czarniawska, 2014; Barzelay, 2007). In light of these complexities, the methodological choice to examine the evaluation practices surrounding issues of nutrition and water is all the more appropriate. The unique context of the research site thus provides appropriate conditions in which to focus on the study's key research concerns on how humanitarian performance evaluation systems might facilitate engagement with the unexpected.

4.3.3. _Data Collection_

Having outlined this study's philosophical underpinnings, its qualitative case study approach and the case selections, this section now turns to the discussion of the data collection methods. In particular, the section explains the three data collection methods it employed: semi-structured interview, observations and shadowing, and document analysis. Furthermore, the section provides specific justifications as to why and how the selection of these methods is appropriate for the aims of this dissertation in terms of its anticipated contributions to the understanding of the research phenomenon.

Consistent with the inductive and exploratory approach that informs this study (Bryman, 2008, 2015; Eisenhardt, 1989), the data collection was initiated with a pilot study in July 2014. To gain a more in-depth understanding of current themes and challenges for humanitarian performance evaluation systems and practices, a three-day professional training workshop for humanitarian practitioners was attended in London. Participants at the workshop included humanitarian field officers, logistics managers and performance evaluation experts from several major international humanitarian organizations. The professional training tackled conceptual issues in humanitarian performance evaluation. Furthermore, it involved scenario planning exercises on the application of Sphere's evaluation system within the context of specific humanitarian crises scenarios, including war zones, natural disasters and protracted humanitarian crises settings. As part of the pilot study, several types of data were collected from different sources, including observations, informal conversations with humanitarian field

workers, and one formal interview with the instructor of the workshop, who is an experienced humanitarian field worker and consultant with two decades of experience in crisis and disaster management. The ordering and coding of this data enabled a first engagement with potential issues and challenges surrounding the design and practice of evaluation systems for humanitarian crisis contexts prior to the commencement of the actual data collection stage (Bryman, 2008, 2015). These insights helped to develop a range of guiding questions to be used later on in semi-structured interviews. Another important take away from the pilot study was that it re-emphasized the centrality of the Sphere Handbook for performance evaluation, planning and coordination practices in large scale humanitarian crises.

In the subsequent data collection, both in the field visits to the refugee camp as well as in the engagement with Sphere's evaluation system designers, the study relied on semi-structured interviews as one of the primary data collection methods (Czarniawska, 2014; Bryman, 2008, 2015; Eisenhardt, 1989). In total, twenty-eight different participants were interviewed from July 2014 to October 2015. Interviewees comprised twenty refugee camp managers, covering a range of different roles in humanitarian crisis management, including emergency coordinators, senior field officers, evaluation officers, engineers, humanitarian consultants and community organizers. Furthermore, interviewees included eight Sphere board members and system designers. While some of the interviewees were narrow experts for the particular case studies on nutrition and water, others were general managers and thus provided insights for the evaluation and governance challenges for two of the case studies in the camp. Furthermore, follow-up interviews with six of the participants, three Sphere system designers and three humanitarian

crisis managers, were successively conducted over a period of twelve months over the phone until November 2015. Almost all interviews were digitally recorded and supplemented by extensive note taking. Upon transcription two interviews were found to be inaudible due to faulty technology. However, the core arguments were recovered through a combination of field notes and memory. Another important element of the data collection was informal conversations before and after the interviews, which informed and provided further insights into the issues raised during the interview (Bryman and Bell, 2011). In several cases, these informal conversations included meetings with refugees, who provided information about certain challenges of living in the camp.

Semi-structured interviews are particularly suitable for an inductive and exploratory research approach, as they permit flexibility while also offering a guide that structures the interview questions across participants to explore different aspects and nuances of the research phenomenon. In her work on research methods, Czarniawska (2014: 30) offers an insightful perspective on the particularities and potential of employing interviews as a data collection method:

> In order to exploit the technique in full, it is essential to understand that interviews do not stand for anything else; they merely represent an interaction that is recorded or inscribed. That is all they stand for, and it is more than enough…the interview can be treated first, as an occasion for eliciting narratives (stories); second, as a special type of observation; and third, as an opportunity to collect samples of the prevalent logic of representation – in other words, of the dominant discourse… although each of the accounts will be unique in the way every interaction is unique, it would be both presumptuous and

unrealistic to assume that interviewees invent a whole new story for the sake of the researcher who happened to interview them.

Following Czarniawska's (2014) insights, the perspective on interviews that is followed in this study recognizes that narratives and conversations still constitute the predominant mode of knowledge production in modern societies (see also Kvale, 1996; Latour, 2005). Accordingly, interviews are ideally suited to explore and elicit different forms of knowledge and perspectives about a particular research phenomenon investigated. The observation that interviews are unique interactions is hereby not considered as problematic, but instead as an opportunity to explore the accounts and particularities of the interviewee's experience about the research setting and phenomenon. While the interview may confirm some of the existing assumptions about the phenomenon, focusing in particular on strange, surprising or unanticipated elements of an account emerging through the interview can be a powerful technique to explore important nuances in rich detail (see also Silverman, 1993; Ahrens & Chapman, 2006; Dambrin and Robson, 2011).

In terms of the technicalities of conducting semi-structured interviews, a mix of different types of questions was used to facilitate the interaction. This mix included questions in relation to introducing specific conceptual concerns of the research. Furthermore, it also included questions to follow-up on particular issues, to clarify specific statements, to confirm what had been said and to elaborate on certain key points (Bryman and Bell, 2011). Even though the precise mix of questions was adjusted to the person that was interviewed, drawing from them permitted a loosely structured approach while at the same

time providing space for the interviewee to express and expand on specific unanticipated and emerging concerns. Two examples of an interview schedule with a Sphere system designer as well as a refugee camp manager are included in appendix 1 and 2. Whenever possible, questions were formulated in an open ended manner to allow the interviewee to lead the interaction into directions s/he considered relevant and informative. Apart from the key leading questions, an additional range of questions was developed that could be spontaneously used to prompt further discussions in case the interview needed additional input to keep up the flow of the interaction.

In terms of concerns relating to formal protocol and research ethics, each interview was initiated with a discussion of several important elements regarding the study. Firstly, each interviewee was asked for permission to record the interview on a digital recorder. Secondly, interviewees were briefed that the data would be handled in an anonymised manner. Thirdly, it was clarified that participants could stop the interview at any point and could decide to withdraw any possible interview statements from the transcript that may have been made during the interview and that they no longer felt comfortable with. Finally, before the interviews began, every interviewee was given the opportunity to ask any questions that they wanted to clarify. As a result, in some instances the interaction began with reversed roles, in which the interviewees interrogated the researcher on specific issues before agreeing to participate. This building of rapport (Bédard and Gendron, 2004; Bryman and Bell, 2011) was particularly important for the senior camp managers who wanted to ensure the integrity of the researcher and the overall approach to data handling. This was also important to discussing and clarifying any security concerns and issues in relation to codes of conduct in the camps.

Another important source of data related to observations and shadowing of organizing and evaluation practices in the refugee camp. The use of multiple sources of data collection is a methodological point in its own right. It not only draws attention to the inherent multiplicity of social research phenomena, but it also highlights that social sciences researchers should never be overly certain in relation to the completeness of the sources and methods used in the investigation (Czarniawska, 2014). Shadowing is a technique that has been widely used in the social sciences (Czarniawska-Joerges, 2007; Latour, 1999), including in management studies (see Mintzberg, 1973) and accounting research (see Chenhall et al., 2013). At a basic level, shadowing implies that the researcher observes participants in the way they perform their work. This observation usually relates to an interest in a specific research phenomenon. In the case of this study, the focus of the observations was placed on a specific performance evaluation technology, namely the Sphere Handbook, and the practices that shape and are shaped by this technology. The use of observations and shadowing are particularly insightful to trace associations and connections that are part of the mundane everyday practices, shaping interactions and constituting the empirical setting within which the research phenomenon is situated (Latour, 2005).

In this study, one set of observations and shadowing (Czarniawska, 2014; Bryman, 2008, 2015) were carried out in relation to engineering teams in the refugee camp. The teams were observed during meetings, in their office work, and in their trips through the refugee camp in which they attempted to deal with specific issues related to supply chains, performance evaluation and governance in relation to the issues of water as well as nutrition. Furthermore,

one of the senior camp managers was shadowed in a full day of working practices, which included interactions with refugee representatives, meetings with security chiefs, conversations with local government officials and journalists. Finally, observations were carried out in coordination meetings in which several of the organizations responsible for the technical delivery of services discussed planning challenges and strategic concerns for the upcoming weeks and months. To keep track of the multitude of information sources and impressions, a detailed notebook was kept, which was subsequently translated into a written narrative (Czarniawska-Joerges, 2007; Rottenburg, 2009). This written narrative not only helped to make sense of the different information sources, but it also helped to organize the data and establish connections, themes and overlaps between the different issues that were raised in relation to the research phenomenon.

Finally, data was collected from a third source, namely from documents (Czarniawska, 2014; Eisenhardt, 1989; Bryman, 2015). These documents were publicly available, and included policy documents, opinion pieces and newspaper articles about the refugee camp and Sphere's performance evaluation system. The research also drew from proprietary documents that were acquired through interviewees and other contacts in the refugee camp and within the humanitarian community, including a variety of evaluation reports and governance frameworks. These documents were not only important to gaining a deeper understanding of the case studies on nutrition and water, but were also crucial to complement the analysis of the design of the Sphere evaluation system and the underlying assumptions and rationales that inform its approach. Furthermore, these documents provided further information on issues that could subsequently be discussed in follow up

interviews. The triangulation between the different data sources thus offered an insightful strategy to enrich the data collection and embrace the multidimensional nature of the humanitarian performance evaluation and organizing practices that form the key focus of this study (Bryman & Bell, 2011; Eisenhardt, 1989).

4.3.4. _Data Analysis_

Building on the previous sections, this section outlines how the different types of data collected for this study were organized and analysed. In a first step, the different data sources, semi-structured interviews, observations and documentary sources were transcribed. Subsequently, the documents were placed into a qualitative data analysis software (Nvivo) as well as into a folder on an external hard drive that was secured with a password and stored at a site that was only accessible to the researcher. The collection of the data into a single space (pooling) enabled a significant first engagement with the empirical material following the interviews. Importantly, it allowed for an initial comparison between the impressions that were formed during the data collection with those formed during the transcription process, enabling the familiarization with the data (Bryman & Bell, 2011; Bryman, 2008). During the transcription, unexpected and unusual elements were flagged for further examination. Furthermore, issues that re-appeared repeatedly were marked for further attention in other transcripts with a view to further exploring them in subsequent interviews. As the pool of transcripts and documents grew significantly over the one-year data collection period, this approach enabled

constant comparative analysis between emergent matters of concern (Glaser & Strauss, 1967; Latour, 2005).

Extending this initial engagement with the collected data, following the guidelines advocated by Czarniwska (2014) and Glaser and Strauss (1967), the analysis proceeded to further order the materials chronologically to identify themes and patterns (see also Ahrens & Chapman, 2006; Eisenhardt, 1989). The data analysis focused in particular on frequently repeated statements in relation to the design and practice of Sphere to explore how and why these statements were made and to compare them with responses from other participants in relation to comparable issues. The analysis then proceeded by relating these themes to issues in the literature, permitting a shift back and forth between empirical data and theory in a triangulating manner. This facilitated a comparison between expectations from the previously identified theoretical insights and the emergent issues in the data (Ahrens & Chapman, 2004, 2006; Cooper & Morgan, 2008; Eisenhardt, 1989).

As the engagement with different themes advanced, the data analysis turned to one of the most central tasks for qualitative data analysis, namely the aim to establish connections between the themes and matters of concern that emerged from the data. In line with these insights, the data analysis continued by re-arranging the data around frequently repeated matters of concern (e.g. unrest in the camp and notions of refugee dignity), connecting them with other important topics (e.g. clashes between the techno-financial framing of humanitarian crises and matters of refugee participation), to comprehend key issues surrounding the evaluation practices in the camp. The use of the

qualitative data analysis software Nvivo was useful at these initial stages of ordering, categorizing and sorting the data as it provided a relatively easy way of exploring and visualizing connections. However, as emphasized by Czarniwska (2014, p.98), all relevant "connections need to be explicated, not merely signalled by arrows in a simplified model." As a result of this challenge, in the further explication of the connections between different matters of concern, the importance of Nvivo reduced as a data analysis tool.

To complement the analysis, the different connections between themes and matters of concern were compiled into a narrative surrounding key problems confronting the interlocutors (Czarniawska, 2004). Such issues included problems surrounding nutrition in the refugee camp, a key challenge confronting the refugee camp managers. A narrative approach does not only offer the advantage of proactively engaging with the data and making the connections between emergent themes concrete and rich in detail, but it also offers the opportunity to compile a theorized account of the issues, thus providing a medium through which the data can be related to literature and theoretical concerns at an early point in the analysis (Latour, 1996; 2005; Rottenburg, 2009; Mol, 2002). The narratives were updated several times as the researcher's familiarity with the several data sources increased and more and more connections between themes, matters of concern and theoretical issues were established (Czarniawska, 2004, 2014). This process of updating the narratives was complemented by presenting the results of the analysis at numerous workshops and conferences to different audiences, including humanitarian practitioners, Sphere evaluation system designers, accounting scholars and the organization theory community. Engaging with the different audiences not only facilitated a further interrogation and validation of

empirical points, it also encouraged reflection about and revision of the broader theoretical implications and contributions emerging from them.

This approach to dealing with empirical material is consistent with the aim to treat research phenomena as matters of concern instead of matters of fact (Latour, 2005), as was outlined in section 4.1.2. on the ontological approach of this research. Conceptualizing research phenomena as matters of concern not only recognizes the multifaceted and complex role of elements and agencies in an empirical setting, it also seeks to embrace and explicate rather than reduce the richness of different types of associations and traces that exist between different elements in the field. Therefore, pursuing a narrative strategy to engaging with data and its analysis is appropriate to take seriously the inherent multiplicity of beings and relationships in the empirical setting (Latour, 2005; Czarniawska, 2004).

One final point of clarification should be made in relation to this dissertation's approach to engaging with and analysing data that is theoretical in nature but has important practical and ethical implications for the manner in which accounts given by interlocutors are re-presented in the analysis and text. As has been noted by Latour (1988) in the *Politics of Explanation*, the construction of a new text is a highly political enterprise involving questions such as: who has the right to speak for whom? What criteria are used to establish this right of representation? Who will be the judge of the account given and by what authority? Furthermore, even when similar criteria, for example coding conventions, are applied, there is always the possibility of several different re-presentations of the same research phenomenon. The production of a text is

therefore always a creation of a new entity, which is connected to its sources by the conventions of data analysis and coding. The implications of this argument are at least twofold. Firstly, acknowledging and embracing the impossibility of full representation recognizes that there is always a gap in the account that cannot be closed. This recognition introduces a measure of humility towards the informants for which it claims to speak, as it acknowledges the multifaceted nature of their being that cannot be reduced to a single text. As Woolgar (2015, ad vocem) noted, "it could be otherwise." Secondly, it recognizes the importance of the medium through which the account is constructed and re-presented. Accordingly, the account does not derive its authority from claiming that it offers a full representation and literal description of all the 'facts' in the empirical setting. Every account as a form of re-presentation always implies a reduction. As highlighted by Czarniwska (2014: 123): "a skilful description depends heavily upon metonymy and synecdoche – on deleting some information – in the hope that readers will fill in the blanks, which should also increase their engagement in the reproduction of the text." Accordingly, such an approach combines the epistemological concern of representation with ethical concerns in relation to the events that are re-presented in the accounts and the way they affect the interlocutors, the reader and broader society (see also Quattrone, 2006; Atkinson, 1990). In light of these insights, the approach towards engaging with, analysing and presenting the data that inform this research project are guided by the aim to balance these different concerns as much as possible.

4.4. <u>Concluding Remarks</u>

This chapter on methodology has described and discussed how the research was carried out. It began by providing insights about the philosophical underpinnings that guided the study. To avoid what has been called 'ontological gerrymandering' (Woolgar and Pawluch, 1985), the study followed advice from Latour (2005) in adopting an ontological approach of keeping the social 'flat.' In line with this flat ontology, the study followed a constructivist epistemology. This epistemological position encourages the investigation to focus on the associations that emerge in the empirical site between entities to trace how social worlds are constructed and how specific elements partake in this construction process (Latour, 2005; Czarniawska, 2014). Based on these philosophical considerations, qualitative case study research was chosen as the appropriate approach to advance the concerns of this thesis. Given the highly problematic nature of developing predictive theory in (social) science research, it was argued that a qualitative case study approach offers an ideal method to advance in-depth understanding about the phenomenon under investigation (see Flyvbjerg, 2001, 2006; Quattrone, 2006).

The chapter then outlined the reasons for selection of the case studies. Three paradigmatic cases (Czarniawska, 2014) were selected to investigate tactics and principles that enable humanitarians to evaluate performance and plan for actions under conditions of chaos, complexity and uncertainty. The first case study investigates the design of the most widely used performance evaluation system that was developed for the sole purpose of engaging with humanitarian crises (Barnett, 2011; Buchanan Smith, 2003; ECBC, 2007). Building on the first case study's analysis, the second and third case studies explore how Sphere shapes performance evaluation practices in a specific humanitarian crisis setting, a refugee camp. In line with the case study

selection strategy, two paradigmatic examples of humanitarian crisis management were chosen, namely the different operational challenges to deliver nutrition and water (Walker, 2016).

Subsequently, the chapter outlined the methods that were employed to collect the data that inform this study. These methods included semi-structured interviews, observations, and the collection of relevant documentary sources from the research settings. Finally, the chapter outlined how the collected data was organized and analysed, which involved thematic coding and establishing connections by constructing theorized narratives between the different emerging matters of concern from the empirical data (Czarniwska, 2014; Latour, 2005; Glaser & Strauss, 1967). Having outlined the central methodological considerations that informed the research, the study is now in a position to present the empirical findings from the three case studies that were conducted.

5. Empirical Findings

5.1. <u>Introduction and Overview</u>

This chapter presents the empirical findings from the three case studies that were conducted as part of the investigation. In doing so, the research is guided by several key concerns. In the previous review of the literature, it was highlighted that instead of focusing on the question of how accounting systems might be implicated in the attempt to 'normalize' disaster situations, the more promising question is how accounting systems might be designed and practiced to facilitate engagement with anomalies and the unexpected by embracing the ambiguity that defines humanitarian operations. Furthermore, it was highlighted that drawing from the concept of heterarchy, with its focus on 'governing through difference,' provides promising insights into a distinct approach towards engaging with heterogeneous forms of evaluation in contexts of uncertainty and complexity. Instead of pre-specifying ideal solutions through top-down designs, heterarchical systems encourage attention to processes that seek to institutionalize forms of criticism and dissonance, foster checks and balances, and recursively create spaces of ambiguity and reflection (Stark, 2009; Lamont, 2012). Combined, these concerns provide a useful starting point for this chapter's analysis.

To explore these issues, the chapter is structured as follows. The first case study is concerned with the analysis of the most widely used humanitarian performance evaluation system, the Sphere standards (hereafter Sphere) (Barnett, 2011; Buchanan Smith, 2003). It begins with a brief overview of the historical conditions and constellations that led to the development of Sphere. The chapter then continues to its main analysis by focusing on four key conceptual problems that Sphere's developers confronted in the design of the evaluation system, which emerged through the empirical research. It is hereby argued that the conceptual answers to these problems set Sphere apart as a distinctive system for the evaluation of performance in situations of humanitarian crisis.

Building on these insights into the conceptual issues surrounding the design of an evaluation system for engagement with humanitarian crises, the second case study focuses its attention on the enactment of humanitarian performance, and the way Sphere shapes this enactment, in a specific disaster setting, namely a large-scale refugee camp. The case study specifically focuses on the challenges emerging in the attempt to manage nutritional aid programmes to address the refugee crisis. The case study hereby indicates how the crisis management practices were shaped by several key elements emerging through engagement with Sphere to make sense of, order and reconfigure understandings of the dynamically unfolding situation in the camp. These elements structure the narrative of the case study and shed light on the question how and why Sphere became an influential tool for engagement with the instability and unpredictability of the crisis setting.

The third case study further pursues these concerns by focusing on the complexities of evaluating actions and governing the delivery of water for the thousands of refugees affected by the humanitarian crisis. Similar to the case study on the delivery of nutrition, the governance of the water networks constitutes one of the most consequential and challenging issues in the humanitarian response, affecting all areas of life in the camp. As the managers are confronted with a specific water governance issue– unprofessional and corrupt organizations and other individuals making the existing system for distribution increasingly unfeasible – Sphere's evaluation system becomes highly influential in detecting challenges and in exploring new solutions for the camp. By focusing on tactics and principles that offer the humanitarian managers insights into confronting these challenges, the case study explores in detail how and through which means Sphere becomes a powerful element in this context.

5.2. Case Study 1: Designing for Crises and Complexity – Sphere's Humanitarian Evaluation System

5.2.1. *Sphere's Background: Developing an Evaluation System for Crises*

The development of performance evaluation systems and metrics in the humanitarian sector gained momentum as a key theme amongst practitioners following the disaster of the 1994 Rwanda genocide. Accentuated by images of death, destruction and suffering, international news media reported on the

chaotic management of the humanitarian response to Rwanda on a previously unmatched scale. Stories emerged on how humanitarians had sheltered, fed and in several instances employed staff associated with Hutu *genocidaires*; how refugee camps, built for the protection of civilians, became sites for perpetrators to re-group and organize military attacks; and how the wastage of resources, insufficient cooperation between agencies, inadequate planning and lack of clean water led to further disease and deaths amongst the affected populations (Rieff, 2002; Barnett, 2011). Among the severe criticisms of the aid operation in Rwanda was the absence of any system against which the performance of humanitarian operations could be evaluated (see JEEAR, 1996). The demand for the development of such a system was further magnified as humanitarian space opened up following the end of the Cold War, which meant that humanitarian aid operations now became truly global, entering into ever more contested and complex environments and thereby raising questions about norms and values of humanitarian practice (Buchanan-Smith, 2003; Barnett, 2011). While the issue of performance evaluation systems and standards had been on and off the agenda of humanitarians for many years, the lessons of the Rwanda response operation was sufficiently alarming enough to initiate and give momentum to an ambitious agenda for humanitarian reform (Rieff, 2002).

As a result, driven by two of the largest European and North American humanitarian umbrella organizations, InterAction and the Steering Committee for Humanitarian Response (SCHR), the Sphere Project was initiated in 1997 with the aim to develop and design a system for the planning and evaluation of performance for humanitarian crisis settings. The process was subsequently described as the largest ever consultation on humanitarian

performance, with hundreds of organizations and individuals participating in a collaborative endeavour (see also Buchanan-Smith, 2003). In line with its considerable ambition in terms of participation, Sphere was also designed to be applicable to the whole humanitarian sector, all areas of humanitarian response and all types of crisis (Barnett, 2011; ECBC, 2007). This ambition is highlighted in one of the introductory paragraphs to the Sphere Handbook (2011:9), which was designed for "a range of situations including natural disasters, conflict, slow-and rapid-onset events, rural and urban environments, and complex political emergencies in all countries." The challenges of constructing a system for humanitarian performance evaluation and creating measurable indicators were (and continue to be) numerous nonetheless. In what follows below, the analysis focuses on several key conceptual problems Sphere's developers confronted in their attempt to design an evaluation system with the sole focus on the management of humanitarian crises.

5.2.2. *Problem 1: The Quantification Conundrum – Technical Measures and Dogma in Humanitarian Performance Evaluation*

From the very beginning of the process of designing Sphere, a variety of influential humanitarian agencies expressed an unequivocal concern with the issue of technical and quantitative performance measures. Designer 1, a key individual driving the establishment and development of Sphere, describes the issue in the following terms:

Several large agencies, including the ICRC [the International Confederation of the Red Cross] and also Medicines Sans Frontiers were very concerned to ensure that Sphere was not moving into doctrinal areas. We had to avoid a dogmatic way of thinking about and working with measures (designer 1).

Defining humanitarian performance along its technical dimension constitutes one of the most influential and at the same time controversial endeavours for Sphere's evaluation system (Barnett, 2011; Buchanan Smith, 2003). A significant share of Sphere's roughly four hundred pages consists in the definition and specification of a range of often quantitative indicators that outline minimum response requirements for humanitarian action in key areas of humanitarian response, including water and sanitation, shelter and site planning, nutrition, and health services. These quantitative measures are not only the most widely known indicators to evaluate humanitarian performance, but they have also been embraced by a large number of humanitarian agencies, ranging from United Nations organizations, including UNHCR or UNICEF, to some of the most influential international NGOs (ECBC, 2007; Barnett, 2011). Designer 4, a former member of the Sphere project board, describes some of the challenges surrounding the development of measurable performance indicators for humanitarian crises thus:

What are the outcomes of humanitarian action? I mean outcomes of humanitarian programmes are incredibly difficult to measure because what are you going to report on? Are you going to report on the number of lives saved? Are you going to report on the number of people who have managed to get all their assets back? I mean what are you going to report on? The trouble is if you are looking on outcomes and on impact, you have a whole problem of attribution and how can you attribute changes to one agency? You can't (designer 4).

Formulating a range of quantitative indicators for humanitarian practice has been controversial for a range of reasons, most of which are consistent with concerns raised in the accounting literature in regards to the quantification of social life. These issues include (1) the significant challenges of attributing changes in performance to a particular individual or organization due to opacity and ambivalence in the construction of performance information (Jordan and Messner, 2012; Dambrin and Robson, 2011); (2) shifting attention away from the heterogeneous political, cultural and security environments that shape response capacities of humanitarian agencies (Everett and Friessen, 2010); and (3) the somewhat arbitrary selection of indicators that might be used for the ex-post legitimation of actions in contexts of uncertainty (Burchell et al., 1980). From the inception of the Sphere Project, its designers were attentive to some of these elements, as stressed in one of the initial policy documents on the establishment of Sphere by the Steering Committee on Humanitarian Response (SCHR):

> To elaborate technical standards, which agencies should seek to implement, without reference in any way to the rights or aspirations of the assisted beneficiaries…risks becoming a self-serving exercise concerned more with agencies' accountability to donors, than the rights of people affected by disaster (SCHR, 1997, p.5).

Some of these concerns can be related to what Strathern (2000:309) termed the 'tyranny of transparency,' which emphasizes the danger that performance metrics might serve as technologies that work contrary to, and draw attention away from, key organizational capacities, including creativity, reflexivity and trust. As Strathern (2000:313) put it: "more information, less understanding,

and in particular *more information, less trust*" (italics from original). Yet, while these arguments constitute important and relevant concerns, they do not alleviate some of the key challenges for both humanitarian practitioners and evaluation system designers. Designer 3 outlines some important challenges:

> Let me first speak as a humanitarian worker and not as a Sphere official. As a humanitarian worker, any standard would be largely impractical without indicators and guidance notes. But the whole development of the notion of control and evaluation is very questionable and problematic in our sector. Allow me to give you a simple example. If a medical doctor continues to kill people, either because of negligence or because of deliberate experimentation, he will be done. But think about scenarios in the humanitarian sector. Take a flood. Contamination of water after a flood often kills more people than the flood itself. Have you come across aid agencies being shut down because of the low quality work they do? In many cases when governments or rebel organizations are not able to provide services of sufficient quality during conflicts, aid agencies take up that role, and they make decisions about people's lives. This is one of our biggest challenges. While most of our principles are morally binding, they are not legally binding (designer 3).

Several key concerns emerge from the above quotes and excerpts. A first concern relates to a simple issue of practicality, which emerges time and again in the descriptions of Sphere's designers and humanitarian practitioners. Without a range of technical indicators, humanitarians would often become quickly overwhelmed by the complexities of large response operations. Sphere's (2011:6) quantitative metrics should not only serve as important "signals" for the ongoing evaluation of decisions, but also inform the planning and budgeting of response operations. Secondly, as stressed by designer 3, the indicators also provide a benchmark for humanitarian agencies providing services to populations affected by a disaster in the absence of functioning

governmental infrastructures. In this sense, Sphere's designers are seeking to provide an evaluative and governance infrastructure for the more than sixty million refugees around the world today, while being conscious that there is no conceptual or practical way to enforce compliance. Thirdly, a central challenge for Sphere's designers was to create a conceptual link between abstract humanist principles that would guide humanitarians in their practices and quantifiable measures that enable engagement with the technical aspects of humanitarian work. Designer 3 and designer 4 further elaborate on this point:

> There has obviously been this challenge that people want to reduce Sphere to a mere set of numbers, of indicators, which is obviously understandable for practitioners. So there have been discussions within Sphere in the past, so why don't we just pull out two pages and why don't we just put all the indicators in one list, so people have it very easily and visually accessible what they need to meet. Sphere would never do that. The idea is not to provide a checklist because then you lose the flavour. So the problem is if you do a half an hour crash course, the tendency for engineers is to ask "what are we actually talking about? What should be the flow of the water?" They can get frustrated when the response is "it's not just that guys." Without the philosophy, the system becomes irrelevant (designer 3).

> When Sphere talks about an ethical framework it talks about the right to protection, to security, the right to receive humanitarian assistance, and the right to a life with dignity. This approach is usually impractical and even unrealistic to implement. We are working in countries like Syria, Sudan, where governments have a very different perspective on these issues. We are however able to do our own programming with it, which is to say taking the humanitarian charter as a start and then looking at the rest of Sphere as kind of an articulation of what that might mean. How do we break that down into specifics? As soon as you are starting to talk about results you want to start talking about numbers and target values. When it comes to humanitarian assistance

it comes down to how many litres of water per day, how many square metres of living space, and everybody forgets that these are indicators (designer 4).

The notions of dogma and doctrinal areas that were raised in the beginning by designer 1 is important in this context. In its extreme form, such an approach encourages the acceptance of performance indicators as taken-for-granted categories to be achieved irrespective of context. In other words, the authority of the indicators would derive from a strong belief in the expert knowledge of the designers of the performance evaluation system in relation to the definition of what needs to be achieved in a 'generic' humanitarian response operation. Such a top-down approach to performance metrics would neatly mirror the checklist approach described as highly problematic by designer 3, in which the users' only task is to fill in information and check the indicator boxes. Yet, while it appears relatively straightforward to point out the shortcomings of such a system, conceptualizing and designing a system that works against such notions constitutes a challenging task.

5.2.3. *Problem 2: Accounting for Gaps - The Role of Distrust and Scepticism in the Performativity of Measures*

Designing for complexity and uncertainty entails recognizing in a first step the inherent limitations of a performance evaluation system in being able to anticipate the diverse range of response requirements across distinct humanitarian crises. In this context, a key question for the designers was whether it is possible to embed within the evaluation system elements and

principles that might help overcome these limitations. The contours of this objective are sketched out in an important introductory section of the Sphere Handbook (2011:11):

> The Handbook is essentially designed as a tool to recognise different contexts and to adapt response programmes accordingly: it guides practitioners in their reflections around reaching a universally applicable standard in a concrete situation or context.

Several key terms stand out from this quote: adaptation, design, contextualization, universal applicability, reflection, and recognition. These terms and the interconnections between them raise interesting questions. How can an evaluation system be universally applicable while requiring contextualization that takes into account the significant variations in deliverables of every humanitarian crisis? In what manner might an evaluation system pro-actively guide reflections? How might an evaluation system be designed so that it facilitates recognition of complex and emergent challenges that cannot be categorized in a predetermined manner? How can an evaluation system encourage adaptation of response programmes that were informed by its own prescriptions in the first place? To approach these issues, in a first move, Sphere shifts attention to procedural elements in humanitarian response management, and pro-actively de-emphasizes the requirement to meet and strictly adhere to all the performance indicators:

> The Handbook does not offer practical guidance on how to provide certain services (the key actions suggest activities to reach a standard without specifying how to do that). Rather, it explains what needs to be in place in order to ensure a life with dignity for the affected population.

It is, therefore, up to each implementing agency to choose a system to ensure conformance with the Sphere minimum standards… Conforming with Sphere does not mean meeting all the standards and indicators (Sphere Handbook, 2011, p. 8).

When you describe and mitigate the gaps between Sphere indicators and the ones reached in practice you are conforming to Sphere and you are being accountable both to your donors and to the people you seek to assist (designer 5).

As emerges from the quote and excerpt, an important aim of Sphere is to encourage users to pay attention to the information gaps that emerge through the interaction with its indicators. In this sense, certain similarities, but also important differences can be observed with recent literature on the role of 'imperfection' and the incompleteness of performance measures. These studies show how managers 'make do' with these incomplete performance measures despite their imperfections by relying on information external to the performance evaluation system (Bürkeland et al. 2010), or by attempting to repair the performance measures or by distancing themselves from them (Jordan and Messner, 2012). In another study, Dambrin and Robson (2011) show how imperfect measures become performative because of their opacity and because users have trust in the system designers. Due to these elements users commonly ignore the gaps that are created by translating contextual information into performance indicators. Sphere's approach to humanitarian performance evaluation exhibits important differences with these studies. As stressed by designer 5, detecting and acknowledging the role of errors and gaps forms a key preoccupation of humanitarians in general and of Sphere's system designers in particular: "We are in the business of error minimization." By suggesting that users focus their attention on such gaps, Sphere attempts

to broaden users' outlook to the range of interconnected factors affecting the situation under analysis. Gaps, inconsistencies and errors are thus not treated as problematic, but are instead embraced and can serve as a source of action and reflection. In the case of Sphere, stimulating a certain degree of distrust and scepticism in the indicators thus appears to serve as a resource to make them performative in the management of crises.

Downplaying the authority of its own measures might at first sight appear counterintuitive for performance evaluation system designers. A downside of such an approach is that it might lead to occasional confusion amongst users. As designer 3 emphasizes in the quote below, it is not uncommon for humanitarian practitioners long for an approach that gives clear and unambiguous prescription:

> So I think this is where the numbers alone miss the whole new ways of looking at notions of well-being, the happiness coefficient, the whole range of things that count. I know some think Sphere is a Magna Charta, but I think it is a catalyst. In today's world the context changes very quickly, and the context should define and will influence what happens. So you take principles and ethics, and you apply them to the new context and something comes out. So it's a question about what sense you make out of it. That's where it becomes a powerful catalyst in the hands of an informed actor on the ground. If you take context out of Sphere, the whole game is gone (designer 3).

Due to the highly heterogeneous challenges across different humanitarian crises, an evaluation system that over-specifies response requirements in a rigid manner would be quickly overwhelmed by emergent complexities and contextual variations. While the issues surrounding technical measures

constituted a crucial challenge for Sphere's designers, another central dilemma was presented by the issue of how to conceptualize the notion of humanitarian value.

5.2.4. *Problem 3: Dealing with a Multiplicity of Humanitarian Modes of Evaluation*

In the above passages on the relationship between technical evaluation criteria and ethical humanitarian principles, an important debate is hinted at, which centres around the following simple question: How can an evaluation system be designed and practiced that does not reduce humanitarian actions to a single techno-financial value dimension? As has been widely noted in the literature, humanitarian value is inherently multiple (Barnett, 2011). Sphere's development of technical principles was broadly considered as an innovation following the aftermath of several failed humanitarian response operations during the 1990s, in particular the Rwanda crisis. However, it was also recognized that a large proportion of deaths during these crises was not because technical evaluation criteria had not been fulfilled, but because of a lack of political will by state actors to protect civilians during conflict. Simply adhering to a range of technical performance criteria would therefore not have been sufficient to improve humanitarian performance. Experience from several humanitarian response operations had also illustrated situations in which agencies were meeting technical indicators but at the same time put affected populations at risk by failing to take into account key requirements for their protection. For example, in the case of Rwanda it was widely reported that several humanitarian agencies had fed and harboured perpetrators of the genocide. As these perpetrators recovered and regrouped in humanitarian

safe spaces, the response operation was deemed to have contributed to prolonging the conflict (see JEEAR, 1996; Barnett, 2011).

All of these lessons had wide ranging influences on the process of designing Sphere's evaluation system. Sphere's response to this challenge was to embrace multiple evaluation dimensions and metrics that are at play during humanitarian action. These include the areas of humanitarian ethics, protection, process quality and technical indicators for operations. Table 1 further illustrates Sphere's key evaluation dimensions and criteria. While the table only offers a selective snapshot of the almost 400 pages that make up Sphere, it provides an outline of the four broad dimensions for the evaluation of humanitarian performance around which the Handbook is designed.

Table 1: Distinct Modes of Evaluation within Sphere

Evaluation Dimensions	Ethical Standards (Humanitarian Charter)	Quality and Process Standards	Technical Standards (Selected)	Protection Standards (selected)
Evaluation Criteria	Right to Live with Dignity. Right to Receive Impartial Assistance. Right to Protection and Security.	Standard 1: People-centered Humanitarian response. Standard 2: Coordination and Collaboration. Standard 3: Response Assessment. Standard 4: Response Design. Standard 5: Performance, Transparency and Learning. Standard 6: Aid Worker Performance.	Nutrition: Indicators: 2100 kcals/person/day; Food Consumption Score; Under 5 Malnutrition Rate. Method: Food distribution is responsive, transparent, equitable and appropriate to context. Water and Hygiene: Quantity: 15 liters/person/day. Quality: No more than 10 fecal coliforms per 100 ml Hygiene: Maximum of 20 people/toilet. Health Care: Crude mortality rate: below 1/10,000 per day Measles vaccination coverage: 95% of children from 6 months to 12 years Medical Infrastructure: 1 health facility per 10,000 population. Shelter and Non-Food Items: Shelter area: 3.5 – 4.5 m² covered area per person.	Principle 1: Avoid exposing people to further harm as a result of your actions. Principle 2: Assist with rights claims, access to available remedies and recovery from abuse Indicator: report on percentage or number of abuse cases reported and responded to. Principle 3: Protect People from Physical and Psychological Harm Arising from Violence and Coercion.

However, engaging with and working with numerous interlinked evaluative dimensions is no easy task and poses challenging demands on users. Designers 3 and 4 describe the complicated nature of engaging with and applying Sphere's philosophy:

> Sphere is about looking at the four technical chapters, about what people have the right to when we talk about the right to humanitarian assistance. But it is also about the 'how' of aid provision. It's also very much about the core standards, it's about the protection principles, it's about all of the qualitative indicators that refer to who has the right to participate in decisions and who should be consulted (designer 4).

> So I think for me, the Sphere handbook, the aspects of the handbook, the right to live with dignity, and then getting on to 'in order to make this happen there are some life-saving sectors you need to address and within the life-saving sectors this is what is needed' (designer 3).

Yet, while Sphere specifies metrics and indicators for each of these evaluation dimensions in great detail, it emphasises that none of these categories should be considered in isolation. The following excerpts draw attention to specific elements of its format and the way the Sphere Handbook should be used:

> All the chapters are interconnected. Frequently, standards described in one sector need to be addressed in conjunction with standards described in others (Sphere Handbook, 2011, p.8).

> I think we need to work on the promotion and the articulation of the inseparability of the indicators and the principles and the ethical aspects of Sphere (designer 3).

The Protection Principles and Core standards are grouped together at the beginning of the Handbook so as to avoid repeating them in each technical chapter. They underpin all humanitarian activity and must be used in conjunction with the technical chapters. They are critical to achieving the technical standards in a spirit of quality and accountability to the affected populations (Sphere Handbook, 2011, p.6).

Hence, while each of Sphere's performance evaluation dimensions builds on the other, the logic of their association is described as circular, which points to its non-hierarchical and process oriented nature. There are several reasons for adopting this approach. Instead of proposing a one-size-fits-all formula to measure properties intrinsic to humanitarian performance in every setting, Sphere specifies that the standard can only be met when its four evaluation dimensions are put in relation to each other in the practice of contextualizing the Handbook. Some of the implications of this approach to evaluation is described by designer 7:

You have to choose which angle you are going to go in and you have to be able to connect them, analyse and think for yourself. It is important to apply Sphere and to think of it from the perspective of all of the cross-cutting issues together as an integrated whole and not just focus on one of the several silo-ed worthy causes that are competing against each other for attention (designer 7).

Sphere's propositions in relation to the development and engagement with performance indicators as well as its process-oriented approach towards performance evaluation provides unconventional suggestions for humanitarians. Its approach entails the abandonment of strict cause-and-effect-thinking (Burchell et al., 1980), a minimalist attitude towards control

(Quattrone and Hopper, 2005), and the pro-active fostering of spaces of ambiguity (March, 1987; Stark, 2009). All of these elements can undoubtedly be uncomfortable for humanitarians who operate under great pressure and seek a more straightforward and prescriptive technique. Against this background, a crucial question relates to notions of compliance, which is a central concept for donors that are commonly remote from the contexts of humanitarian response operations. In other words, in light of Sphere's adaptive approach to performance evaluation, how does it address notions of compliance and control? This question is addressed in the final part of this chapter.

5.2.5. *Problem 4: Towards Adaptive Compliance and Control – Treating Delocalized Knowledge with Caution*

In both the accounting and the development literature, issues of compliance and control are frequently treated as a problem of action at a distance (see Robson, 1991, 1992; Miller 1990). In his study on technologies of representation in development cooperation, Rottenburg (2009: xxiii) describes the issue as follows:

> Both development cooperation itself and the organizational structures it is supposed to set up aim to establish reliable technologies for remote sensing, monitoring and control, which enable organized action from a distance that is independent of local loyalties and priorities.

Rottenburg's (2009) quote insightfully connects the issue of action at a distance with the delocalization of knowledge. If donors want to ensure that their plans, on which funding requirements are commonly based, are complied with, they need to develop technologies that help them accumulate knowledge over the actions of implementing agencies in remote places. To make activities comparable over distinct settings, donors rely on the production of inscriptions that are largely independent of local cultural, contextual or political factors. The transformational processes that form an intricate part of producing inscriptions are thus linked to a delocalization of knowledge (Latour, 1999; Rottenburg, 2009).

In line with its adaptive approach, notions of long distance control and de-localized knowledge do not constitute Sphere's central concern and are treated with caution. In the following excerpts from the Handbook and quotes from designer 6 and 3, some key elements of Sphere's approach towards compliance and control are described:

> The Sphere Project does not operate any compliance mechanism. There is no such thing as signing up to Sphere, a Sphere membership or any process of accreditation. The Sphere Project has consciously opted for the Handbook not to be prescriptive or compliance oriented (Sphere Handbook, 2011, p.8).

> Sphere was always opposed to a centralized verification of performance against the standards (designer 6).

> Sphere will not be promoting certification because we believe in the voluntary approach (designer 3).

Many unforeseen performance requirements commonly emerge during the course of humanitarian response operations. The idea to pre-specify fixed and delocalized performance objectives against which agencies would be required to report is thus often unrealistic at best and at worst can encourage agencies to blindly comply to donor requirements while ignoring important conditions and needs on the ground. Against this background, rather than focusing on the production of de-localized knowledge in a remote centre of calculation, Sphere's approach to performance evaluation practices seeks to stimulate context-specific and practical judgement and communication through engagement with the different evaluative dimensions and indicators in the Handbook. The following quotes and excerpts highlight this approach:

> Sphere is used in many different ways. So while the donors may not ask you to report on Sphere standards they might ask you how you plan to adhere to Sphere standards. So you have to explain it to them, which is a good thing (designer 3).

> When users treat the Handbook as if it had all the answers, it loses its flavour. It becomes powerful in the hands of an informed actor on the ground… It is not about a perfect scenario where everything has been achieved. Sphere was not only about humanitarian action itself but it was also about humanitarian activism (designer 3).

Complying with Sphere thus does not entail the fixed application of static metrics for compliance, but, on the contrary, its rules might be employed as rough guides which retain flexibility and encourage improvisation. Against this background, Sphere is described as a common language and a type of

performance evaluation lexicon that enables multiple stakeholders and groups to communicate with each other:

> The Sphere Handbook…offers the humanitarian sector a common language for working together towards quality and accountability in disaster and conflict situations (Sphere Handbook, 2011, backside of cover).

> It was absolutely necessary for us to have some kind of lexicon that enables us to discuss and evaluate quality. And that needs to be linked to a discussion also about accountability and certainly I think the latter thing struck me as being the casualty of the absence of standards and effectively it is impossible to hold anybody accountable to anything when you don't have a common language related to performance (designer 1).

As emerges from these quotes, Sphere is conceptualized to provide a template for the terms of engagement between the heterogeneous stakeholders in humanitarian response operations. As stressed previously by designer 3, while donors might not always require agencies to report on Sphere's indicators, it is a common requirement to explain how agencies plan to adhere to Sphere. One of the possible roles of Sphere is to provide a lexicon that facilitates a discussion and communication about contextual parameters which help define and evaluate humanitarian response requirements and actions. In line with the notion of Sphere as a lexicon or dictionary, the following quotes further highlight the importance of creating narratives to explain the complexities of response operations, instead of narrowly focusing on technical expertise:

In relation to the issue of communication and information, we relied on something Plato said. He said: 'Those who tell stories rule societies.' That's important. Because the stories can move governments, the stories can move aid agencies and the stories can move the media. You control, you regulate you respond to reality based on information. So it's about the power of information. When you look at the Sphere standards, when the last handbook was launched in London I was invited to make a speech in 2011. And what struck me was that Sphere should continue doing what it does. But I also recognized that the 4 areas were not the only lifesaving areas. I wanted to put in a chapter on "Information as a Life Saver" because today appropriate information at the appropriate time is a life-saver (Designer 3).

Sphere is part of a broader discourse about the role of humanitarian aid. It needs to be kept within a much bigger debate about how aid is understood, how aid is impacting affected communities. Probably the most important and defining issue is listening to how those at the receiving end actually experience this engagement (Designer 1).

These quotes insightfully highlight the challenges of constituting meanings through the creation of narratives, and how Sphere's designers envision its role in these processes. Designer 3 strongly implies that the narratives should be developed by agencies and other informed parties on the ground, rather than by distant donors. While this approach is consistent with an adaptive approach to compliance, it also raises important questions. How does it influence the process of negotiating the meaning of compliance for the multitude of organizations involved in providing services in humanitarian response operations? What are the shortcomings of this flexible and adaptive approach? How are organizations that operate outside the contextually established boundaries of performance, for example highly corrupt organizations, controlled, checked and held to account? It is towards questions

such as these that case studies two and three on the enactment of Sphere in a particular humanitarian crisis now turn.

5.3. Case Study 2: In Search of Dignifying Qualities – Organizing Nutritional Supplies in a Large Scale Refugee Camp

5.3.1. Introduction to the Case Study: Organizing Nutritional Supplies for Zaatari Refugee Camp

This case study focuses on techniques and principles that allow humanitarian disaster response managers to evaluate actions and deal with the complexity presented by the operational requirements of managing Zaatari refugee camp. In particular, it focuses on exploring how Sphere's evaluation system is translated and informs camp managers in confronting common challenges, ranging from detecting emergent issues, exploring options and recognizing and exploiting contextual opportunities. The case study takes place in a period of transition from the beginning of the disaster as a result of the Syrian Civil War towards a more permanent and protracted state. As Zaatari's refugee population grew rapidly to over 100,000 people within less than a year, it also started to experience significant security issues, becoming one of the most unstable locations in all of Jordan.

One of the key areas around which conflicts and violence frequently surfaced was the issue of managing nutritional programmes, which is one of the most

important and consequential areas of managing humanitarian crises. The following quotes highlight some important dimensions of this context:

> The main thing at that time was that around food there were big tensions. Every time we distributed food we had demonstrations, often violent ones, it was absolutely crazy (senior camp manager 1).

> There were times when it was extremely dangerous for anybody to enter the camp. Mr. [Interviewee KK] started joking with us, calling us 'the rabbits' because we sometimes had to run like rabbits when refugees were throwing stones at us. The refugees did not want to tell us their true name. They suspected that we cooperate with the Syrians, that we might tell them where they stay. So as a result we often did not know how many people lived in the camp. That gave us a lot of problems for security reasons, but also for camp administration… It took a long time to establish trust (field officer 1).

Some of the initial challenges of setting up a nutritional supply programme in the camp were addressed by reference to a few simple indicator calculations from the Sphere Handbook (2011). The nutritional programme's value was priced alongside a variety of dimensions that were perceived to be crucial. The following quotes highlights the importance Sphere has in the daily activities of humanitarian response operations:

> What Sphere does… is to provide all the different evaluation criteria for what we are doing […]. Think about the technical domain, food, water, shelter, health. The Handbook provides specific indicators on what to consider and whether we are coming short. But it is not only that. There are also chapters for the evaluation of process quality, and the ethical considerations […] and there is protection monitoring […]. Sphere provides us with specific indicators to see whether we might be causing the refugees more harm (field Consultant).

When thousands of people started coming over the border every night, in this initial response stage everything was driven by technical indicators. Are we able to provide for these masses of people? Without that, everything else becomes useless. All the initial calculations were based on the 2100 kcal indicator (senior camp manager 1).

The 2100 kcal indicator is one of Sphere's most basic and widely used technical performance indicators. It provides an estimate of the minimum requirements per person per day for a disaster-affected population to be used for the planning, budgeting and monitoring of nutritional aid rations. In the context of Zaatari's nutrition programmes, managers used the indicator to construct a road map linking up technical concerns, operational budgets and fundraising appeals (see table 2). The road map not only quantifies the response's overall monetary requirements for specific periods of time and its overarching key objectives but it also provides a template for the break-down of each agency's financial requirements. Senior camp manager 2 provides a detailed explanation of how the process of linking Sphere to financial planning works in practice and how operational costs are broken down per beneficiary, for inputs, outputs and operational matters:

While Sphere does not have a financial component, it guides financial planning. So if we want to achieve Sphere we need to budget for the indicators and we need to budget for a way how we achieve these indicators. So we have a budget and it is coordination, operations and management cost added to quantity, quality checks and delivery costs. So you see there is an input based cost calculation and an output cost calculation. In most cases the difference between this is the operational cost. So we take all the inputs, which is for example we take $ 2,000,000 to achieve X amounts for the refugees that is the input cost per refugee. And then we go to the output base and we say each refugee receives X

goods multiplied by X amounts delivery cost and quality checks which translates into X amounts per unit. Once you get the difference between input and output cost you get the operational cost, which is how much we spend on staff and training and so on. And this is where it makes an interesting calculation, because the output costs generally gives you a whole sum of everything while the input costs are related to the indicators in Sphere. This is just to show how Sphere is very important for our financial management (senior camp manager 2).

However, while the road map provided a space that allowed camp-managers to translate the many complexities they were facing into more abstract and relatable parameters, none of these explained or addressed the unrest Zaatari was experiencing:

In fact, there was a discrepancy between the issue that for once we were able to deliver on the main life-saving standard indicators and yet Zaatari was a very unsafe place with very unhappy people. So in fact, the technical measures did not match with the psyche of the people living in that place. There was something wrong. This was what nobody understood. When I came everybody said: we achieved everything so why are they unhappy? Why are they rioting, why are they throwing stones? (senior camp manager 1).

As follows from the testimony of senior camp manager 1, that the camp was experiencing a far-reaching struggle for authority that had to be taken into account if progress was to be made. This struggle violently exposed the limitations of a reductionist and linear approach to managing and measuring humanitarian performance.

Table 2: Road-mapping Nutritional Refugee Assistance for Refugees

Agency	Total Jan-Dec 2014	Jan-Jun 2014	Jul-Dec 2014
ACF	750,000	450,000	300,000
ACTED	1,700,000	1,020,000	680,000
Caritas	2,188,644	1,271,186	847,458
FAO	6,500,000	4,400,000	2,100,000
JHCO	4,012,500	2,407,500	1,605,000
UNHCR	1,489,199	893,519	595,680
WFP	305,050,000	141,848,250	163,201,750
WVI	500,000	300,000	200,000
Total	**322,120,343**	**152,590,456**	**169,529,887**

Source: UNHCR (2014a, p.49).

5.3.2. *The Role of Tensions between Evaluative Principles in Interrogating and Detecting the Unexpected*

As was previously emphasized, instead of pursuing an unrealistic vision by attempting to fully anticipate and control the range of interconnected issues at play, humanitarian evaluation systems need to be judged as much by their ability to challenge taken-for-granted categories and re-appraise expectations as by their capacity to provide stable indicators and calculations. For the management of humanitarian crises, it is of critical importance to work with

systems that embrace and even routinize the exercise of questioning, doubt and reflection (March, 1987; Quattrone, 2015b; Mouritsen, 2016). The clash between the approach taken in Zaatari and the unrest that followed offers an opportunity to investigate and reflect on these issues. It forced the managers to re-consider what was at stake around an issue that appeared relatively straightforward: the organization and delivery of food and nutritional aid items.

To explore the challenges surrounding the camp's problems, in the cluster meetings where all agencies working on nutrition regularly meet, Zaatari's managers went back to the basics of Sphere's philosophy. Senior camp manager 1 describes some elements of this process:

> The big question now became: how can we support people in that particular desire to allow difference in a situation in which we have to deal with masses and big numbers and general logistics. The problem is that you are rarely equipped to deal with all of Sphere's different interconnected performance requirements. It exposed that the refugees simply had a different conceptualization of the space. They had a different idea how their settlement should look like. It didn't match with our planning, with our thinking (senior camp manager 1).

The evidence raises several issues. Instead of offering a clear-cut prescription, in the engagement with Sphere a number of new questions emerged. One question was to understand how Sphere's evaluation requirements are connected, how they might be balanced and what sort of trade-offs existed between them in the particular setting of Zaatari. In particular, Sphere's requirement to treat its distinct evaluative dimensions as inseparable is

mentioned as an important factor. As was outlined in the case study on Sphere's design, while each of Sphere's performance evaluation criteria builds on the other, the logic of their association is conceptualized as circular. The heterarchical arrangement of Sphere's evaluative principles hereby provided the effect of working against a tendency to oversimplify response requirements and shifted the focus on detecting gaps in understanding. This point is further emphasized by senior camp manager 2:

> Having these multiple performance criteria competing for our attention is improving performance because it is keeping all of these checks in play and it is keeping the humanitarian body on guard on what we are supposed to be doing (senior camp manager 2).

While the previous financial plans were based on an effort to make pre-defined expectations work, this process became about challenging and re-defining expectations in a more open-ended manner. As Senior camp manager 1 put it:

> This stuff is so complex and the problems we encounter on an everyday basis so complicated that it can be quite overwhelming without a reference point (senior camp manager 1).

Even though many of the managers involved in the response operation had significant amounts of operating experience in humanitarian crises, some of them more than twenty years, Sphere began to serve as a guide to make an inventory of options and provided the terms of engagement for the managers to confront this novel situation.

As Zaatari's managers sought to gain a better understanding of the factors contributing to the camp's unrest, a range of specific issues emerged. While the above problem description evidenced challenges generated by an over-regulated manner of attempting to control people's food consumption, it also simultaneously demonstrated a significant absence of control and a lack of ability to respond to vulnerabilities in the camp. Even though Zaatari's managers had begun to conceptualize the operation's key requirements by linking the technical evaluation dimension to financial concerns, contextualizing and enacting Sphere in the camp's nutrition sub-working group meetings resulted in the emergence of other tensions between its evaluation dimensions. Senior camp manager 1 describes the process as follows:

> This was quite an unusual scenario. We appeared to meet most of our technical indicators and at the same time we lost the camp to gangs, protests and bandits. We delivered more aid, and the result was more problems. When we began to talk about this in the cluster meeting we asked: have you looked at the other performance chapters? Did you look at the process standards? Did you look at the protection metrics? The answer was silence. This was an important moment (field manager 1).

The quotes highlight the general challenge of balancing requirements from the technical, ethical and process evaluation dimensions within the context of mass distributions and the need to construct a liveable space in the camp that allows for difference and dignity. Even though the managers had employed one of Sphere's important performance indicators, they did not follow the standard's philosophy in which value is not conceptualized as essentialist but

as relational. In Sphere's terms, performance evaluation is not about representing pre-given notions of value but about instigating processes of evaluation. Evaluation practices in humanitarian contexts do not exclusively focus on economic value. All economic goods in Zaatari camp are inseparably moral goods. While a focus on technical and financial elements is crucial to sustain a response operation, an excessive reliance on these dimensions can undermine performance by encouraging oversimplification.

Several key points can be summarized at this juncture. In the process of investigating challenges and conducting an inventory of options, the tensions and ambiguity between Sphere's evaluative dimensions played an important role for several reasons. Firstly, focusing on tensions and trade-offs proactively fostered doubt and encouraged attention to inconsistencies and unknown elements in the camp. Unlike in Chenhall et al.'s (2013) study, which focuses on the question of how accounting systems might forge compromise between different evaluation criteria, it was precisely because actors did not pursue compromise that tensions became informative. While the initial technical and financial plans were crucial to conceptualize the response operation, the plans also lulled the managers into thinking that the response operation might evolve in the anticipated manner. The emerging ambiguities and tensions between Sphere's evaluative principles worked against such a deterministic attitude, thus sharpening the managers' senses towards possibilities and problems. In other words, while compromise normalizes and stabilizes, focusing on tensions encouraged a focus on anomalies and gaps (Weick and Sutcliffe, 2007).

Secondly, the tensions did not simply emerge out of coincidence or due to the complexity of the setting, but developed through the interaction with Sphere's format and approach. In particular, senior camp manager 2 highlights how the process of connecting performance requirements on the drawing board raised questions and led to discussions. This observation suggests that tensions between evaluative principles are not informative per se, but must be actively fostered in an organized and strategic manner. While this description is in line with the notion of heterarchy, it differs from Stark's (2009) ethnographies, in which multiple evaluative principles appear to form part of all of his research settings and play a similar role in each of them. Sphere thus attracted users because of its capacity to challenge taken-for-granted plans and assumptions and evoke additional meanings beyond what was knowable at this point (March, 1987). Thirdly, tensions and ambiguity between evaluative principles became informative for Zaatari's managers at a particular point in time, namely in their exploration of shortcomings, gaps and options. This points to the importance of timing in the engagement with evaluative tensions. While several accounting studies have pointed towards the enabling potential of deliberately ambiguous accounting systems, this temporal element does not form an important focus of these studies (see Cooper et al., 1981; March, 1987; Stark, 2009; Chenhall et al., 2013; Quattrone, 2015b; Revellino and Mouritsen, 2015). However, as demonstrated in the following section, exploring and interrogating the tensions, links and ambiguities between evaluative principles was not the only element that was necessary to assess performance and seek solutions for the challenges of Zaatari.

Evaluation as a Participatory Process

As emphasized in the section on the development of Sphere, designing for complexity and uncertainty entails recognizing the inherent limitations of a performance evaluation system in being able to anticipate the diverse range of response requirements across distinct humanitarian crises. However, it takes a specific evaluative infrastructure to discover emergent issues before they become unmanageable. While Sphere specifies that it can never be met without involving affected populations in the decision-making and evaluation process, this was one of the key elements missing from the response. This guideline further builds on an important implication of navigating the intrinsically dynamic humanitarian environment: the impossibility of pre-specifying the problem, to identify pre-exiting alternatives and optimal solutions. Without an in-depth understanding of the concerns of and demands from the various groups in the community, and in particular its most vulnerable members, no remotely constructive account of the camp's challenges would be possible. Senior camp manager 1 and evaluation officer 1 explain some of the challenges managers were facing in this regard:

> Remember there are no communities at the beginning, you need to build them. So representation is dangerous because you are creating power structures as well. But there are ways of involving the refugees into decisions without creating abusive power structures. You have to build a community first (senior camp manager 1).

> I mean there were no rules in the camp. The refugees had a total rejection of rules   because they associated that with dictatorship…When I started in the camp, communication between

humanitarians and refugees had completely broken down (senior camp manager 1).

When they came, they were fleeing violence and more. Most of them are traumatized and shocked of what they have seen. They were running for their lives. As such, they were not psychologically stable (field manager 2).

As stressed below, as the managers began to enact Sphere's prescription of community participation, the performance evaluation activities began to no longer exclusively reside in the fenced trailer offices of the agencies in the humanitarian village at Zaatari's outer edges. Instead they started to evolve into more dialogical forms:

Left to our own devices we tend to shut out the people who are receiving our assistance. They are not seen as an equal stakeholder, they are not seen as a, they are seen as, you know, communicating with communities or consulting with them is sort of a luxury and an add-on if you have got a bit of extra time. So, you know, we wouldn't need to have things like Sphere if people fundamentally kind of believed and saw the importance of regarding people at the centre of humanitarian action, as equal stakeholders in the outcomes (evaluation officer 1).

As a result of this engagement, an additional feedback loop was added to the performance evaluation activities in the camp. By placing the affected populations at the centre of the humanitarian response, Sphere encourages the set-up of parallel structures for information collection and processing throughout the refugee camp:

Now there is a very significant dynamic in the camp for community participation which Sphere advises us to do and what should be common sense humanitarian practice. All the partners meet with the refugees in specific groups, one for males and one for females, in which they speak about what requests they have and how we respond to them. So refugees are holding us accountable, they are participating in the decision-making in the camp and also we are holding them accountable, telling them 'if you expect us to do this then these are your responsibilities.' So this means we are keeping this feedback mechanism in place, which is telling us how we are performing on a regular basis (Senior camp manager 2).

So we involved them [the refugees], having really constructive sessions with them on how to improve performance and services. And they came up with suggestions on how to distribute the food items in a more equitable manner (Senior camp manager 1).

Several important elements emerge from the quotes. Firstly, the limitations of evaluation systems for the management of humanitarian crises is stressed. Due to the messiness of humanitarian settings, the most crucial information is usually not to be found in the information systems. To overcome these limitations, setting up mechanisms to conceive of evaluation as a participatory process with the refugees was an important factor to gaining additional information sources. Senior manager 1 hereby describes the importance of the refugees in coming up with innovative ideas for the organization of the camp. Secondly, these participatory mechanisms had the effect of building trust between humanitarians and refugees. Thirdly, the participatory mechanisms started a discussion of the allocation of tasks in the camp through which the refugees would take a more pro-active role. Fourthly, the participatory mechanisms provided a platform for refugees to hold agencies accountable and vice versa. This process that was guided through engagement with Sphere did not only flatten accountability relationships in the camp, in which refugees

otherwise had little say about decisions affecting their lives, but it also made it more difficult for agencies to hide from criticism. Through this process, Zaatari's managers slowly uncovered further important issues:

> We discovered that we distributed many items that people did not need or did not like. So these people were selling these commodities cheaply. I mean you simply cannot live with 6 or 7 items per day. So people also have to sell some of this stuff to get the items that they really need. That in fact made people very vulnerable to the crooks (senior camp manager 1).

As emerges from the evidence, the fixed way of budgeting for, organizing and distributing nutritional supplies did not suffice for people to make a living. The lack of choice and the inflexibility made it necessary for the refugees to trade with some of their aid items to get access to elements they needed. The result of this was the emergence of a significant black market for goods and services. In the specific context of the camp, this created additional power structures in which gang-like groups formed and these controlled and traded access to rare supplies:

> There was the problem that some big guys managed to get hold of some of the ration cards of people that were not in the camp anymore. In a certain way, we contributed to the rise of these gangs, because they financed themselves through the stuff we distributed. As the gangs became stronger, the weaker people started losing out. That meant that they were selling the food rations for these people making about 30.000 dollars each day (Senior camp manager 1).

These gangs, which financed themselves by selling goods that were donated to assist Zaatari's refugees, were partly responsible for the problem of the camp getting out of control. Hence, Zaatari's managers found themselves in the paradoxical situation that the aid they were delivering actively contributed to the formation of crime and generated further vulnerabilities in the camp.

The community participation meetings that were shaped through the guidance of Sphere's evaluation system served as a form of lateral performance information gathering and started a debate about the organization of important spaces in the camp of which nutritional programme performance formed an essential one. A key problem was to strike the right balance between control and self-organization:

> The question is how much space and freedom can you give someone and where do you limit the capacity? We had to re-think the relationship between the space, the services, the culture and the economy of the camp (senior camp manager 1).

To allow more flexibility and to reduce their footprint on the ground, a more direct, reliable and durable relationship between humanitarians and refugees had to be established.

As emerges from the analysis in this section, setting up mechanisms for evaluation as a participatory process was itself valuable for several reasons. Firstly, rather than conceiving of evaluation as a linear process in which pre-specified categories are assessed, Sphere encouraged the set-up of mechanisms

that fostered doubt and encouraged a focus on processes that exposed anomalies, inconsistencies and errors (Cooper et al., 1981; March, 1987; Quattrone, 2015b). This is of particular importance since most crucial information is not to be found in the formal information systems due to the complexity of humanitarian crises. While Sphere's performance categories guided evaluation processes, the categories remained open for re-definition and re-combination. In other words, Sphere's performance categories served to organize the terms of engagement but did not constitute a fixed end to be achieved. Secondly, as was emphasized in the chapter on Sphere's development, one of the key aims of its designers was to prevent a further alienation and stigmatization of refugees. Conceiving of evaluation as a participatory process not only provided additional information sources for the managers in relation to vulnerabilities in the camp, but also offered a channel for refugees to voice concerns and criticise agencies. In this context, the interaction amongst heterogeneous groups with different types of expectations reduced oversimplification and helped to make the most of the distributed intelligence throughout the camp (Weick and Sutcliffe, 2007; Stark, 2009).

5.3.4. _Adapting Humanitarian Performance: Re-combining Incomplete Templates for Evaluation_

The first two sections of this chapter were largely concerned with techniques and mechanisms that helped Zaatari's managers to detect emergent issues and explore options in the engagement with Sphere. However, the response operation also had to be adapted to address some of the camp's urgent

challenges. This section explores evaluation techniques that allowed Zaatari's managers to adapt to and exploit contextual opportunities. The following three quotes highlight several important elements in the beginning of this process:

> The reality is that Sphere is not written in stone. We have to contextualize it. Sphere is not perfect […] The way we worked with Sphere became more about adapting and contextualizing its recommendations. It wasn't actually about meeting everything in the Handbook. There might be some situations where we are unable to meet Sphere. And what we then do is to go back to the drawing board, look at the standard, look at the implications why we are not able to meet the standard, what is the problem, what is missing (senior camp manager 2).

> There are lots of bits about the Sphere standards that are not perfect. But that is mainly because of the challenges humanitarians have to face …humanitarian work is very complex, and to boil Sphere down to something that is supposed to not look complex can be very difficult in regards to application in complex environments (humanitarian consultant 1).

> Having a book that provided a place for discussion, not a place for decisions, not a place for right or wrong, was probably one of the most valuable things Sphere has done (evaluation officer 3).

In the section on the development of Sphere, it was frequently pointed out that its designers actively sought to prevent a dogmatic attitude to the engagement with its metrics and evaluative dimensions. In the above quotes, specific reference is made to the enabling role of imperfection in the practice of Sphere (see Dambrin and Robson, 2011; Jordan and Messner, 2012; Busco and Quattrone, 2015). Sphere's evaluative dimensions and metrics are described as

rough and open templates for the engagement of humanitarian crises. The perceived imperfection of these templates appears not as a factor to be eliminated, but instead as a precondition for contextualization, adaptability and recombination. Since Sphere's evaluative dimensions are conceptualized as highly interconnected, adaptive improvisation and tinkering is actively encouraged. Such an unconventional approach to performance categories is not trivial. As Weick and Sutcliffe (2007:58) emphasize:

> Although categories are unavoidable, we can carry them more lightly. If you want to carry a category more lightly, you need to both believe and doubt it. If you can do that, you are that much closer to wise action.

Using performance categories and building on their known imperfections in such a way can have several effects for organizations operating in crisis situations. Firstly, they can be treated as a pretext or method to investigate blind spots. Secondly, they can serve as an organizing mechanism to question the expectations that underlie these categories and compare these with observations made. Thirdly, building on the two former points, they can provide a means to question whether contextual observations and variations suggest the emergence of *new* adjusted performance categories.

In line with these insights, as Zaatari's managers continued to enact Sphere prescriptions, several key issues for the re-organization of the nutritional aid supply system emerged, differing significantly from the initial technical-financial framing of the issue:

The key driver for change was to give people their dignity back, which means they will behave normally. Sphere asks us to contextualize what dignity in Zaatari could mean. You heard about the violence. Violence will significantly reduce if people have dignity (field manager 2).

As the technical requirements of resource efficiency became associated with Sphere's ethical evaluation dimension, the need to re-organize nutritional supply programmes suddenly became not only a practical need in order to deliver the programmes without violence but it was also a matter of giving back to refugees a sense of dignity. The practicalities of setting up and running nutritional supply programmes thus had to be conceptualized as part of this broader narrative of the refugees' struggle. The lack of choice and participation reinforced the notion of 'begging for food' that became strongly linked to a further loss of dignity. Financial efficiency hence became associated with the concept of human dignity:

Meeting Sphere's quantitative indicators was never a big problem in principle...but speaking of the concept of dignity, Sphere is quite clear on that…so let's say, I lost my home, I lost half of my family members, and I ran away with my wife and my daughter. Now I am not used to someone giving me food and say 'now you are going to eat this and that'. I am used to buying my own food. Now I am in a camp, so of course I have to queue and it's not even the food that I like. Begging for food that they don't want to eat is not dignity for the refugees. And this was where we lost them (field manager 2).

Sphere's ethical evaluation dimension specifies the right to live with dignity as its first criterion. In line with this approach, Sphere puts the challenge to search for the meaning of human dignity in a specific disaster situation as one of the crucial tasks confronting humanitarians. Since dignity is not defined in

essentialist terms, it can only be created relative to Sphere's other evaluation dimensions:

> So typically when you are working with people the right to a life with dignity then suddenly gets kind of fluffy. It's hard to articulate what exactly that means. Then you hold up the [Sphere] handbook and say 'this entire book is an attempt to articulate what that means.' So the right to a life with dignity is this broad overarching concept but I think it is useful because you never remember the nitty-gritty bits. But it only really makes sense when you articulate it in detail in terms of practical implications for your work in the camp (field officer 2).

To be able to move ahead, bridges between the different evaluative principles had to be built. The incomplete and interconnected nature of Sphere's evaluative dimensions encouraged managers to pay attention to the links between these elements, helping them to create an integrated conceptualization of operational issues instead of focusing on isolated concerns. Thus, working with Sphere meant keeping evaluative principles incomplete and therefore open to contextualization and recombination.

Against this background, an innovative solution emerged which would have been impossible if Sphere had prescribed a tightly coupled and inflexible system for evaluation (Weick, 1976, 1988; March & Olsen, 1975). After a consultation with a private contractor, an operations team came forward with the suggestion to install supermarkets inside the refugee camp, a solution that had not been previously attempted elsewhere. Zaatari's managers quickly realized the potential of this new way of managing nutrition, which would reorganize the previously hierarchical and fixed manner through which

nutritional supplies had been delivered to refugees. Firstly, it allowed the contextual recombination of two important performance requirements: Sphere's demands of financial and technical efficiency and its ethical evaluation dimension. Accordingly, organizing nutritional supplies through supermarkets allowed the community to regain a sense of choice and dignity:

> They push their trolley, and they buy the food they actually want. Now, you may see it as a normal thing, so what? But no, if you go to the supermarket and you see the pride refugees feel there. So the father will get together with the mother and they take their kids and they go shopping. Before they felt like begging for food. But when you go to the supermarkets you can understand it. The dignity and pride of 'I am buying this because I want it' (field manager 2).

Organizing nutrition through supermarkets also freed Zaatari's managers from having to handle mass distributions. This not only "reduced overhead costs significantly" (senior camp manager 2), but also allowed for their re-investment in food ration allowances, increasing them from $25 to $45 per person per month. This facilitated the integration and re-conceptualization of several other key performance requirements that had emerged:

> It means that humanitarians reduce their footprints on the ground. The cost and associated operational input required for this was becoming too high (senior camp manager 2).

> Instead of refugees having to queue up for half a day to get food rations we moved away from this. Now we don't do in kind food donations. We now give them a voucher they can redeem at the two supermarkets, Tazweed or Safeway [located inside the camp] (field manager 2).

Once you have calculated this [the food rations] it is their business for god's sake. This is where we have an incredible arrogance to tell people what to do, to monitor, to measure. That is a major problem, the attitude of the aid business to, I would say, its victims (senior camp manager 1).

So a key issue was how to establish a place that would not be attacked, that would be safe. So an important issue was security (Senior camp manager 1).

Without mass distributions, the gatherings that led to some of the most violent situations the camp had experienced could be avoided, thus contributing to greater security. Not only was the organized space of the supermarket easier to police, but the elimination of the feeling of 'begging for food' made refugees less likely to react violently during food distribution. Moreover, organizing nutrition through supermarkets prevented unwanted goods landing on the black markets in the camp that were controlled by gangs. Refugees' exposure to gang-like groups running black markets had been previously identified by managers as contributing to additional vulnerabilities in the camp. Finally, putting in place a supermarket significantly strengthened trust between humanitarians and refugees. In this context, Zaatari became the first ever refugee camp in which nutritional supplies were organized through two unconventional supermarkets. This contextual recombination of existing resources was subsequently praised as a model innovation, one to be implemented in refugee camps around the world (Betts *et al.*, 2015).

5.4. Case Study 3: Evaluating, Controlling and…Democratizing – Sphere's Role in Re-shaping Water Network Governance

5.4.1. *Introduction to the Case Study: Mapping the Complexities of Water Network Governance in Zaatari Refugee Camp*

Building on the insights from the previous chapter, this case study takes a closer look at the enactment and translation of Sphere's evaluation system within the context of Zaatari refugee camp. It specifically focuses on techniques employed by humanitarian disaster response managers to evaluate actions and maintain operations while dealing with emergent and unexpected issues. As stressed earlier, the ability to confront crises commonly depends on technologies and expertise that developed before disaster strikes (see Weick and Sutcliffe, 2007). These techniques do not necessarily entail that managers are able to prevent problems from occurring or avert all human deficiencies. However, they might provide a basis for problem detection and solving as well as serve as a guide for reasoning.

The key challenge confronted by Zaatari's managers in this case study is the organization of water supply chains for the refugee camp. Zaatari's managers faced the problem of how to deal with and control corrupt agencies in the absence of a clearly defined and enforceable command structure.

> Water creates the biggest tensions in and around the camp. We have been experiencing continuous problems with all of Sphere's dimensions of managing water, including quality, contamination, taste, equitable access and safety (senior camp manager 2)

> Of course the Sphere standards have to be applied in every situation. For instance, for water there are indicators that we cannot go below. For water, the technical Sphere indicator would be between 15 and 20 litres per person per day. In this sense, Sphere guides the financial management and controlling of this operation. But the key issues in this operation were not really related to our technical numbers or finance (field manager 2).

The administration and organization of water supply chains constitutes the single most consequential task for Zaatari's managers, reaching into all possible areas of life in the camp. These include health care, sanitation and drinking, financial planning and budgeting, as well as infrastructure development. At the same time, with its location in the dry Jordanian desert, managing water supplies for the camp constitutes a significant risk factor for conflict between host communities, refugees, the Jordanian government, donors and the aid agencies:

> One of the biggest headaches was the sub-distribution of water. A significant problem is that it is extremely difficult to break it down into controllable operational details because they [the agencies] can do whatever they want. Control in this context is a big issue. The question that arises here is what is the governance of such a camp? If we talk about camp management and camp coordination: so here I am as a sort of unrecognized mayor in a sense, not having direct authority over 90 percent of the resources. It is as if the mayor of Vienna has no authority over the various departments to run the city (senior camp manager 1).

There is no one single actor in this camp who has a permanent state of power towards others. You see what I mean, you have to look at every single event by itself (field officer 3).

Supplying water to the tens of thousands of refugees in Zaatari is a vast operation. Senior camp managers 1 and 2 describe some of the basic complexities and key parameters surrounding the daily task of running and governing water in Zaatari:

Bringing water into the camp is a huge operational challenge: around 3600 cubic metres [3. 6 million litres] per day (senior camp manager 1).

Nearly one hundred water tankers come into the camp every day to truck in the water. We are doing 240 trips of water per day. To reach the target number, they often have to come two or three times. The amount varies between 3.6 and 4 million litres on a daily basis. The variation is because of several reasons. The amount of people in the camp varies. Also the government has a big say in how much water we are allowed to use. We need to keep Mafraq's [the district within which Zaatari is located] hydrological balance in tact (senior camp manager 2).

The quotes highlight some of the complexities of managing the camp's water operations. Transporting the required quantity of around 3.6 million litres per day to supply the camp constitutes a massive logistical challenge, requiring constant interactions between humanitarian agencies, private contractors, the Jordanian government and refugees. As described above, Sphere's technical evaluation indicator of 15 litres of water per person per day provided a simple benchmark around which considerations surrounding financial planning and controlling could be initially coordinated. However, as the camp managers

were conceptualizing and implementing water supply solutions for Zaatari's refugee population, a variety of challenges emerged, reaching beyond technical and financial concerns. Senior camp manager 1 highlights this point through a specific example:

> The NGO, [organization X], was pretending that it was checking everything that was going in and out, water quality when it was coming in and whether there was anything left when the truck was going out, including smuggled people and so on. Of course, they were not doing that properly. Every time I confronted them with it, they were saying 'no, it's all fine.' So you have to deal with this issue of institutional cover up to protect your contract and of course there are personal interests involved (senior camp manager 1).

> If somebody like [organization X] has a contract with [organization Y], receiving 20 million dollars to truck water in and out. If you receive this kind of money you do not want to be seen to be doing something wrong. You cover up because you want to look good with the one you are having a contract with. That organization is involved in the 20-million-dollar business. It frequently allowed water tanks to go in three times with the same water (senior camp manager 1).

The above-described factors demonstrate that relations of distributed authority are defining elements of Zaatari's setting. While the camp experienced problems with the sub-distribution and quality of and equitable access to water, there was no clear manner through which the operations of these stakeholders could be aligned and their actions controlled. As Zaatari's managers were wrestling with these immense complexities, the governance challenges they were facing manifested themselves forcefully in an alarming fashion:

> We caught this water tanker, which really delivered water mixed with faeces. The water tanker was also used as a sewage tanker, or it was a sewage tanker the day before. My team physically saw that when the water was delivered there was at first actually shit coming out of the truck. Unfortunately, we discovered that this was not really an isolated incident. This meant that we could no longer trust the agencies we were working with (senior camp manager 1).

The detection of the truck with contaminated water epitomized the biggest possible indicator of ethical and technical failure, suggesting that the current form of governing the water distribution was severely defective. Yet, as emerges from the quotes above, even in the face of compelling evidence of poor performance and corruption, little could be achieved through the prevailing system of controlling and governing the camp. While Zaatari had experienced a whole range of challenging issues in regards to the conduct of its daily life, including unrest and violent protests, the detection of the delivery of water mixed with faeces provided a turning point. Issues of contaminated water do not only entail a health and safety risk for the refugee population but also to surrounding communities as contagious diseases can quickly spread beyond the fences of the camp.

5.4.2. *Interrogating Evaluative Tensions and Exploring Solutions for the Camp*

As emphasized by Ramalingam (2013: 158): "a key characteristic of complex systems is for effects to propagate in ways that produce wide-ranging and long-lasting secondary consequences that may have little to do with the initial trigger." In line with Ramalingam's (2013) description, the challenges of Zaatari's water networks are characterized by the complex interconnected

nature of its agencies and the emergence of unexpected challenges involved in its management. In this context, Sphere gradually evolved into assuming an increasingly powerful role. Senior camp manager 1 describes some key elements of this process:

> What Sphere does of course is to set up the monitoring and evaluation criteria for what we are doing. But there was a real issue with Sphere being monitored at the agency level. What is going on with the water can affect health and nutrition, etc. So in situations where agencies have divided up: we will do water and you will do health, you know, this did not make sense. They were just monitoring on their own. There has to be a higher meaningful level on which monitoring is carried out. For us, the Sphere standards provided a basis for whole operation monitoring (senior camp manager 1).

Confronting Zaatari's issues involved a necessary shift from an individual organization-centred to a more integrated approach to performance evaluation, taking into account the characteristics of Zaatari's complex socio-technical network. To explore alternative options for Zaatari, a working group was created to translate Sphere into a framework to monitor and evaluate performance from a camp-level perspective. This working group included all agencies involved with water governance. Senior camp manager 2 describes how Sphere's evaluation system provided a template for wider engagement with the camp's problems:

> To practicalize Sphere we put overall sector working groups in place. For example, as a result of the problems we had with water, we set up a small working group to work on standards for WASH [Water, Sanitation and Hygiene] facilities. We looked at the Sphere standards, we looked at the conditions on the ground and we looked at the results

of our monitoring systems and then we started to discuss required actions and recommendations on what do we need to do. Then in the WASH sector coordination meeting they are reviewed and then presented to camp management (senior camp manager 2).

The core stakeholders working in WASH in Zaatari are categorized mainly into four. We have the implementing partners, [organization x and organization y], that are actually dealing with the day-to-day delivery of water. And then we have the coordinators, mainly UNICEF, UNHCR. And we have some external support, mainly the City of Amsterdam. And we have the government as a whole, the local municipality and the national government. So this combination of four stakeholder groups actually determines the majority of things. What Sphere did for us, it gave these very different stakeholders a way of engaging through a common tool, a shared vocabulary (senior camp manager 2).

A crucial purpose of the sector working groups was to elicit a process of questioning future options for the camp's water network governance. This process was guided by, in a first step, translating Sphere's evaluation dimensions into an integrated matrix, visualizing factors affecting the entire camp (see figure 3). The matrix not only facilitated a process that allowed the managers to assess their understanding of the challenges surrounding each of the evaluation dimensions, but it also provided an illustration of the inherent tensions and necessary trade-offs between them.

Figure 3: Interrogating Tensions and Exploring Solutions

Technical and Social Criteria	Definition	Risk Score Communal	Risk Score Communal +	Risk Score HH
Equitability	Ability to meet the adequate amount of water needed by all groups/HHs/individuals.	10	5	1
Accessibility	Ability to maximise adequate access to water supply as per minimum standard agreed at camp level.	10	5	1
O&M	Ability to minimise technical and operational challenges in O&M from boreholes to points of use.	5	5	10
Water losses (not misuse or wastage)	Ability to minimise water losses from boreholes to points of use.	1	5	10
Water quality	Ability to prevent water contamination or re-contamination from boreholes to points of use.	1	5	5
Implementation	Ability to minimise the social challenges during implementation (social and technical)	5	5	10
	Overall Technical risk score for the 3 options	32	30	37
WEIGHTING ACROSS ALL OPTIONS - Community acceptance and engagement	Ability to meet the preferences by the population in terms of service level, not only in design phase but also in implementation and in operation phase.	3	2	1
	Risk score for the 3 options	96	60	37

(UNHCR, 2014b, p.66)

As follows from the matrix, the competing evaluation dimensions drawn from Sphere involved issues of equitable distribution of water, access by all vulnerable groups, quality, misuse and abuse by agencies and refugees, "social" factors affecting implementation, and preferences by the refugee community. Importantly, technical and social evaluation criteria are not treated as separate but as mutually constitutive. Thus, the familiar value vs. values distinction that has shaped the social sciences since Parson's Pact (see Stark, 2009) became irrelevant as Zaatari's managers translated Sphere to the context of the camp. As described below, this strategy had several purposes:

> Sphere was our reference and our benchmark to develop our new evaluation frameworks. What the framework does is it guides us in the

question of how we can achieve the quality required as per Sphere standards considering the Zaatari context. The framework gives us an action plan that tells us 'this is the way we have to work to achieve the Sphere standards (senior camp manager 2).

When you explore options, it also depends on the host country. I mean the water needs for a country like Jordan, their water producing capacities are already very stressed, it's a very water scarce country which means that while you want to alleviate the suffering of the people as much as possible you also don't want to upset the hydrological balance of the country. So this means that you have a minimum limit and somehow a maximum limit also. So putting all of this into a basket requires looking at Sphere, but at the same time it requires us to look at the person, the human being, and say 'well this is what Sphere wants us to supply but this is not appropriate for this context. This is the background of why we have to adapt the measures and contextualize them. Once it moves beyond the sector to camp management, who are usually managers and not technical engineers, this is where it serves as a road map. If we want to achieve this, this is what we want to do, this is a problem, and that is a problem (senior camp manager 2).

In this context, rather than employing Sphere as a fixed set of metrics to evaluate performance, its approach and philosophy became a guide to construct a novel governance framework for the context of Zaatari. To make the tensions and trade-offs between the evaluation dimensions more concrete, the managers scored them against three possible future options for Zaatari that would move away from the problematic trucking system. While each of the distinct requirements, e.g. equity, accessibility, quality and other technical factors, is scored against the different options, they are subsequently weighted against the ability to gain community acceptance and engagement.

The emerging options are specified in the top right corner of the matrix: firstly, a communal solution in which water would be delivered to central points around the camp and shared by approximately eighty people. Secondly, a communal plus solution, which in addition to the communal option would involve the construction of extra water ports around the camp with a distance of fifty metres between each water port. Thirdly, the construction of a more permanent infrastructure with water connections for every household of approximately six persons. The latter was called the household solution. While the former two options constituted mere extensions of the trucking option, the household option would entail a radical decentralization in the way the distribution of water is governed in the camp.

One of the results of this process of searching and questioning was that the trucking as well as the communal and communal plus solutions were judged to be insufficient to address the key problems of the water network. While it was evident that developing infrastructure would not only be costly and require several years of construction, a distinct way of governing the distribution of water in Zaatari would also entail a far-reaching change in approach. To complement the inventory of options, the managers further conducted a financial viability analysis, which comprised projected cost components of total capital and operational costs for all options, the results of which are shown in figure 2 (see appendix 3 and 4 for further breakdowns of cost-projections).

Consolidated Costs

OXFAM						
OPTION	CAPITAL COSTS			Annual Power + O&M Costs		
	CIVIL (USD)	E&M (USD)	TOTAL (USD)	Elec Cost (USD)	O&M (USD)	TOTAL (USD)
Trucking	$ -					
Communal	$ 3,934,478.92	193,523.00	4,128,001.92	24,387.80	518,414.38	542,802.18
Communal +	$ 3,936,177.72	362,228.00	4,298,405.72	28,345.01	1,660,086.26	1,688,431.27
HouseHold	$ 9,907,185.42	362,228.00	$10,269,413.42	28,345.01	329,656.09	358,001.11

ACTED						
OPTION	CAPITAL COSTS			Annual Power + O&M Costs		
	CIVIL (USD)	E&M (USD)	TOTAL (USD)	Elec Cost (USD)	O&M (USD)	TOTAL (USD)
Trucking	$ -			2,282.39	3,933,269.08	3,935,551.47
Communal	$ 1,798,101.98	376,300.00	2,174,401.98	6,736.23	354,476.14	361,212.37
Communal +	$ 2,179,491.26	376,300.00	2,555,791.26	6,736.23	3,543,485.66	3,550,221.89
HouseHold	$ 6,420,123.88	376,300.00	6,796,423.88	6,736.23	338,934.37	345,670.61

(UNHCR, 2014b, p.66)

The analysis of the consolidated costs indicated that the communal solution would be the cheapest long-term option, while the communal plus solution would result in the highest expenses when the projected net present value of the combined capital and operational costs is taken into account. In turn, the household solution would require by far the largest capital investments for the construction of a more permanent infrastructure, but its operational costs were projected to be the second lowest out of the three options. However, while financial analysis complemented the process of searching for and questioning

the viability of different options, it played a secondary part in evaluating the problems of the camp:

> Analysis strongly indicates that communal and communal+ require additional risk mitigation costs and would not be adequate based on lack of community acceptance (Zaatari Water Network Technical Working Group, 2014, p.80).

The working group thus not only concluded that "social factors" would require additional unquantifiable risk mitigation costs for the communal and communal plus solution, but also that community acceptance for these solutions would complicate the operationalization of these options. Therefore, the only immediately viable option was a radical reform of the trucking option, which is explored in the next section.

At this point several issues can be summed up, which share important similarities and differences with the previous case study on the organization of nutritional programmes in the camp. Firstly, Sphere played an important role in the process of conceptualizing the camp as an integrated social space that cannot be managed by merely looking at the sum of its parts, namely its individual agencies and challenges. Translating and visualizing the range of Sphere's evaluation dimensions into the matrix in the sector working groups helped the managers to move away from an agency level evaluation to a more holistic approach. In this context, technical and social evaluative dimensions were not treated as separate but as mutually constitutive, thus rendering an artificial dichotomy between value and values irrelevant (Stark, 2009; Kornberger et al., 2015). Secondly, similar to the case study on nutritional

programmes, engaging with Sphere's evaluative dimensions stimulated attention to tensions and trade-offs, challenging assumptions and expectations underlying the management of the response operation to this point. This practice was of particular relevance for the managers in the course of interrogating and creating future solutions for the camp. To further explore the evaluative tensions, the managers employed a scoring mechanism. However, neither the evaluative dimensions, nor the scoring mechanism was taken at face value. Instead they were used to stimulate questioning, expose blind spots and examine new possibilities, thus not only representing what was known but transforming what was knowable (see March, 1987; Weick and Sutcliffe, 2007; Stark, 2009). This process was influential and informative precisely because it drew attention *beyond* what could be categorized and represented through the evaluation system (Quattrone, 2009; Jordan et al., 2016). Thirdly, Sphere provided a shared vocabulary for communication and negotiation between the highly heterogeneous organizations and actors in the camp. While Sphere's prescriptions were interpreted in a flexible manner, they provided an influential template for recording observations and were a guide for problem solving. However, as demonstrated in the following section, while these elements were insightful for Zaatari's managers when exploring options and alternatives, further mechanisms were required to address the problems of the camp.

5.4.3. *Participatory Evaluation as a Mechanism for Exploration and Control*

In the previous case study, participatory evaluation played the important role of providing an additional mechanism for exploring solutions and making the

most of the distributed intelligence throughout the camp. In this case, as Zaatari's managers faced the question of how they might construct a governance system that controls unprofessional and corrupt agencies in the camp, it led to several more radical propositions:

> So we considered to have them [the refugees] actually monitoring and measuring the quantity and quality of water getting into each district. So number one was to recognize that they were at least partially right. The moment we recognized that the refugees were partially right it all changed. This might not be exactly what Sphere says in terms of participation, but this what we moved towards (senior manager 1).

> So for instance what does the right to participate mean? It means you have got to meaningfully involve affected populations in designing humanitarian response, which is how Sphere defines it. Fine, fine, I can act on that (field officer 3).

> So now we have regular formal and informal consultations with refugees. The formal consultations take the form of regular meetings with the refugee representatives, whereby we consult them on various issues, from the quality of the services in the camp to future plans. We can involve them in the planning, distributions for example. I mean, of course, the refugee participation in the decisions that affect their life is an important priority of Sphere. This is the only way the assistance in made sustainable. A refugee camp is imagining a new community settling down. If you apply all the standards in an environment that itself is not conducive you will not be necessarily meeting the needs of the people. So it's very important that refugees are involved in the planning, in choosing types of assistance so it can be improved (senior camp manager 3).

In line with these insights, at the centre of the new strategy stood a fundamentally revised form of governance and ownership for water

management that draws from several key elements of Sphere's philosophy and methodological approach. In the governance of water related issues, an active shift away from the humanitarian agencies was encouraged. Some of these elements were inscribed into a novel framework of water governance, which emerged from the negotiations of the technical working group. The first section of the new water framework, 'governance and ownership', gives concrete meaning to Sphere's process standard of a people centred humanitarian response for the context of Zaatari:

> Users actively participate in the development of new facilities and WASH interventions and in the management of existing facilities (Zaatari water framework, 2014, p.1).

> To determine the performance of the WASH committees, WASH blocks will be monitored at least twice every week by the lead agency… Meetings are to be held with 100% of WASH committees scored as underperforming within 2 weeks of the assessment and an action plan for improvement agreed…WASH Partners monitor and evaluate performance of WASH Committees and provide support to them as necessary (Zaatari water framework, 2014, p.2).

> Communities are informed of ways they can report their WASH committee members with complaints and problems, making the committees accountable to their communities through activities such as community mobilisation, flyers, FGDs, surveys etc… (Zaatari water framework, 2014, p.2).

The governance framework spreads out the tasks of monitoring, evaluating and controlling performance criteria for water management away from agencies towards Zaatari's community. As follows from the excerpts and quotes above, the framework envisions: (1) mechanisms for the refugee

community to hold humanitarian agencies to account; (2) mechanisms through which communities can hold their own representatives to account; (3) mechanisms through which the agencies can hold the communities to account; and (4) mechanisms through which agencies can hold other agencies to account. The governance framework thus rearranges the previously hierarchical mode of managing water and provides a mechanism through which the refugee community would be included for the first time in the evaluation and decision-making processes. This development was not trivial. Up to this point, the refugee community had been treated as passive aid recipients. The quotes below describe some of the initial developments following the gradual operationalization of this approach:

> Now we have the district level monitoring, then the street level monitoring. The street level monitoring works like this. We bring the water in from the entry level to the street level of the camp that is the responsibility of our partners, [Organization X and Organization Y]. From the street level onwards there is a group of people from the community, the water fillers, who are actually refugees. So these people will now physically follow the trucks, taking the right amount of water from the tanks to fulfil the demand in their places. So they are monitoring the main water supply and they are the people that raise red flags when there are issues with water (senior camp manager 2).

> The important thing is that there is a real-time monitoring and feedback process. So the refugees do not have to wait until the end of the month to take action. So if they find a problem immediate action can be taken (senior camp manager 3).

> I mentioned that the responsibility of [organization X] and [organization y] is to run the WASH committees. We are having community gatherings in every district of the camp, almost on a day-to-day basis, discussions with them (evaluation officer 3).

Figure 5 shows a copy of a performance sheet that each of the water fillers had to complete every time a water delivery took place. The performance sheet is organized around a variety of relatively simple metrics, including water quantity and quality, timeliness of delivery, location of delivery and identification of the delivery truck. Despite its simplicity, it provides a direct channel for refugees not only to engage in monitoring and evaluation activities, but also as a means of resistance since refugees can reject the delivered water and raise direct concerns. Sharing the sheet with the responsible WASH committee also provides a mechanism for additional feedback loops, in which the refugees are no longer directly dependent on the water providers but can go through established procedures for mediating emergent concerns. In addition, the WASH committee provide an open forum for the agencies to hold each other accountable. If an agency claims that their performance evaluation reports show no issues, refugee participants can provide contrary evidence if necessary. This process enables other agencies to follow up and increase pressure on underperforming or corrupt agencies in a timely manner.

Figure 5: Performance Sheet for Zaatari's Street Level Evaluation Procedures

While the monitoring and evaluation practices were partly shifted away from agencies, having refugees participate in the water governance was significantly more than a technical exercise. One of the fundamental reasons for the creation of Sphere, apart from providing a tool for reflective evaluation in humanitarian settings, is the advancement of the rights of people affected by disasters (Buchanan Smith, 2003; Barnett, 2011). As emphasized by Stockton (2005, p.5), who was a key leader in the development of Sphere, the principle of participation is one of its cornerstones:

> There can be no exclusions from this principle [of participation and informed consent] without simultaneously reducing the affected person to sub-human status (Stockton, 2005, p.5).

The question of refugee participation is not only fundamental but also complex due to two interrelated reasons. Firstly, there are no well-established community structures in refugee camps. A sense of community has to be carefully developed over time. Secondly, while refugees are recipients of relief aid, these services are delivered in the absence of well-established accountability structures. Refugees commonly do not have the same status as citizens, who might exercise judgement on the quality of service provision on a regular basis by exercising their right to vote. Against the background of these issues, the notion of refugee representation became a complex and central concern:

> I mentioned earlier that we had these street leaders in place who were in charge of water. And these people they are influential, they are the

kind of Mafia of water. We had incidents where they were selling water, where they were punishing people because they have all the power where the water goes etc… (senior camp manager 2).

[Trustworthiness] is a big concern and challenge. Initially we had what we called the street leaders. These were very influential people. But we then doubted how trustworthy they were. At the same time, we needed to keep these people (field manager 2).

We would invite the street leaders and have a discussion with them. These street leaders, of course they are self-appointed. After we engaged in discussion with them we realized that it's probably not a very good idea to keep in direct contact with them because it was empowering them and there is a risk that they will bribe for information, act as information gatekeepers (field manager 2).

As became evident, the majority of refugees were not only vulnerable to malpractice from humanitarian agencies that were underperforming or corrupt, but also to abusive power structures emerging from within Zaatari's community. In this context, the management of water was recognized as a highly political problem for the camp. Access to and control over water gives specific groups in the camp influence over others. This issue was of particular importance given that Zaatari had experienced significant problems with mafia-like groups that had seized aid items and converted them into power and influence to the detriment of the vulnerable, handicapped and elderly.

Pursuing New Performance Categories: Adaptability, Relational Value and Imperfection

The above sections have highlighted how the experimentation with Sphere facilitated the initiation of a process of moving away from the failures of a top-down, linear command approach to governing water in Zaatari. Despite the emerging challenges, the reorganization of water governance constituted an important advance, making it significantly more difficult for corrupt agencies or individuals to cheat. A strong implication from the findings reported is that in the evaluation of performance during humanitarian response operations, important insights and changes were produced by techniques that encourage continuous exploration, doubt and questioning instead of exact repetition. Practices and techniques that favoured such exploration, including the interrogation of evaluative tensions and the notion of participatory evaluation, became highly influential in this context.

However, to address the challenges that emerged through this process, the existing performance requirements had to be further adjusted to the changing conditions. As Weick and Sutciffe (2007:67) emphasize, performing reliably in conditions of high uncertainty and complexity require the capacity "to *sense* the unexpected in a *stable* manner and yet *deal with* the unexpected in a *variable* manner" [italics from original]. While performance categories are unavoidable to make sense of disaster contexts, they can also draw attention away from important elements that they do not capture and by locking in expectations that are inappropriate for the specific disaster context. Thus, to prevent managers from taking Sphere's performance evaluation categories at face

value, specific mechanisms are needed to counteract oversimplification and encourage flexibility, adaptability and judgement. This requirement of working with performance categories and at the same time being sceptical of them is a crucial but demanding task for humanitarians. This section focuses on techniques and practices that enable humanitarian response managers to adapt performance categories, assumptions and actions. The following quotes highlight several aspects against this background:

> Let me give you an example on how flexibility matters. As I speak now the government is giving us something like a cap, the amount of water we can provide in Zaatari on a daily basis. And this amount slightly falls below the actual needs on the ground at the moment as per the assessment of our partners…This changes the technical indicators and with that the entire puzzle of other issues. As we modify our response, the measures have to change. We cannot treat them as if they were set in stone. Sphere gives us a base in relation to what we have to meet and connect, but also modify. It guides us in this process of negotiating with partners but also with ourselves. There is no one solution for this (senior camp manager 2).

> The path of a humanitarian operation is often chaotic and there is very imperfect information about what's going on. So major decisions have to be made on the basis of judgement in very uncertain conditions. The problems we faced around water are no different in this regard. And those are decisions about programme focus in terms of location, in terms of content, and also in terms of what sort of funding to go for. That means that major amounts of money are being allocated on the basis of judgement rather than hard data. Any ongoing work on performance management fits into that. We didn't use Sphere as a hard edged set of tools against which we measure performance (field officer 3).

Even though significant advances in the reorganizations of the camp had been made, urgent issues still lined up on the agendas of Zaatari's managers. These issues included addressing the abusive power structures emerging from within Zaatari's community; the protection of vulnerable groups; adjusting the technical indicators to fit the contextual requirements, taking into account the limits on water provision demanded by the government; and strengthening the refugee's capacity for self-organization. As stressed in the above quotes, a change in one of these issues entails a shift in the entire puzzle of other issues, which therefore required great care in adapting response requirements. The following quotes highlight several elements of the process:

> We know that if we supply X-amount of water per day as per Sphere we are fulfilling an important part of our job. But to ensure equity, to ensure that the weaker people in those communities are also receiving the amount of water that they need, we had to put many, many mechanisms in place. Our community outreach mechanisms provided the raw data from which we could then continue working and modifying our performance frameworks (senior camp manager 2).

> And then working really closely with the people we found out that the refugees had really clever ideas of how to manage and measure the water and how to distribute it more equally. We came up with different quotas for each street, based on the number of families and children, disabled and elderly. We, with our sort of stupid humanitarian thinking, we say that everybody has to have the same amount. We came up with a more transparent process of dividing private vs. public tanks. To accept these flexibilities is very hard for humanitarian agencies. So then we worked towards further changing the system (senior camp manager 1).

As follows from the quotes, the process of engaging with Sphere's measures involved as much a process of discovering possible objectives as acting on

them (March, 1971; Cooper et al., 1981). By introducing new quotas for each street, the managers moved away from a one-size-fits-all approach to measuring water delivery, thus adapting Sphere's technical indicators according to contextual factors. This was a crucial element to balance the limits on delivery imposed by the government with issues of equitable distribution, protection and health. However, as described above, the interconnected nature of Sphere's evaluative dimensions encouraged managers to pay attention to the links and tensions between these elements, instead of focusing narrowly on isolated concerns (Stark, 2009). This approach was particularly crucial to breaking down the problem of equitable distribution into more specific and relatable parameters. Against the background of the abusive power structures emerging from within Zaatari's community, the following adaptations emerged:

> So to address the issue of abusive structures in the camp we had to modify the participatory mechanisms we initiated. We created what we called a 'rotational map.' So within each block receiving water, they had to meet and agree on twelve members who would monitor water supply. Among those members, we would have six members that would monitor water supply and six members that would stand by in this process. And this is done on a rotational basis to ensure that every household that wants to be represented in this can be represented and also to ensure that people do not dominate the water. So whatever you do wrong, you can do it wrong only on a fortnightly basis. And in addition we recruited an army of 27 staff whom we call 'water monitors.' They are field assistance. So they also form part of a monitoring system that goes to the community to check on a daily basis, doing spot checks to verify that the amount of water written on the board is provided to this amount of people and that everybody who is supposed to receive water is receiving it. These adjustments were extremely meaningful for our ability to improve water governance (senior camp manager 2).

As stressed above, the continuous engagement with Sphere's openly imperfect system provided avenues for the involvement of new contributors, namely different refugee groups, who in turn added their own knowledge, interests and goals (see Garud et al., 2008; March, 1971; Stark, 2009). Performance evaluation practices were thus conceived as an ongoing, collaborative and emerging process that played a central role in the reorganization of the camp's water governance and the discovery of new possible solutions to its problems. As emerges from the evidence presented, the adaptations strengthened the refugees' capacity for self-organization, which at the same time constituted a crucial requirement in developing an enhanced ability to control corrupt agencies and ensure a more equitable and safe distribution of water as a crucial resource. This process, far from being predetermined, emerged through engagement with particular techniques and practices of Sphere's evaluation system. While solutions in the management of humanitarian crises are always temporary, the above factors helped to address some of the key issues the camp had experienced.

This section's analysis was mainly concerned with techniques and practices that enable humanitarian response managers to adapt performance categories, assumptions and actions. In line with these concerns, the above analysis indicates four key elements of how engaging with performance evaluation categories in this way might be possible and informative. Firstly, put provocatively, while 'complete' performance categories normalize and draw attention to the familiar, 'incomplete' categories anomalize and draw attention to the unusual, the odd and the unexpected. Given that 'incomplete' categories can never be fully trusted, they can serve the purpose of expanding the attention of the user towards signals that might indicate trouble or a need for

adaptation. As stressed in the review of the literature, while stimulating questioning and reflection is not commonly a virtue associated with accounting systems, for the management of humanitarian crises it is of critical importance to embrace systems that routinize the exercise of questioning, doubt and imagination (see March, 1987; Quattrone, 2015b; Revellino and Mouritsen, 2015). Secondly, as follows from the above quotes, the incomplete and interconnected nature of Sphere's evaluative dimensions encouraged managers to pay attention to the links between these elements, helping them to create an integrated map of operational issues instead of focusing narrowly on isolated concerns. Thirdly, incompleteness can serve as an incentive to create ideas on how design can be developed and further improved. In this sense, performance evaluation practices in Zaatari involved as much a process of discovering possible objectives as acting on them (see March, 1971; Cooper et al., 1981). Fourthly, the continuous engagement with Sphere's openly incomplete system also provided avenues for the involvement of new contributors, who in turn added their own interests, goals and value criteria. Incompleteness thus facilitates the conceptualization of performance evaluation design as an ongoing, collaborative and emerging process (see Garud et al., 2008).

5.5. <u>Concluding Remarks</u>

This chapter has outlined three case studies focusing on tactics and principles that enable humanitarian performance evaluation systems to facilitate engagement with the unpredictability and complexity of humanitarian crises. In so doing, it has explored the question of how accounting systems can

embrace the ambiguity of highly complex settings and make a resource of it (Cooper et al., 1981; March, 1987; Weick & Sutcliffe, 2015). To conceptualize how ambiguity may become a resource for accounting systems, in the previous chapters the notion of heterarchy (Stark, 2009; Lamont, 2012), defined as 'governance through difference,' was mobilized as an insightful theoretical approach that stimulates criticism, fosters checks and balances, and recursively creates spaces for reflection. Furthermore, it was stressed that theoretical understanding in relation to elements that may work in favour of, and against, the notion of heterarchy is underdeveloped (see Lamont, 2012, Holm, 2010). From the empirical material several key concerns emerged that enrich our understanding in relation to the question of how the notion of heterarchy may inform accounting systems. These key concerns are briefly summarized below.

The first case study has explored specific challenges and conceptual responses in relation to the design of Sphere as the most widely used performance evaluation system for the management of humanitarian crises. The analysis of the empirical material indicated that in the design of Sphere, several key elements stood out that set it apart as a distinctive system for humanitarian performance evaluation. These elements included the conceptualization of multiple, inseparable and interconnected modes of evaluation, measures that openly embrace incompleteness and a focus on accounting for gaps, and an emphasis on a more flexible and adaptive form of control instead of advocating the production of de-localized knowledge in a remote centre of calculation.

Building on these findings, case study two and three explored how Sphere's evaluation system shaped humanitarian response management in a large-scale refugee camp, focusing on the operational complexities of organizing nutritional aid programmes and the governance of water respectively. As the prescriptions of the evaluation system were translated into the refugee camp setting, several key elements emerged through the interaction with Sphere, highlighting how it asserts influence in the management of humanitarian crises. These elements included the generative role of evaluative tensions in interrogating the unexpected and in exploring future solutions for the camp, the organized openness of the evaluation system that was stimulated through participatory mechanisms that fostered both inclusion and control, and the role of incomplete templates for evaluation in adapting to emerging challenges and exploiting contextual opportunities. While these elements unfolded in a distinct manner in each of the two case studies on nutrition and water, they emerged strongly as key tactics and techniques that enabled engagement with the humanitarian crisis setting. Drawing from the empirical insights of this chapter, the following discussion chapter connects these concerns with the relevant literature and develops an argument for how heterarchical principles of performance evaluation may be theorized for engagement with highly unstable settings such as humanitarian crises.

6. Discussion

6.1. <u>Introduction</u>

Having detailed the empirical findings in the previous chapters, this chapter turns to the theoretical implications and the study's contributions. One of the key guiding questions of this research that emerged from the literature review was how accounting systems might be designed and practiced to facilitate engagement with anomalies, inconsistencies and the unexpected in the management of humanitarian crises. To serve as such 'anomalizing accounts,' it was argued that it is crucial to takes seriously the limitations and incompleteness of accounting systems in such contexts and attempt to build on them. These guiding concerns framed the study's theoretical focus on exploring how humanitarian performance evaluation systems and practices enact notions of ambiguity (March, 1987; Cooper et al, 1981; Mouritsen, 2016), incompleteness of accounting information (Jordan and Messner, 2012; Dambrin and Robson, 2011) and a multiplicity of distinct modes of evaluation (Chenhall et al, 2013; Stark, 2009; Coslor, 2016). One of the key analytical concepts that was mobilized for the combined exploration of these issues was the notion of heterarchy, which is further developed in this chapter (Stark, 2009; Lamont, 2012).

What emerged strongly from the findings chapters is that the interaction with particular techniques and principles of Sphere's performance evaluation system shaped the engagement with unexpected and emergent issues as well as the complexity of the camp. To explain how Sphere becomes a performable technique attuned to the highly dynamic humanitarian environments, in this chapter the study theorizes four interlinked procedural principles that emerge from the case studies, which further advance the notion of heterarchy (Stark, 2009; Lamont, 2012). These principles include: (1) In-built tensions between evaluation dimensions; (2) Open and participatory design; (3) Relational value and incompleteness; and (4) Enacting minimalist control through a community of practitioners. In this chapter, it is argued that the theorization of these heterarchical principles contribute to scholarly understanding of *how* accounting systems can foster techniques that make ambiguity a resource and enhance their evocative power (Cooper et al., 1981; March, 1987; Quattrone, 2015a; Revellino and Mouritsen, 2015) to engage with complex settings such as humanitarian crises.

To tackle these issues, the chapter adopts the following structure. Section 6.2 outlines and discusses four heterarchical principles for performance evaluation that emerged from the field research and relates these to the accounting literature. To present these principles the section is divided into several subsections, ranging from 6.2.1 to 6.2.4, each of which relate to a specific heterarchical principle for performance evaluation. Following the discussion of the individual principles, sub-section 6.2.5 gives an overview and further discusses the interaction between these heterarchical principles for performance evaluation. The chapter ends with concluding remarks in section 6.3.

6.2. Making Ambiguity a Resource: Heterarchical Principles for Performance Evaluation

6.2.1. *In-Built Tensions between Evaluation Dimensions*

Attention to the potentially productive role of tensions between divergent principles for evaluation constitutes one important part of the recent interest in the practice of heterarchies of value (Stark, 2009; Lamont, 2012). However, as was also discussed in section 3.4, scholarly understanding remains underdeveloped in relation to the conditions, principles and technologies that might work in favour of, or in contradiction to, the practice of heterarchies (see Lamont, 2012). While it is not surprising to find multiple notions of value in organizations and society (see Annisette and Richardson, 2011; Chenhall et al., 2013; Gehman et al., 2013), it is less clear how heterarchical tensions might be productive. As emphasised by Holms (2010), a key criticism in relation to the theoretical understanding of the notion of heterarchy is that its practice might contribute to chaos and facilitate strife. A limited number of studies have begun to explore the role of tensions in accounting and governance. For example, Quattrone (2015a) highlights how the material solution of a Jesuit cash box served as a mediating technology between economic and religious concerns. However, despite these recent advances, understanding remains limited in relation to the question of how coordination is achieved despite the dissonance that constitutes a defining feature of heterarchies (see Lamont, 2012).

This study's findings suggest several insights to further develop the notion of heterarchy for accounting research and beyond (Stark, 2009; Lamont, 2012). Instead of fostering static compromise (Chenhall et al. 2013), the tensions between Sphere's evaluative dimensions became productive for Zaatari's managers because they allowed for the challenging of expectations in relation to the response operation and facilitated the interrogation of future possibilities. In the case study on water, Zaatari's managers used Sphere's evaluation criteria to build a matrix that integrates overall response requirements for the entire camp, moving the performance evaluation activities away from an individual agency focus to a more holistic approach. In this process, the tensions and trade-offs between Sphere's evaluative dimensions emerged as a powerful mechanism to explore benefits as well as shortcomings in relation to several future possibilities and options for the camp. The interaction with Sphere's evaluative dimensions was thus less about ensuring co-existence between evaluative dimensions (Chenhall, et al., 2013), but instead about fostering an attitude of questioning and exposing blind spots (Cooper et al., 1981; March, 1987). This emphasis on tensions and 'anomalizing' is important as it contributes to scholarly understanding of *how* accounting systems might embrace and foster techniques within their designs that draw attention beyond the limitations of their prescriptions (Jordan et al., 2016; Revellino and Mouritsen, 2015; Quattrone, 2015a). As has been noted in the findings, while compromise normalizes and stabilizes, focusing on heterarchical tensions encourages a focus on anomalies and knowledge gaps.

The focus on tensions, interrogation and anomalizing is highlighted in figure 5, which visualizes the process-oriented conceptualization of humanitarian value that emerged through the engagement with Sphere. As indicated by the

question marks in figure 5, the relationship between the different evaluation dimensions is always uncertain and might frequently lead to tensions between them as users are induced to question and balance their competing criteria.

Figure 6: Tensions, Interrogation and Sphere's Procedural Approach to Humanitarian Performance Evaluation
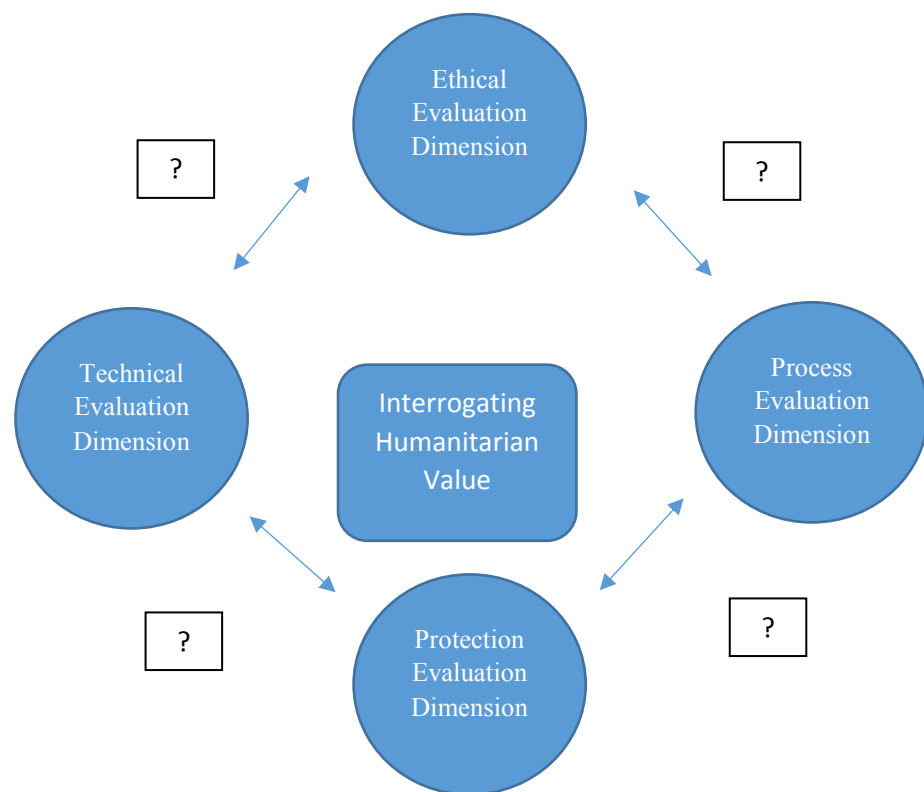
Another key factor in relation to the notion of heterarchy that emerged from the research was that tensions between evaluation principles were not productive per se, nor was any type of tension equally valuable. In Stark's (2009) ethnographies tensions form a productive element in all of the organizations that were studied and apply equally to each of the described organizational problems. In other words, apart from the qualification that

rivalry between evaluative frameworks must be principled, no conceptual distinction is made between tensions that are productive and other types of tensions that simply imply chaos and confusion. This study's research suggests a narrower scope. In the process of engaging with Sphere, tensions between its evaluative principles became productive at a specific point in time, namely as Zaatari's managers were exploring future options for the camp. This highlights the significance of notions of timing and temporality, which have so far not featured in the scholarly work on heterarchies or in the literature that explores the enabling potential of deliberately ambiguous accounting systems (see Cooper et al., 1981; March, 1987; Quattrone, 2015a; Revellino and Mouritsen, 2015). Notions of timing and temporality are particularly important in relation to the concern that the practice of heterarchies might result in chaos and strife (Holms, 2010; Lamont, 2012). The case studies reported in this thesis reveal that it was undoubtedly a common and complicated challenge to gain an appropriate sense of timing in the attempt to reorganize the response operation. However, instead of producing chaos, exploring tensions between Sphere's evaluative tensions attracted users because of its role in challenging taken-for-granted assumptions and evoking additional meanings beyond what was knowable at this point (March, 1987; Jordan et al., 2016), thus sharpening the manager's senses towards possibilities and problems. These insights suggest that there might be different temporal phases that influence whether heterarchical principles are productive or destructive.

In line with these insights, the findings suggest a further factor that was influential in making the evaluative tensions productive (Stark, 2009; Lamont, 2012). Unlike in Stark's (2009) ethnographies, tensions did not emerge due to

the complexity of the setting, but developed through the interaction with Sphere's format and approach, highlighting the importance of the materiality of the Sphere Handbook. In the case study on nutrition, the material aspects of the Handbook were frequently highlighted as an important factor to guide discussions, produce visualizations and connect performance requirements. Similarly, as Zaatari's managers sought to build a matrix to integrate overall performance requirements in the camp, the Handbook provided an influential template for problem solving and to score the different evaluation criteria within different scenarios, highlighting the importance of stimulating "interpretative flexibility" (Qu and Cooper, 2011, p.347). This indicates that tensions between evaluative principles were fostered in an organized and strategic manner through the interaction with the Sphere Handbook. Attention to the materiality of the Sphere Handbook in stimulating heterarchical tensions thus highlights the importance of the format and design of performance evaluation systems and the 'silent' epistemology that is embedded in them (see Pollock and D'Adderio, 2012; Quattrone, 2000). As indicated in the following section however, fostering and harvesting tensions between evaluative criteria is only one, albeit important, element of the theorization of heterarchical principles for performance evaluation.

6.2.2. *Open and Participatory Design*

Designing for complexity and uncertainty entails recognizing the fundamental limitations of performance evaluation systems in anticipating the wide variety of performance requirements across distinct humanitarian crises. As the intrinsically dynamic humanitarian environment makes it impossible to pre-

specify the exact nature of problems or to identify pre-existing alternatives and solutions, humanitarians require technologies that enable them to discover and contain emergent issues before they become unsurmountable (see Weick and Sutcliffe, 2015). Cooper et al., (1981:187) precisely articulate the ensuing theoretical challenge for accounting systems to operate in such uncertain and anarchistic environments: "without a pre-existing goal structure, can there be prescription and design?" The previous section began to outline how the heterarchical principle of in-built tensions between evaluative principles might provide several insights in relation to these theoretical challenges. In this section, another factor in the development of heterarchical principles for performance evaluation is developed, namely the notion of 'open and participatory design'.

As pointed out by Stark (2009:16), heterarchies seek to make "assets of ambiguity by keeping open multiple ways of redefining, and hence recombining and redeploying, resources." However, the question of how this openness might be organized and unfold is less clear. The notion of 'open and participatory design' highlights how such openness might be practiced as part of performance evaluation systems. As emerges from the findings, the interaction with Sphere's prescription of refugee participation evolved into a powerful mechanism to move away from conceiving performance evaluation practices as a linear process in which pre-specified categories are simply assessed and ticked-off. By allowing different forms of community participation to emerge, Zaatari's managers were able to attract and gain insights from heterogeneous groups with different types of expectations, thus reducing oversimplification (Weick and Sutcliffe, 2015) and making the most of the distributed intelligence throughout the camp (Stark, 2009). This process

contributed significantly to the revision of prescriptions and plans about how nutritional aid and water governance were organized. This was of particular importance since, as was emphasized in the findings, many crucial pieces of information were not to be found in the formal information systems due to the messiness of the situation on the ground, which constitutes a common element in humanitarian crisis management (Ramalingam, 2013; Sargiacomo, 2015).

Moreover, the notion of open and participatory design was in many instances a valuable instrument for community building. As highlighted in the findings, by engaging weaker, isolated and vulnerable groups in the camp, the participatory mechanisms that were set up through the interaction with Sphere provided a powerful tool to build and increase trust between refugees and humanitarians. The building of trust was vital to gaining additional sources of information and to providing regular channels of communication for the refugees to be able to hold agencies accountable, thus making it more difficult for unprofessional and corrupt agencies to hide from criticism. This process helped uncover mafia-like structures within the refugee community and started discussions on the allocation of tasks through which the refugees would take a more proactive role in many of the decisions and practices affecting their day-to-day lives. While notions of community building may appear somewhat distant from more conventional accounting practices, historically they constituted a central element of the accounting craft (see Puyou and Quattrone, 2016), and appear to be of particular importance in situations of crisis. In this sense, conducting performance evaluation as open and participatory practice served the purpose of giving back to refugees a sense of dignity, facilitating the re-definition of the ties and relationships between Zaatari's community, or in other words its 'socie-ties' (Latour, 2005).

To summarize, while Stark (2009) highlights the importance of system openness as an important element for heterarchies, scholarly understanding of how such openness might be theorized and enacted is underdeveloped and poses a significant challenge for the design and practice of accounting systems (Cooper et al., 1981; Cooper, 1983; March, 1987). In particular, this study provides several insights into how such system openness might be conceptualized as part of the development of heterarchical principles for performance evaluation.

Firstly, the principle of 'open and participatory design' shifted the focus of the managers towards processes that allowed them to question what was unknown and unknowable beyond the limitations of the accounting system (Jordan et al., 2016; Weick and Sutcliffe, 2015; Quattrone, 2015a). The study hereby highlights *how* the openness of Sphere's evaluation system was guided in an organized manner, which allowed for the repeated questioning and criticism of some of its own prescriptions. An important example in this regard was how the participatory mechanisms led to a challenging and re-thinking of the initial techno-financial framing of the organization of nutrition towards a more inclusive approach that was more respectful of notions such as human dignity.

Secondly, the notion of 'open and participatory design' was powerful to further include vulnerable groups in the camp. In doing so, it gave concrete meaning to the ethical principle that disaster-affected people and refugees should not be treated as passive recipients with no agency of their own. In this

sense, engagement with Sphere contributed to the inclusion of marginalized groups and facilitated a conceptualization of performance evaluation as a collaborative and emerging process. Against this background, by developing the principle of 'open and participatory design', this study further contributes to enriching the concept of heterarchy (Stark, 2009; Lamont, 2012) and how its emphasis on 'governance through difference' can be conceptualized as part of performance evaluation systems.

6.2.3. *Relational Value and Incompleteness*

The previous two sections addressed questions on how heterarchical principles for performance evaluation can facilitate the interrogation of humanitarian value and the challenging of expectations and deliverables in response operations. Both of these principles provide important insights into how humanitarians might engage with uncertainty and the unexpected in crisis situations, providing an element of stability in the face of chaos. However, dealing with the rapidly unfolding challenges and emerging problems of humanitarian crises also requires high degrees of flexibility and adaptability, which cannot be addressed by statically relying on pre-specified performance indicators (March, 1981; Weick and Sutcliffe, 2015). Following Stark (2009:4), such adaptive processes, which constitute an important element of the theoretical concerns surrounding heterarchical systems (Lamont, 2012), involve the highly challenging task of recognizing and re-assembling "what is not yet formulated as a category." In line with this issue, another central challenge for humanitarian performance evaluation systems is not only to serve as a technique to interrogate and embrace blind spots, but also to

proactively encourage adaptation to unplanned challenges. This section focuses on theorizing how heterarchical principles for performance evaluation can inform such adaptive processes. For this purpose, the principle of 'relational value and incompleteness' is developed.

In the process of dealing with and recovering from unexpected challenges, two interrelated themes in relation to the interaction with Sphere's evaluation system stood out from the findings. A first theme was that, in line with the notion of heterarchy, the interconnected nature of Sphere's evaluative dimensions encouraged managers to pay attention to the links between these elements, helping them to conceptualize operational issues in an integrated manner in which a shift in one of the dimensions affects all others. Since Sphere's evaluative dimensions are not conceptualized as separate but instead as mutually interdependent, improvisation, tinkering and a search for novel associations based on contextual requirements was actively encouraged. This insight draws attention to the specific manner in which accounting systems can foster interpretative flexibility (Cooper et al., 1981; Qu and Cooper, 2011; Jordan et al., 2016), without losing coherence. The practical effort of improvising encouraged reflection about the way Sphere could be met for the requirements of the camp, working against an overly strict and dogmatic manner of engaging with its metrics. Thus, while Sphere's evaluation dimensions offered a template for engaging with the complexities of humanitarian response operation, these templates are designed to be continuously adapted. In other words, instead of providing a ready-made package of pre-established options, Sphere can offer its users a way of re-configuring what possible options could exist in specific disaster contexts (Stark, 2009).

Such a recombination of existing resources into a novel solution was particularly evident through the development of Zaatari's supermarkets. The ensuing reorganization of the nutritional response operation connected several emerging concerns by not only providing refugees with an increased sense of choice and dignity, but also in cutting operational costs and reducing violence and demonstrations surrounding the delivery of food aid. The contextualization and association between Sphere's ethical and technical evaluation requirements thereby resulted in a novel solution for the camp, i.e. the reorganization of the nutritional response operation. Figure 6 visualizes this idea in a simplified manner.

Figure 7: Recombining Incomplete and Interdependent Evaluation Dimensions
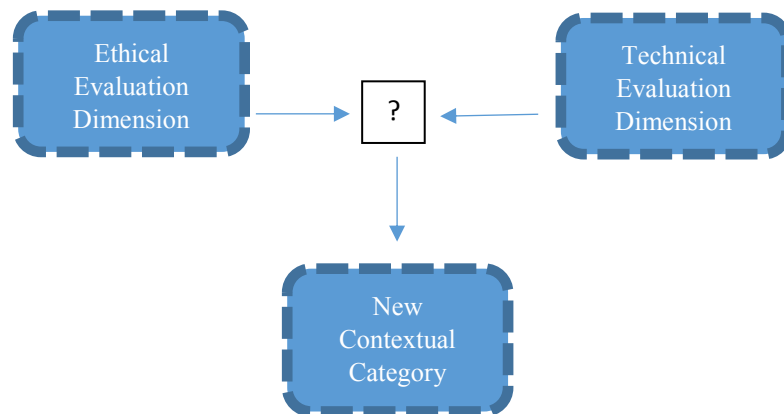
However, the relational nature of Sphere's evaluation dimensions was not the only element that contributed to the adaptability of Sphere's evaluation system. Another important factor in making Sphere a performable technique

in the management of humanitarian response operations was that its users were actively aware of its imperfect and incomplete nature and were therefore less susceptible to being blinded by the limitations of its prescriptions. While several recent studies have stressed how the incompleteness and imperfection of performance measures might not necessarily be problematic (see Andon et al., 2007; Dambrin and Robson, 2011; Jordan and Messner, 2012), the idea of proactively fostering incompleteness as a resource for accounting systems is less well developed (see Cooper et al, 1981; March, 1987; Quattrone, 2015a).

In line with this recent interest, several ways in which incomplete performance measures might be informative emerged from the findings, which add to and differ in some respects from the above studies. Given that Sphere's openly 'imperfect' categories could never be fully trusted, users employed them as a means to investigate knowledge gaps and to focus on the detection of signals indicating a need for adaptation. Instead of ignoring the gaps that were created by translating performance data into indicators (Dambrin and Robson, 2011), Sphere's incomplete metrics drew attention to, and problematized, such information gaps and inconsistencies. Using performance measures and working with their open and evident incompleteness in this manner served the purpose of questioning the expectations embedded in these categories and comparing these with observations made. Thus, performance measures can be carried in a more flexible manner, which implies a need to "both believe and to doubt [them]" (Weick and Sutcliffe, 2007, p.58). In this sense, as suggested by figure 6 above, working with openly incomplete performance measures can serve as an incentive to question whether contextual discrepancies suggest the emergence of novel adapted performance categories (Stark, 2009). In the context of engaging with the complexity of Zaatari refugee camp, using

Sphere's metrics as means rather than ultimate objectives involved as much a process of discovering possible goals and adjustments as acting on them (see March, 1981; Jordan and Messner, 2012). In line with these insights, incompleteness of evaluation principles was not a factor to be eliminated, but a precondition for contextualization, adaptability and recombination.

6.2.4. *Enacting Minimalist Control through a Community of Practitioners*

Another principle emerging from the findings relating to the engagement with Sphere's evaluation system was that it encouraged a minimalist approach to control (Quattrone and Hopper, 2005; Stark, 2009; Law, 1997). It is now well established in the literature how accounting systems are integral to asserting action at a distance (Robson, 1992) and shaping the calculable space through which populations can be subjected to control and surveillance (Miller and O'Leary, 1987; Miller and Power, 2013; Sargiacomo, 2015). However, less is known in relation to how such notions might be implicated in stigmatizing and marginalizing specific groups in organizations and society, including for example the poor or disaster affected people, e.g. refugees (Walker, 2008, 2016). Furthermore, as noted by Quattrone and Hopper (2005: 761):

> Isolating individuals in calculable spaces and emphasising differentiation – characteristic of modern times – does not necessarily produce unitary, homogeneous, and totalitarian control…This reveals shortcomings of studies within large organisations in terms of a centralisation vs. decentralisation dilemma or action at a distance.

As emerged from the case studies, due to the complexity of Zaatari's setting, in which no unequivocal command structure between agencies existed and no actor had a permanent state of power over the others, notions of control were always emerging and changing. Therefore, in line with Quattrone and Hopper's (2005) quote, instead of focusing on the analysis of how centres of calculation emerge and what they do, it is more promising to focus on exploring how Sphere implicated in shaping control practices within the camp (see also Dechow and Mouritsen, 2005; Quattrone and Hopper, 2006; Walker, 2016).

As emerged from the findings, by 'consciously' rejecting compliance enforcement with its evaluation system, organizations can conform to Sphere without meeting some of its prescriptions, but are required to report on how and why they might have been unable to do so (Sphere, 2015). However, despite the absence of a centralized certification body that acts at a distance, Sphere asserted influence in an influential and yet flexible manner. Instead of imposing a top-down form of control, a more bottom – up approach emerged over time. For example, one of the defining elements of the problem of water governance in Zaatari was that it was impossible to predefine the exact nature of the control problem. As Zaatari's managers began to confront the issue of how to engage corrupt agencies, even in the face of compelling evidence of malpractice, it was highly challenging to enforce any form of formal or unitary control (Quattrone and Hopper, 2005; Dechow and Mouristen, 2005; Quattrone, 2015a). As a result, a working group was created to translate Sphere into a framework to monitor performance from a camp-level instead of an individual agency level perspective. This new governance framework rearranged the previously hierarchical mode of managing water in the camp

and expanded the task of controlling away from agencies towards Zaatari's community. In particular, it included mechanisms for the refugee community to control agencies and vice versa.

An equally adaptive approach to control can be witnessed in the example of the conceptualization and construction of Zaatari's supermarkets in the second case study on nutrition, which not only allowed the refugee community to regain a sense of normalcy, of choice and of dignity but also reduced overhead costs and violence in the camp. The emergence of these solutions would have been impossible if Sphere had prescribed a tightly coupled and inflexible system of control (Weick, 1976, 1988; March and Olsen, 1975).

Thus, while notions of control do not constitute the most important concern advanced by Sphere, its prescriptions provided an influential mechanism to allow for various forms of control to be negotiated based on contextual requirements, constraints and opportunities. In this sense the importance of the Sphere Handbook emerged time and again. It served as a carrier of Sphere's method and approach into specific disaster settings and thereby facilitated the development of a community of practitioners that sought to contextualize its approach to the particular environment of each humanitarian crisis. Sphere was thereby frequently referred to as a language (see Cooper et al., 1981; Boland et al., 2008) that shaped and allowed for communication between the community of practitioners. Accordingly, members of this community check and balance each other and thereby impose limits to what can reasonably be expected within a given humanitarian crisis context. Such a

minimalist approach to control recognizes the diversity of organizations, issues and contexts that define humanitarian practice and seeks to embrace the ensuing need for adaptation. Combined, these insights specifically address questions of how minimalist notions of control might be practiced as part of heterarchies, contributing to scholarly understanding of the enabling and constraining elements (Holms, 2010; Stark, 2009; Lamont, 2012) of heterarchical evaluation systems.

6.2.5. *Overview of and Interaction between Heterarchical Principles in Engagement with the Unexpected*

The previous sections have described how heterarchical principles for performance evaluation can be theorized and practiced. In doing so, the study advances knowledge in relation to the question of how performance evaluation systems can productively foster ambiguity and proactively build on it to enhance their evocative power (Cooper et al., 1981: March, 1987; Chenhall et al., 2013; Quattrone, 2009, 2015a). Furthermore, the study showed that evaluative tensions are not sufficient to practice heterarchies. By developing the four heterarchical principles – in-built tensions, open and participatory design, relational value and incompleteness, and enacting minimalist control through a community of practitioners – the study further enhances theoretical understanding about elements that enable and constrain the practice of heterarchies (Stark, 2009; Lamont, 2012). Against this background, table 3 provides a brief overview of the different roles that each of the heterarchical principles played in the management of humanitarian

crisis, highlighting how interaction with the Sphere's evaluation system facilitated engagement with the unexpected.

Table 3: Heterarchical Principles for Performance Evaluation

Exploration

In-built Tensions between Evaluation Principles

- Highlights requirement for balancing between evaluative principles.
- Tensions break reductionism (e.g. strict focus on technicalities).
- Production of visualization invites questioning and exploration.

Adaptation

Relational Value and Incompleteness

- Stimulates the production of novel, contextually adjusted value criteria.
- Incompleteness fosters tinkering, which in turn invites organizing work towards "perfection".
- Focus on recombination/seizing contextual opportunities.

Open and Participatory Design

- Conceives evaluation as an ongoing, participatory process.
- Ensures the attraction of diversity of users (e.g. affected populations).
- Procedural emphasis helps overcome design limitations.

Enacting Minimalist Control through a Community of Practitioners

- Provides a shared language and dictionary for evaluation.
- Facilitates the enactment of control in pragmatic interaction.
- Notions of control and constraints always subject to negotiation.

The table also makes a conceptual distinction in relation to two key processes performed through the principles that frequently emerged in the findings and the discussion, namely exploration and adaptation. In the evaluation practices that guided the humanitarian response operation, meaningful insights were commonly created by techniques that encouraged continuous exploration and

questioning instead of strict repetition. To detail how these techniques facilitated such processes and engaged users in an open yet organized manner, the left side of the table outlines the principles that favoured such exploration. However, to confront emergent challenges in humanitarian crisis settings, exploration and interrogation were not sufficient. Humanitarians also had to be able to quickly adjust to emergent challenges and interpret performance requirements in a flexible manner. In line with these concerns, the right side of the table gives an overview of principles and techniques that enabled response managers to adapt performance categories, assumptions and actions.

On the exploration side, the table highlights how the principle of in-built tensions helped to work against a tendency to oversimplify response requirements, encouraged continuous questioning, and stressed the importance of balancing the different requirements posed by Sphere's evaluative dimensions. Complementing this focus on exploration, the principle of open and participatory design ensures that the performance evaluation system made the most of the distributed intelligence throughout the camp, thereby overcoming the limitations of its own design in anticipating the range of unpredictable response requirements (Ramalingam, 2013; Garud et al., 2008; Weick and Sutcliffe, 2015). Furthermore, the principle ensures that refugees were not treated as passive recipients of aid with no voice in decisions affecting their lives.

On the adaptation side, the table summarizes how the principle of 'relational value and incompleteness' shapes adaptability in humanitarian response operations. The relational conceptualization of Sphere's evaluative

dimensions encouraged attention to their inter-sensitivity and interdependent nature. Moreover, due to the openly incomplete nature of Sphere's evaluative dimensions, Zaatari's managers used them in a flexible manner to investigate knowledge gaps and to focus on signals indicating a need for adaptation and contextualization (Cooper et al, 1981; March, 1987; Busco and Quattrone, 2015). Finally, the second principle in the adaptation column 'enacting minimalist control through a community of practitioners' stresses how notions of control emerged in the humanitarian response operation despite the lack of an independent body to verify compliance (Quattrone and Hopper, 2005; Dechow and Mouritsen, 2005). The principle highlights how Sphere's minimalist approach towards control contributed to the emergence of a community of practitioners that not only used it as a medium to deliberate and communicate about complex problems, but also to impose checks in relation to what could reasonably be expected within a given humanitarian crisis context.

However, while being conceptually distinct, the heterarchical principles for performance evaluation did not work in a strictly isolated or even separate manner. Instead, they frequently overlapped, complemented and interacted with each other. This is visualized through the circular arrows in the middle of table 3. For example, in the case study on water governance, the principle of 'open and participatory design' served as a pre-condition to the exploration of future possibilities and options but was subsequently complemented by the notions of 'relational value and incompleteness' as well as 'the enactment of minimalist control' in order to adapt response requirements. A similar trend can be observed in the case study on nutrition. Here the exploration of heterarchical tensions between Sphere's evaluative principles was an

important requirement in moving away from a techno-financial approach to managing nutrition and in subsequently adapting the response operation to develop a more inclusive innovation – the establishment of Zaatari's supermarkets – that was more respectful of human dignity. In line with these insights, the four heterarchical principles facilitated a recursive process of alternating between exploration of emergent and unknown issues and adaptation to contextual opportunities and challenges.

6.3. Concluding Remarks

This chapter has further developed the central themes that were outlined in the findings chapters. A specific focus was placed on how accounting systems might be designed and practiced to facilitate engagement with the unexpected in the management of humanitarian crises. In so doing, the chapter followed up on gaps in the literature that open up questions on the potentially productive role of deliberately ambiguous accounting systems to enhance their evocativeness and interpretive flexibility, which were shown to remain under-theorized (Cooper et al, 1981; March, 1987; Quattrone, 2015a, b; Revellino and Mouritsen, 2015). To further develop understandings of the notions of incompleteness and ambiguity in the accounting literature, the study mobilized the notion of heterarchy (Stark, 2009; Lamont, 2012), interpreted as 'governance through difference'.

The four heterarchical principles, which were presented in this chapter, further contribute to scholarly understanding of how heterarchies might be practiced and how they can inform accounting research. More specifically, the

four heterarchical principles for performance evaluation were classified around the conceptual pair of exploration and adaptation. The principles of 'in-built tensions' and 'open and participatory design' were particularly influential in organizing processes of exploring and interrogating the unexpected. Complementing these notions, the principles of 'relational value and incompleteness' as well as 'minimalist control enacted through a community of practitioners' were conceptualized as important elements to facilitate adaptation to emergent contextual opportunities and to constrain the flexibility of heterarchical performance evaluation techniques. While the four heterarchical principles were described as conceptually and analytically distinct, it was also highlighted that they commonly interacted with each other, thereby fostering a recursive process of exploring and adapting to emergent challenges in humanitarian response operations. In line with these insights, the chapter provided a detailed theorization of *how* accounting systems might embrace techniques that make a resource of the inevitable ambiguity that defines humanitarian crises.

7. Conclusion

7.1. <u>Introduction and Overview</u>

Humanitarian crises, and consequently the management of refugee camps, are major contemporary socio-political challenges. This study has argued that, due to their intrinsic complexity and instability, such settings present unique opportunities to theoretically advance knowledge in relation to the roles accounting can play in such extreme environments. In particular, the study explored several ways in which accounting systems can embrace the complexity and ambiguity that define humanitarian crisis contexts and attempt to proactively build on them. To investigate these issues, the research analysed how the most widely used system to evaluate performance in humanitarian crises, the Sphere Handbook, facilitated engagement with the unexpected. In terms of its theoretical framing, the study drew from and further developed the concept of heterarchy, interpreted as governance through difference (Stark, 2009; Lamont, 2012).

The study makes three main contributions that are formulated in detail in this chapter. The first contribution is to the literature on ambiguity as a resource to enhance the evocative power of accounting systems (Cooper et al., 1981: March, 1987; Chenhall et al., 2013; Quattrone, 2009, 2015). Based on the findings of the case studies, the study shows how ambiguity not only actively

promoted an attitude of scepticism, which meant that users were less likely to be blinded by the evaluation system's shortcomings, but managers used the openly incomplete measures as adaptable templates for the engagement with the humanitarian crisis. The second contribution of this study is to scholarly understanding of the concept of heterarchy, which offers insights into a distinct notion of engaging with dissonant principles of evaluation in contexts of uncertainty beyond notions of compromise (Stark, 2009; Lamont, 2012; Berthoin-Antal et al., 2015). The thesis suggests that tensions and dissonance are neither enough to sustain heterarchies, nor is any type of tension equally valuable. By drawing from and further developing the notion of heterarchy, the research shows how heterarchical principles can provide promising insights for evaluation systems, thereby enhancing scholarly understanding about elements that enable and constrain the practice of heterarchies. Based on these insights, the third contribution of this study is to the emerging accounting literature on humanitarian response operations (Sargiacomo, 2015; Everett and Friessen, 2010; Taylor et al., 2014; Lai et al., 2014). By problematizing issues of complexity and ambiguity as key challenges in humanitarian response operations, the study contributes to scholarly understanding of how accounting technologies can be implicated in the adaptive management of humanitarian crises (see Ramalingam, 2013; Weick and Sutcliffe, 2015).

Aside from the theoretical contributions, this concluding chapter also outlines several limitations of the thesis. These limitations arose partly from the practical challenges of engaging with and conducting research in humanitarian crisis settings and refugee camps. They also relate to methodological limitations. Furthermore, the chapter highlights the practical

implications of this study. These include implications and suggestions for the designers of humanitarian performance evaluation systems as well as for the way humanitarian organizations engage with internal control and information systems. The chapter concludes with possibilities for future research, which relate to challenges of performance evaluation systems in extreme settings and to the further development of the notion of heterarchy.

7.2. <u>Contributions</u>

7.2.1. <u>*Contribution 1: Making Ambiguity a Resource for Accounting Systems*</u>

The first contribution this study makes is to the literature on ambiguity as a resource to enhance the evocative power of accounting systems. While recent studies have shown interest in the role of incompleteness and ambiguity in the design and practice of accounting systems (e.g. Dambrin and Robson, 2011; Jordan and Messner, 2012), the idea of proactively fostering ambiguity as a resource has been explored to a lesser extent (Cooper et al., 1981: March, 1987; Busco and Quattrone, 2015). The point of making a resource of ambiguity is not to renounce any form of measurement; to the contrary, an important theme running through the findings chapters was how important the engagement with performance measures was for the managers in making sense of the unfolding challenges they confronted on a daily basis. The point is therefore to develop the "interplay, rather than the strict opposition, of [the ideal of] algorithmic formulation and judgement within calculative practices" (Jordan et al., 2016, p.4). In other words, the argument is about developing different

forms of engaging with performance evaluation systems that enable a more reflective and flexible manner of working with measures.

Enriching this strand of literature, the study contributes to scholarly understanding by detailing *how* evaluation systems might foster ambiguity as a resource for engagement with complex settings such as humanitarian crises. As emerges from the findings and the discussion, Sphere's evaluation system does not become performative because users trust in or 'make do' with its incomplete metrics (Dambrin and Robson, 2011; Andon et al., 2007), nor because managers attempt to repair or distance themselves from the imperfect measures (Jordan and Messner, 2012). Instead, the conceptualization of Sphere as an openly incomplete technique for performance evaluation actively stimulated scepticism of its metrics. In its engagement with the complexity of the camp, Sphere became influential precisely because users were not overly trusting in its prescriptions and were therefore less inclined to be deceived by the shortcomings of its measures.

Such differences with Dambrin and Robson's (2011) and Jordan and Messner's (2012) studies are not trivial. The explicit incompleteness of the evaluation system was crucial insofar as it worked against a tendency to oversimplify response requirements and quickly classify issues into something expected and familiar. As was frequently emphasized, it is important for humanitarian performance evaluation systems to work against a treacherous sense of security and complacency that can arise if measures are followed in an unreflective and de-contextualized manner, e.g. a check-list approach. It is only through these elements that the evaluation system was able to draw

managers' attention to emergent problems and anomalies and thereby overcame its limitations in anticipating the range of unpredictable response requirements.

Another important element that promoted ambiguity in an organized manner was engagement with the tensions between Sphere's evaluative dimensions. Contributing to previous research that raises the possibility of embracing tensions (Stark, 2009; Chenhall et al., 2013; Quattrone, 2015a), this study highlights that tensions are not productive per se, but must be proactively fostered in a strategic way. Focusing on the uncertain gaps and tensions between Sphere's evaluation dimensions forced the managers to move away from a strict focus on the technicalities of the response operation, thereby breaking reductionism and inviting exploration into blind spots and unknown elements. However, the ensuing ambiguity was not random and did not lead to chaos (Holm, 2010), but was instead facilitated through the engagement with the Sphere Handbook, which served as a carrier of its method into the field operations that most of the managers were familiar with.

The refugee camp was itself a confusing and challenging space, with a multitude of contradictions forming a common part of everyday operational practices. Working with evaluative tensions in an uncoordinated manner would have resulted in simply more confusion by overburdening the analytical capabilities of the managers. In line with these insights, fostering evaluative tensions in an organized manner is important as it contributes to scholarly understanding of *how* accounting systems might embrace ambiguity

(Revellino and Mouritsen, 2015; Busco and Quattrone, 2015) and promote techniques that enhance and transform what is knowable (March, 1987).

Furthermore, working with an openly incomplete performance evaluation system also meant that users engaged with the measures in a more flexible manner. This point manifested itself in several different ways. Instead of treating the performance evaluation system as a fixed goal to be achieved, managers commonly used measures as imperfect templates for engagement with the varying demands of the response operation that always required contextual adjustments. This process of adjusting proactively invited tinkering and an active search for novel, contextually appropriate performance criteria. Using the performance measures as imperfect templates also encouraged managers to question some of the expectations and assumptions that underpinned the measures and comparisons with observations in the camp, which is consistent with an approach that prioritizes error minimization. Finally, the principle of open and participatory design encouraged the attraction of a diversity of users (e.g. affected populations), to take advantage of the distributed knowledge and creativity in the camp. Combined, these elements ensured that the evaluation metrics remained open for contextualization and adaptation, which constituted important factors for confronting the emergent problems and dynamism of the response operation. While these insights build on the literature on the role of incompleteness in the performative nature of accounting metrics, they also contribute to scholarly understanding of *how* the proactive use of incomplete measures can enhance the evocative power of accounting systems (March, 1987; Quattrone, 2009; Qu and Cooper, 2011).

The second contribution of the study concerns the concept of heterarchy. Against the background of a renewed theoretical interest in processes of valuation and evaluation (see Boltanski and Thevenot, 2006; Kornberger et al., 2015; Berthoin-Antal et al., 2015), heterarchy was introduced as an insightful concept for engaging with heterogeneous notions of value in contexts of uncertainty (Stark, 2009; Lamont, 2012). Instead of methodically separating different ideal(ized) notions of value (see Boltanski and Thevenot, 2006), the notion of heterarchy seeks to advance meaningful ways of organizing the interplay between different evaluative principles to enable reflective judgement and the re-combination of existing resources. However, it is precisely in relation to the question of how such an interplay between divergent evaluative principles might be organized that the notion of heterarchy is underdeveloped and under-theorized (see Holm, 2010; Lamont, 2012). While it is probably a common feature in most organizations to have divergent principles for evaluation, Stark's (2009) argument leaves open how their interaction may be coordinated and made productive. In line with this argument, the study also indicates that the notion of heterarchy requires additional elements to be brought to life.

Through the theorization of the four principles in the discussion – in-built tensions, open and participatory design, relational value and incompleteness, and enacting minimalist control through a community of practitioners – this study contributes to and enriches understandings of the elements that enable and constrain the practice of heterarchies in several ways (Stark, 2009; Lamont,

2012). The point that evaluative tensions are not productive per se, but have to be stimulated in an organized manner through the engagement with evaluation systems was outlined in section 7.2.1. However, it is important to highlight that this point addresses a notable gap in Stark's (2009) theorization. The study indicated that tensions are neither enough to sustain heterarchies, nor is any type of tensions equally valuable. Evaluative tensions only became meaningful in phases of exploring future possibilities for the camp through the interaction with the format of Sphere's evaluation system. In turn, when the managers entered a phase in which they adapted the response requirements to contextual challenges and opportunities, the importance of heterarchical tensions faded. In these phases of adaption, other elements became influential, including the principles of 'relational value and incompleteness' as well as 'minimalist control enacted through a community of practitioners'. While the former principle was central to enabling the organized recombination of existing resources, the latter provided an element of flexible control and constrained the possibilities of what could be done with the evaluation system. Combined, the four principles this study has theorized thus contribute in a meaningful manner to scholarly understanding of the elements that enable the practice of heterarchies, beyond a focus merely on evaluative tensions (Stark, 2009; Lamont, 2012).

In line with these insights, another important contribution to the conceptual understanding of heterarchies arising from the research is the need to pay attention to different temporal phases and how these phases shape interaction with heterarchical principles. The four principles theorized in this study did not only play different roles, but also became insightful during different stages of the humanitarian response operation. As was highlighted in section 6.2.5.,

the interaction between the heterarchical principles was shaped by a recursive alternation between phases of exploration on the one hand and adaptation on the other. These insights suggest that there is a need to allow for different phases during which tensions are opened up, while being mindful that an unconstrained engagement with the range of possibilities that emerge through heterarchical systems will lead to chaos. Notions of timing and temporality have not formed an important part of the theorization of heterarchies to this point (Stark, 2009; Lamont, 2012; Holm, 2010). By drawing attention to these concerns, the study thus provides an insightful contribution to the theoretical interest in how heterarchical principles can be embraced and facilitated in a proactive manner.

7.2.3. *Contribution 3: Accounting for Humanitarian Crises*

The third contribution this study makes is to the emerging accounting literature on humanitarian crises. Against the background of the increasing occurrence of humanitarian and natural disasters around the world, recent studies have stressed that it is important to advance scholarly understanding of the possible roles accounting systems might play in the distribution of scarce resources, including food and water, and the management of dislocated populations in such unstable settings (Walker, 2016; Sargiacomo, 2015). With its focus on the question of how performance evaluation systems might engage with and inform management practices in a large-scale refugee camp, this study directly speaks to the concerns expressed in this accounting literature.

The predominant focus of the existing studies on accounting in humanitarian crises has been on explaining how notions of long distance control are reconstituted through accounting technologies (Sargiacomo, 2015) and how humanitarians employ distinct accountability scripts to legitimate their actions in an ex post manner (Everett and Friessen, 2010; Taylor et al., 2014; Lai et al., 2014). While this thesis builds on this emerging strand of accounting research, it took a different perspective on the role of accounting in humanitarian crises. The study showed that a central challenge for humanitarian performance evaluation systems is to foster strategies that embrace the variations of such dynamic settings and enable the continuous modification of performance requirements. In other words, the study highlighted the need to theoretically advance the question of how designers might develop accounting systems that foster the reflective capacities of evaluation systems in the context of humanitarian crises.

Against this background, the study provides rich evidence of how heterarchical principles for performance evaluation contributed to a search based on raising important questions, problematizing anomalies and interrogating blind spots, rather than unreflectively following pre-packaged performance and control categories (Burchell et al., 1980; Quattrone, 2015a). By focusing on issues of complexity, ambiguity and incompleteness as key notions in humanitarian response operations, the study thus theorized how accounting technologies can be implicated in the adaptive management of humanitarian crises. This point is also particularly important for the theoretical debates emerging in the area of humanitarian crisis management. As emphasized by Ramalingam (2013: 138), "in the way aid agencies learn, strategize, organize, and perform, there is a clear manifestation of an obsession

with organized simplicity, with the ideas and principles of nineteenth century physical sciences". With its specific focus on the design and practice of evaluation systems, this study has contributed to scholarly understanding of how accounting technologies can shift away from such a preoccupation with organized simplicity.

7.3. <u>Limitations, Practical Implications and Future Research</u>

7.3.1. *<u>Limitations</u>*

As highlighted in the methodology section, through the selection of the research setting – a large scale refugee camp – the study confronted a particular challenge. Due to the interconnected governance arrangements of refugee camps, the most insightful interactions and negotiations do not take place within one single organization but instead between different organizations and between organizations and refugees. To trace performance evaluation practices, the study therefore adopted a strategy of following the actors and challenges (Czarniawksa Joerges, 2007; Latour, 2007; Stark, 2009). This strategy was pursued through field visits, observations and shadowing, interviews and document analysis. While this strategy was necessary to gain an in-depth understanding of the key performance evaluation issues that the managers in the camp confronted, to investigate and analyse the intricacies and particularities of the research setting, a long-term ethnographic study would have been the ideal method for data collection. Such an ethnographic study would have enabled the researcher to collect more observational data

and to shadow the different work teams in the camp in a more detailed manner, thereby complementing the interview data and document analysis.

Unfortunately, pursuing such a long-term ethnographic study was impossible due to several factors. Firstly, even though a longer stay for research had been initially planned, the security situation in the Middle East made these plans impossible to carry out. At the time of the research visits, the Syrian civil war was escalating and terrorist threats were rising, in particular in the border region where the camp was located. These factors led to a shortening of the research stay not only for reasons of personal safety, but also because the Jordanian Ministry of the Interior did not allow foreign researchers to remain prolonged periods in the camp. Another factor that complicated a long-term ethnographic study related to constraints in terms of funding and time. Given that the field research was self-funded, financial constraints would have also complicated the performance of a long-term ethnographic study. These financial and time constraints also further engagement with performance evaluation practices in refugee camps in different geographical areas around the world. Conducting such a comparative analysis could have complemented the data in a meaningful and interesting way.

Another important limitation of the research relates to the voice of refugees and disaster-affected people in the case studies. Due to several ethical and practical constraints, this research was strictly focused on the analysis of the performance evaluation practices in humanitarian aid organizations and does not include interviews with refugees. A first reason for this was that the researcher does not have appropriate training and qualifications that would

enable him to carry out research with psychologically and physically affected participants, in this case refugees. As a result, due to ethical concerns it was decided not to carry out interviews with refugees for the study. Furthermore, from a practical perspective, it would also have been challenging to conduct extensive interviews with refugees due to language translation and access issues, as well as time constraints in the research setting. Therefore, while the researcher met refugees during the camp visits, their views and stories are only implicitly included. While no direct quotes or examples are used, meeting with some refugee families on an informal basis significantly enriched the researcher's understanding of the camp setting, and it also gave a humbling insight into the devastating effects of the Syrian civil war on the lives of these families. Combined, these features result in the voice of refugees only being heard through second-hand accounts.

7.3.2. *Practical Implications*

Apart from its theoretical contributions, the study also has several practical implications. The first concerns the designers of humanitarian performance evaluation systems. The thesis has provided novel insights into techniques and principles that enable performance evaluation systems to confront the challenging setting of humanitarian response operations. Each of these principles – in-built tensions, open and participatory design, relational value and incompleteness, and minimalist control enacted through a community of practitioners – enhances understanding of how humanitarian performance evaluation systems can confront processes of exploration and adaption in an organized manner. In line with these insights, one of the practical implications

for system designers is to further develop the principles that emerged through this study and make them more explicit in the format and methodology of humanitarian evaluation systems.

Implementing these recommendations is more challenging than it may seem. In fact, a common theme running through the findings was that Sphere's evaluation system was frequently reduced to its technical dimension and indicators. Narrowing the evaluation requirements to merely technical concerns is an appealing incentive for humanitarian managers overwhelmed by the complexities of response operations. However, it was also a source of significant problems in both case studies, contributing to a further alienation and stigmatization of disaster-affected populations. System designers will therefore have to further consider how such a reductionist approach to humanitarian performance evaluation can be prevented and how the principles that emerged from this research can be embedded in a more explicit manner into humanitarian evaluation systems.

The implications of this study therefore suggest that the flexible and adaptive approach that distinguishes Sphere as an influential system of performance evaluation in humanitarian response operations should be protected against attempts to make it more compliance oriented. The debate about how humanitarians might build institutions and professional bodies that enforce compliance amongst the broad range of aid agencies with different objectives, funding and cultural backgrounds dates back at least to the beginning of the 1990s (see Barnett, 2011). However, the findings of this study suggest that such attempts should be treated with caution. Given the degrees of variation and

unpredictable response requirements, building such a compliance-oriented regime runs the risk of reducing the ability for reflective judgement, local adaptation and contextualization. This does not mean that humanitarians do not need specific guidelines for performance evaluation and quality assessments. Yet, as shown in the case studies, compliance and control can also be achieved through more minimalist means. While it is commonly assumed that control is best enforced through rigid hierarchies and tight coupling between objectives, measures and organizational actions, this study has shown how a more flexible approach could provide a powerful instrument to negotiate and enact contextualized notions of control as appropriate to the specific disaster setting.

In line with these arguments, another practical implication of this study concerns humanitarian practitioners and organizations. The four heterarchical principles for performance evaluation this study theorized highlight how humanitarian agencies can mobilize ambiguity in a strategic manner to enable reflective decision-making. To enhance their adaptive capabilities, the findings of this research suggest that humanitarian organizations should embrace such principles and embed them as explicitly as possible into their own internal performance evaluation, control and information management systems. While Sphere's evaluation system offers a widely known approach and template for engaging with humanitarian environments, it is also indisputable that all humanitarian organizations operate their own tailor-made evaluation systems. The more they can move away from simple check-list approaches and incorporate tactics and principles that enable adaptive management, the more they appear to be equipped to face the dynamic contexts of humanitarian response operations.

7.3.3. _Future Research_

One important question arising from this study that could be explored in future research is to what extent the principles for performance evaluation – in-built tensions, open and participatory design, relational value and incompleteness, and minimalist control enacted through a community of practitioners – might enable engagement with the unexpected in areas different from humanitarian crises. One significant feature of the study was its focus on an environment that is not only defined by high degrees of dynamism and complexity, but also where the stakes are extremely high because small mistakes can have significant consequences, ranging from loss of life to violent conflicts. These features clearly do not apply to every organization to the same extent. Thus, while the principles of performance evaluation have the potential to provide important insights for any organization in relation to the capacity to engage with the unexpected, the principles appear to be particularly relevant for organizations that have to perform under extraordinarily risky and challenging circumstances. In line with this argument, future studies could explore how the principles of performance evaluation hold up in settings that share some of these characteristics. For example, these may include areas such as health care, military operations, or public sector institutions that deal with preparedness and the management of natural and other types of disasters. In this context, it would also be insightful to explore whether further principles need to be developed and theorized to complement or even replace the ones this study has proposed.

Another related area of interest pertains to the notion of heterarchy, which has been further developed in this thesis as an insightful concept for theorizing different ways of mobilizing multiple notions of value to confront complex and uncertain settings. It was shown that embracing different value criteria and keeping them at play was a crucial to working against an overly reductionist and technical approach to managing humanitarian response operations, which, as emerged from the case studies, can run the risk of having dehumanizing effects. One area where further exploration of the notion of heterarchy is promising relates to questions of how other organizations that differ from the ones discussed in this study may embrace multiple evaluation criteria and benefit from their interaction and tensions. These concerns are of particular relevance for organizations such as those in health care or the public sector, which have to address and balance a variety of value criteria beyond merely financial metrics.

Another question emerging from the study that arises in relation to the notion of heterarchy was how to benefit from its productive ambiguity without producing confusion, conflict or chaos. The study showed that heterarchical tensions are not productive per se, but must be fostered in an organized manner through the mediation of an evaluation system. Furthermore, the study raised the issue that there may be different temporal phases in which the interrogation of evaluative tensions may be insightful, for example when organizations explore new opportunities and make an inventory of options. The relationship between notions of timing, temporality and heterarchical tensions opens up further potential for research, which were touched upon by this study, but there remains scope for further investigation.

7.4. <u>Concluding Remarks</u>

This final chapter has offered reflections on the study's objectives and design, reiterated the theoretical contributions, discussed limitations, practical implications and possibilities for future research. By developing four heterarchical principles for performance evaluation – in-built tensions between evaluative principles, open and participatory design, relational value and incompleteness, and the enactment of minimalist control through a community of practitioners – the study has investigated how accounting systems can meaningfully embrace the ambiguity and complexity of humanitarian crisis settings and make a resource of it. In so doing, the concluding chapter outlined that the study has contributed to three different strands of literature. Firstly, the thesis contributed to the literature on ambiguity as a resource to enhance the evocative power of accounting systems by theorizing *how* ambiguity can be productively generated to enable engagement with the unexpected and detect blind spots (see Cooper et al., 1981: March, 1987; Quattrone, 2015a, b; Revellino and Mouritsen, 2015). Secondly, the study has advanced the notion of heterarchy in a twofold manner. It has unpacked how heterarchical tensions can become productive through the mediation of an evaluation system without leading to confusion and chaos. Furthermore, the thesis has theorized additional principles that are necessary to sustain the openness of heterarchies in a procedurally organized manner (see Stark, 2009; Lamont, 2012). Thirdly, departing from the existing accounting literature's focus on controlling humanitarian crises (Sargiacomo, 2015; Sargiacomo et al., 2014; Everett and Friessen, 2010), the thesis has theorized how accounting technologies can serve as anomalizing devices for the adaptive management of crises.

As highlighted in this concluding chapter, the thesis has not only sought to advance theoretical understanding of the enabling role of heterarchical principles for performance evaluation. The thesis opened with a quote from the former UN High Commissioner for Refugees Antonio Gueteres, calling for a fundamental change in the way humanitarian crises are tackled from a political and administrative level. As the total number of displaced people has continuously increased to the highest levels since World War II (see UNHCR, 2015), the management of humanitarian crises has become an ever more important topic, particularly given their major socio-political implications. In light of this important contemporary challenge, the thesis offered practical insights on two levels. The idealized objective to build mechanisms that strongly enforce compliance with a narrow range of metrics across crisis settings has long formed part of the debate surrounding humanitarian performance evaluation systems (see Barnett, 2011). The findings of this thesis suggest that such proposals, while understandable, may not only be unrealistic but also hazardous by encouraging an overly static 'checklist approach' incapable of addressing the significant variability in humanitarian crisis settings. Departing from such a reductionist position, the study has explicated tactics and principles that can inform design processes of humanitarian performance evaluation systems with the aim to enable reflective decision-making and adaptive management. Furthermore, based on the empirical findings, the study has argued that humanitarian organizations should embed elements such as the four heterarchical principles into their internal evaluation and control systems to strengthen their adaptive capacities. To search for new solutions to the socio-political challenges posed by humanitarian crises, policy makers and administrators must look beyond familiar value categories. Heterarchical systems bring ambiguity. Yet it is in

this ambiguity that a new configuration of values that sustain communities can emerge.

References

Ahrens, T., & Chapman, C. S. (2004). Accounting for flexibility and efficiency: a field study of management control systems in a restaurant chain. *Contemporary Accounting Research*, *21* (2), 271-301.

Ahrens, T., & Chapman, C. S. (2006). Doing qualitative field research in management accounting: positioning data to contribute to theory. *Accounting, Organizations and Society*, *31* (8), 819-841.

Andon, P., Baxter, J. & Chua, W.F. (2007). Accounting change as relational drifting: a field study of experiments with performance measurement, *Management Accounting Research*, *18* (2), 273-308.

Andon, P. & Free, C. (2012). Auditing and crisis management: the 2010 Melbourne storm salary cap scandal. *Accounting, Organizations and Society*, *37*, 131-154.

Annisette, M. (2009). Race and ethnicity. In J. R. Edwards, & S. P. Walker (Eds.), *The Routledge companion to accounting history* (pp. 451-469). Abingdon: Routledge.

Annisette, M., & Richardson, A. J. (2011). Justification and accounting: Applying sociology of worth to accounting research. *Accounting, Auditing & Accountability Journal*, *24* (2), 229–249.

Apadurai, A. (2011). The Ghost in the Financial Machine. *Public Culture. 23* (3). 517 – 539.

Atkinson, P. (1990). *The ethnographic imagination: Textual construction of reality.* London: Routledge.

Baker R. (2014). Breakdown of accountability in the face of natural disasters: the case of the Hurricane Katrina. *Critical Perspectives on Accounting*, *25* (7), 620-632.

Barnett, M. (2011). *Empire of Humanity: A History of Humanitarianism.* New York: Cornell University Press.

Barzelay, M., (2007). Learning from second-hand experience: methodology for extrapolation-oriented case research. *Governance*, *20* (3), 521-543.

Beckert, J. (2011). Coordination through Dissonance. In Review Symposium on David Stark: The Sense of Dissonance. *Socio-Economic Review. 9*. 597 – 612.

Bédard, J., & Gendron, Y. (2004). Qualitative research on accounting: some thoughts on what occurs behind the scene, In C. Humphrey, & B. Lee (Eds.), *The Real Life Guide to Accounting Research: A Behind-The-Scenes View of Using Qualitative Research Methods* (pp. 191-206). Oxford: Elsevier Ltd.

Berger, P. & Luckman T. (1991). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge.* United States of America: Anchor Books.

Berthoin-Antal, A., Hutter, M. & Stark, D. (2015). *Moments of Valuation: Exploring Sites of Dissonance.* Oxford: Oxford University Press.

Betts, A., Bloom, L., & Weaver, N. (2015). *Refugee Innovation: Humanitarian Innovation that starts with Communities*. Available at: https://www.rsc.ox.ac.uk/files/publications/other/refugee-innovation-web-5-3mb.pdf Accessed on 4 November 2016.

Beunza, D. & Stark, D. (2003). The organization of responsiveness: Innovation and recovery in the trading rooms of Wall Street. *Socio-Economic Review, 1* (2), 116-135.

Boland R. J. (1979). Control, causality and information systems requirements. *Accounting, Organizations and Society*, *4* (4), 259-271.

Boland R. J. (1989). Beyond the objectivist and the subjectivist: Learning to read accounting as text. *Accounting, Organizations and Society*, *14* (5/6), 591-604.

Boland, R. J., Sharma, A. K. & Afonso, P. S. (2008). Designing management control in hybrid organizations: The role of path creation and morphogenesis. *Accounting, Organizations and Society*, *33* (7), 899-914.

Boltanski, L. & Thevenot, L. (1999). The sociology of critical capacity, *European Journal of Social Theory*, *2* (3), 359-77.

Boltanski, L. & Thevenot, L. (2006). *On Justification: Economies of Worth*. Princeton: Princeton University Press.

Borton, J. (2004). The joint evaluation of emergency assistance to Rwanda. *Humanitarian Exchange. 26* (March), 14-18.

Borton, J. (2008). *How feasible is it to monitor the overall performance of the humanitarian system*. Available at: http://www.alnap.org/resource/5675 Accessed on April 12th, 2014.

Boyns, T., & Edwards, J. R. (1996). The development of accounting in mid-19th century Britain: A non-disciplinary view. *Accounting, Auditing & Accountability Journal*, *9* (1), 40–60.

Briers, M. & Chua, W.F. (2001). The role of actor-networks and boundary objects in management accounting change: a field study of implementation of activity-based costing, *Accounting, Organizations and Society*, *26* (3), 237-69.

Bryman, A. (2008). *Social Research Methods.* Oxford, Oxford University Press

Bryman, A., & Bell, E. (2011). *Business Research Methods*, Oxford, Oxford University Press.

Bryman, A. (2015). *Social Research Methods.* Oxford, Oxford University Press.

Buchanan-Smith, M. (2003). How the Sphere project came into being: a case study of policy-making in the humanitarian aid sector and the relative influence of research. *Working Paper 215.* London: Overseas Development Institute.

Burchell, S., Clubb, C., Hopwood, A.G., Hughes, J. & Nahapiet, J. (1980). The roles of accounting in organizations and society. *Accounting, Organizations and Society*, *5* (1), 5-27.

Burchell, S., Clubb, C. & Hopwood, A.G. (1985). Accounting in its social context: towards a history of value added in the United Kingdom, *Accounting, Organizations and Society*, *10* (4), 381-413.

Burrell, G., & Morgan, G. (1979). *Sociological Paradigms and Organizational Analysis: Elements of the Sociology of Corporate Life*, Portsmouth, UK: Heinemann.

Bürkland, S., Mouritsen, J., & Loova, R. (2010). Difficulties of translation: Making action at a distance work in ERP system implementation. Working paper.

Busco C., & Quattrone, P. (2015). Exploring how the balanced scorecard engages and unfolds: articulating the visual power of accounting inscriptions, *Contemporary Accounting Research*, *32* (3), 1236-1262.

Callon, M. & Latour, B. (1981). Unscrweing the Big Leviathan: How actors Macrostructure Reality and How Sociologists Help Them Do So. In Knorr-Cetina, K and Cicourel, A. (Eds). *Advances in Social Theory and Methodology: Towards an Integration of Micro- and Macro-Sociologies* (pp. 277 – 303). Boston: Routledge and Kegan Paul.

Callon, M. (1986). Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of Saint Brieuc Bay. In J. Law (Ed.), *Power, Action and Belief: a new Sociology of Knowledge? Sociological Review Monograph* (pp. 196-233). Routledge and Kegan Paul.

Callon, M., Lascoumes, P. & Barthe, Y. (2009). *Acting in an uncertain world*. Cambridge, MA: The MIT Press.

Carruthers, B.G. (1995). Accounting, Ambiguity and the New Institutionalism, *Accounting, Organizations and Society*, *20* (4), 313-328.

Carruthers, B.G. & Espeland, W. (1991). Accounting for Rationality: double-entry bookkeeping and the rhetoric of economic rationality, *American Journal of Sociology*, *97* (1), 31-69.

Chandler, A. D., & Daems, H. (1979). Administrative coordination, allocation and monitoring: A comparative analysis of the emergence of accounting and organization in the USA and Europe. *Accounting, Organizations and Society*, *4* (1/2), 3–20.

Chapman, C. S., Cooper, D. J., & Miller, P. (2009). *Accounting, Organizations and Institutions*, Oxford: Oxford University Press.

Chenhall, R. H., Kallunki, J.-P., & Silvola, H. (2011). Exploring the Relationships between Strategy, Innovation, and Management Control Systems: The Roles of Social Networking, Organic Innovative Culture, and Formal Controls. *Journal of Management Accounting Research*, *23* (1), 99-128.

Chenhall, R., Hall, M., & Smith, D. (2013). Performance measurement, modes of evaluation and the development of compromising accounts. *Accounting, Organizations and Society*, *38* (4), 268-287.

Chua, W.F. (1986), Radical developments in accounting thought, *The Accounting Review*, *61* (4), 601-32.

Cooper D. J. (1983). Tidiness, muddle and things: commonalities and divergencies in two approaches to management accounting research. *Accounting, Organizations and Society*, *8* (2/3), 269-286.

Cooper, D. J., Hayes, D., & Wolff, F. (1981). Accounting in organized anarchies: understanding and designing accounting systems in ambiguous situations. *Accounting, Organizations and Society*, *8* (2/3), 269-286.

Cooper, D. J. & Morgan, W. (2008). Case study research in accounting. *Accounting Horizons*, *22* (2), 159-178.

Coslor, E. (2016). Transparency in an opaque market: evaluative frictions between 'thick' and 'thin' price data in the art market. *Accounting, Organizations and Society*, Article in press.

Crumley, C. L. (1995). Heterarchy and the analysis of complex societies. *Archeological Papers of the American Anthropological Association*, *6* (1), 1-5.

Czarniawska-Joerges, B. (2007). *Shadowing and other Techniques for Doing Fieldwork in Modern Societies*. Copenhagen: Copenhagen Business School Press DK.

Czarniawska, B. (2009). *A Theory of Organizing.* Cheltenham and Northampton: Edward Elgar Publishing.

Czarniawska, B. (2014). *Social Science Research: From Field to Desk.* London: Sage Publications.

Dambrin, C. & Robson, K. (2011). Tracing performance in the pharmaceutical industry: ambivalence, opacity and the performativity of flawed measures. *Accounting, Organizations and Society, 36* (7), 428-455.

Davis, B. and McNeilage, P. F. (1990). Acquisition of correct vowel production: A quantitative case study. *Journal of Speech, Language, and Hearing Research. 33* (1). 16-27

Dechow, N. & Mouritsen, J. (2005). Enterprise resource planning systems, management control and the quest for integration, *Accounting, Organizations and Society*, *30* (7/8), 691-733.

Dewey, J., (1915), 'The logic of judgments of practise (Part I)', *The Journal of Philosophy, Psychology and Scientific Methods*, *12* (19), 505–523.

Dewey, J., (1923). Values, liking, and thought, *The Journal of Philosophy, 20* (23*)*, 617–622.

Dewey, J., (1939). *A Theory of Valuation*, Chicago: University of Chicago Press.

Diamond, J. (1996). The roots of radicalism. *The New York Review of Books*, 4-6.

DiMaggio, P. J. & Powell, W. (1983). The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48* (2), 147-160.

Ebrahim, A. (2005). Accountability myopia: Losing sight of organizational learning. *Nonprofit and Voluntary Sector Quarterly, 34* (1), 56–87.

ECBP (2007). The good enough guide: impact measurement and accountability in emergencies. London: Emergency Capacity Building Project.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14* (4), 532-550.

Elliott, A., & Turner, B. S. (2012). *On society*. Cambridge: Polity Press.

Espeland, W., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology, 113* (1), 1–40.

Everett, J. & Friesen, C. (2010). Humanitarian accountability and performance in the theatre d'absurde. *Critical Perspectives on Accounting, 21* (6), 468–485.

Ezzamel, M. (2012). *Accounting and Order*. Abingdon: Routledge.

Flyvbjerg, B. (2001). *Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again*. Cambridge, UK: Cambridge University Press.

Flyvbjerg B. (2006). Five Misunderstandings about case study research, *Qualitative enquiry*, *12* (2), 219-245.

Fourcade M., & Healy K. (2007). Moral view of market society. *Annual Review of Sociology. 33*, 285–311

Garud, R., Sanjay, J. & Tuertscher, P. (2008). Incomplete by design and designing for incompleteness. *Organization Studies*, *29*, 351- 371.

Geertz, C. (1995). *After the Fact: Two Countries, Four Decades, One Anthropologist*. Cambridge, MA: Harvard University Press.

Gehman, J., Trevino, L., & Garud, R. (2013). Values work: a process study of the emergence and performance of organizational values. *Academy of Management Journal*, *56* (1), 84–112.

Gendron, Y., Cooper, D.J. & Townley, B. (2007). The construction of auditing expertise in measuring government performance, *Accounting, Organizations and Society*, *32* (1/2), 101-29.

Glaser, B & Strauss, A. (1967). *The Discovery of Grounded Theory. Strategies for Qualitative Research.* Chicago, Aldine.

Gray, R. H., Bebbington, J., & Collison, D. (2006). NGOs, civil society and accountability: Making the people accountable to capital. *Accounting, Auditing and Accountability Journal*, *19* (3), 319–348.

Greer, S., & Neu, D. (2009). Indigenous peoples and colonialism. In J. R. Edwards, & S. P.Walker (Eds.), *The Routledge Companion to Accounting History* (pp. 470-484). Abingdon: Routledge.

Harrell-Bond, B. (1986). *Imposing Aid: Emergency Assistance to Refugees.* Oxford: Oxford University Press.

Helgesson, C & Muniesa, F. (2013). For what it's worth: An introduction to valuation studies. *Valuation Studies, 1* (1), 1-10.

Hines, R.D. (1988), Financial accounting: in communicating reality, we construct reality, *Accounting, Organizations and Society*, *13* (3), 251-61.

Holm, P. (2010). Review: the sense of dissonance: accounts of worth in economic life. *Administrative Science Quarterly*, *55* (2), 332-334.

Hopper, T., Storey, J. & Willmott, H. (1987), Accounting for accounting: towards the development of a dialectical perspective, *Accounting, Organizations and Society*, *12* (5), 437-56.

Hopwood, A.G. (1976). Editorial. The path ahead. *Accounting, Organizations and Society*, *1* (1), 1-4.

Hopwood, A.G. (1977a). Editorial. *Accounting, Organizations and Society*, *2* (1), 1-2.

Hopwood, A.G. (1977b). Editorial. *Accounting, Organizations and Society*, *2* (3), 187-188.

Hopwood, A.G. (1977c). Editorial. *Accounting, Organizations and Society, 2* (4), 277-278.

Hopwood, A.G. (1978a). Editorial. Accounting research and the world of action. *Accounting, Organizations and Society*, *3* (2), 93-95.

Hopwood, A.G. (1978b). Editorial. *Accounting, Organizations and Society*, *3* (3/4), 189-190.

Hopwood, A., (1983). On trying to study accounting in the contexts in which it operates. *Accounting, Organizations and Society*, *8* (2/3), 287–305.

Hopwood, A.G. (1985). The tale of a committee that never reported: Disagreements on intertwining accounting with the social. *Accounting, Organizations and Society*, *10* (3), 361- 377.

Hopwood, A.G. (1987). The archaeology of accounting systems. *Accounting, Organizations and Society*, *12* (3), 207-234.

Hopwood, A. G. (1990). Accounting and organisation change. *Accounting, Auditing & Accountability Journal*, 3 (1), 7-17.

Hopwood, A. G. (2005). Editorial. *Accounting, Organizations and Society*, *30* (7/8), 585-586.

Hopwood, A.G. (2009). Editorial. Reflections and projections – and many, many thanks. *Accounting, Organizations and Society*, *34* (8), 887-894.

Hoskin, K. W. & Macve, R. H. (1986). Accounting and examination: a genealogy of disciplinary power, *Accounting, Organizations and Society*, *11* (2), 105-136.

James, E. (2008). *Managing Humanitarian Relief: An Operational guide for NGO's*. UK: Practical Action Publishing.

Jay, J. (2013). Navigating paradox as a mechanism of change and innovation in hybrid organizations. *Academy of Management Journal*, *56* (1), 137–159.

Jeacle, I., & Carter, C. (2011). In TripAdvisor we trust: Calculative regimes and abstract systems. *Accounting, Organisations and Society*, *36* (4), 293–309.

JEEAR. (1996). *The international Response to Conflict and Genocide: Lessons from the Rwanda Experience*. Available at: http://www.alnap.org/resource/2517 Accessed on March 20th, 2015.

Jordan, S., Messner, M., & Becker, A. (2009). Reflection and mindfulness in organizations: rationales and possibilities for integration, *Management Learning*, *40* (4), 465-473.

Jordan, S. & Messner, M. (2012). Enabling control and the problem of incomplete performance indicators. *Accounting, Organizations and Society*, *37* (8), 544-564.

Jordan, S., Jorgensen, L., & Mitterhofer, H. (2013). Performing risk and the project: Risk maps as mediating instruments. *Management Accounting Research*, *24* (2), 156-174.

Jordan, S., Mitterhofer, H. & Jorgesen, L. (2016). The interdiscursive appeal of risk matrices: collective symbols, flexibility normalism and the interplay of 'risk' and 'uncertainty'. *Accounting, Organizations and Society*, article in press.

Jørgensen, B. & Messner, M. (2010). Accounting and strategizing: a case study from new product development. *Accounting, Organizations and Society*. *35*, 184–204.

Justesen, L. & Mouritsen, J. (2011). Effects of actor-network theory in accounting research. *Accounting, Auditing & Accountability Journal, 24* (2). 161-193.

Karpik, L. (2010). *Valuing the Unique: the Economics of Singularities.* Princeton: Princeton University Press.

Knight, Frank H. (1921). *Risk, Uncertainty and Profit*. Kissimmee, Fla: Signalman.

Kornberger, M., & Carter, C. (2010). Manufacturing competition: How accounting practices shape strategy making in cities. *Accounting, Auditing & Accountability Journal*, 23 (3), 325–349.

Kornberger M., Justesen L., Mouritsen J. & Madsen A.K. (Eds.) (2015). *Making Things Valuable*, Oxford: Oxford University Press.

Kornberger, M. (forthcoming). Making strategy count: rethinking value in strategy research. *Organization Studies.*

Kuhn, T. S. (1987). What are scientific revolutions? In L. Kruger, L. J. Daston, & M. Heidelberger (Eds.), *The probabilistic revolution, Vol. 1: Ideas in history* (pp. 7-22). Cambridge, MA: MIT Press.

Kvale, S. (1996). *InterViews: An Introduction to Qualitative Research Interviewing.* Thousand Oaks, CA: Sage.

Lai, A., Leoni, G. & Stacchezzini, R. (2014). The socializing effects of accounting in flood recovery. *Critical Perspectives on Accounting, 25* (7), 579-603.

Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annual Review of Sociology*, *38*, 201–221.

Latour, B. (1987), *Science in Action*, Cambridge, MA: Harvard University Press.

Latour, B. (1988). The politics of explanation: An alternative. In: Woolgar S. (ed.) *Knowledge and reflexivity: New frontiers in the sociology of knowledge* (pp.155-176). London: Sage.

Latour, B. (1996). *Aramis, or the Love of Technology*. Cambridge, MA: Harvard University Press.

Latour, B. (1999). *Pandora's Hope*. Cambridge, MA: Harvard University Press.

Latour, B. (2004a). *The Politics of Nature*. Cambridge, MA: Harvard University Press.

Latour, B. (2004b). Why has critique run out of steam? From matters of fact to matters of concern. *Critical Inquiry*, *30* (2), 225-48.

Latour, B. (2005). *Reassembling the Social*. Oxford: Oxford University Press.

Law, J. (1997). Heterogeneities. Paper presented at the meeting on Uncertainty, Knowledge and Skill, 6–8 November, 1997 at Limburg University.

Llewellyn, S. & Northcott, D. (2005). The average hospital, *Accounting, Organizations and Society*, *30* (6), 555-83.

March, J. G. (1971). The technology of foolishness, *Civilokonomen*, 7-12.

March, J. G. & Olsen, J. (1976). *Ambiguity and Choice in Organizations*. Bergen: Universitetsforlaget.

March, J. G. (1981). Exploration and exploitation in organizational learning. *Organization Science*, *2* (1), 71 – 87.

March, J. G. (1987). Ambiguity and accounting: the elusive link between information and decision making. *Accounting, Organizations and Society*, *12* (2), 153-168.

March, J. G. (2006). Rationality, foolishness, and adaptive intelligence. *Strategic Management Journal*, 27 (3), 201-214.

Marx, K., (1976) *Capital: A Critique of Political Economy, vol. 1* Harmondsworth: Penguin Books.

Maxwell, D., Walker, P., Church, C., Harvey, P., Savage, K., Bailey, S., Hees, R., & Ahlendorf, M. L. (2008). Preventing corruption in humanitarian assistance. Berlin: Transparency International/Feinstein International Center research report.

McCulloch, W. (1945). A heterarchy of values determined by the topology of nervous nets. *Bulletin of Mathematical Biophysics, 7*, 89-93.

Mennicken, A & Miller, P. (2012). Accounting, territorialization, and power. *Foucault Studies, 13* (1), 4-24.

Miller, P., & O'Leary, T. (1987). Accounting and the construction of the governable person. *Accounting, Organizations and Society*, *12* (3), 235-265.

Miller, P. (1990). On the interrelations between accounting and the state. *Accounting, Organizations and Society*, *15* (4), 315-38.

Miller, P., & Rose, N. (1990). Governing Economic Life. *Economy and Society*, *19* (1), 1-31.

Miller, P. (1991). Accounting innovation beyond the enterprise: problematizing investment decisions and programming economic growth in the UK in the 1960s. *Accounting, Organizations and Society*, *16* (8), 733-62.

Miller, P., & Napier, C. (1993). Genealogies of calculations. *Accounting, Organizations and Society*, *18* (7/8), 631–647.

Miller, P. & O'Leary, T. (1993), Accounting expertise and the politics of the product: economic citizenship and modes of corporate governance, *Accounting, Organizations and Society*, *18* (2-3), 187-206.

Miller, P., & O'Leary, T. (1994). Accounting, "economic citizenship" and the spatial reordering of manufacture. *Accounting, Organizations and Society*, *19* (1), 15-43.

Miller, P. (1994). Accounting as social and institutional practice: an introduction. In A.G. Hopwood and P. Miller (eds) *Accounting as social and institutional practice* (pp. 1-39). Cambridge: Cambridge University Press.

Miller, P. (1998). The margins of accounting. *The European Accounting Review, 7* (4), 605–621.

Miller, P. & O'Leary, T. (2007). Mediating instruments and making markets: capital budgeting, science and the economy. *Accounting, Organizations and Society*, *32* (7-8), 701-734.

Miller, P., & Rose, N. (2008). *Governing the Present: Administering Economic, Social and Personal Life*, Cambridge, UK: Polity Press.

Miller, P. & Power, M. (2013). Accounting, organizing, and economizing: connecting accounting research and organization theory. *Academy of Management Annals*, *7* (1), 557-605.

Mintzberg, H. (1973). *The Nature of Managerial Work.* Englewood Cliffs, NJ: Prentice Hall.

Mol, A. (2002). *The Body Multiple: Ontology in Medical Practice.* United States of America: Duke University Press.

Morgan, G. (1988). Accounting as Reality construction: Towards a New Epistemology for Accounting Practice, *Accounting, Organizations and Society*, *13* (5), 477-485.

Mouritsen, J., Hansen, A. & Hansen, C.Ø. (2009), Short and long translations: management accounting calculations and innovation management, *Accounting, Organizations and Society*, *34* (6-7), 738-54.

Mouritsen, J. & Kreiner, K. (2016). Accounting, decisions and promises. *Accounting, Organizations and Society*, *49* (2), 21-31.

Muniesa, F. (2011). A flank movement in the understanding of valuation. *The Sociological Review*, *59* (2), 24-38.

OECD (2014). Paris Declaration and Accra Agenda for Action. Available at: http://www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm Accessed on 27 January 2014.

O'Dwyer, B., & Unerman, J. (2008). The paradox of greater NGO accountability: A case study of Amnesty Ireland. *Accounting, Organizations and Society*, *33* (7 - 8), 801-824.

O'Leary, S. (2016). Grassroots accountability promises in rights-based approaches to development: the role of transformative monitoring and evaluation in NGOs. *Accounting, Organizations and Society.* Article in press.

Oxford Dictionary (2016a). Definition of Crisis. Available at: http://www.oxforddictionaries.com/definition/english/crisis Accessed on March 20th , 2016.

Oxford Dictionary (2016b). Definition of Ontology. Available at: https://en.oxforddictionaries.com/definition/ontology Accessed on September 17th, 2016.

Pollock, N. & D'Adderio, L. (2012). Give me a two-by-to matrix and I will create the market: rankings, graphic visualization and sociomateriality. *Accounting, Organizations and Society*, 37 (8), 565-586.

Pollock, N. & Campagnolo, G. (2015). Subitizing Practices and Market Decisions: the role of Simple Graphs in Business valuations. In Kornberger M., Jusesen L., Mouritsen J. and Koed Madsen A. (Eds.) *Making things valuable* (pp. 89 – 108). Oxford University Press.

Puyou, F. & Quattrone, P. (2016). Figures, not Numbers: Accounting, Legitimacy, and the Making of *Socie-ties.* Working Paper

Power, M. (2004). *The Risk Management of Everything.* London: Demos.

Power, M. (2007). *Organized Uncertainty: Designing a World of Risk Management.* Oxford: Oxford University Press.

Preston, A., Cooper, D.J. & Coombs, R.W. (1992). Fabricating budgets: a study of the production of management budgeting in the National Health Service. *Accounting, Organizations and Society*, 17 (6), 561-93.

Preston, A. (2006). Enabling, enacting and maintaining action at a distance: a historical case study of the role of accounts in the reduction of the Navajo herds. *Accounting, Organizations and Society*, 31 (6), 559-78.

Qu, S. & Cooper, D. (2011). The role of inscriptions in producing a balanced scorecard. *Accounting, Organizations and Society*, 36 (6), 1-19

Quattrone, P. (2000). Constructivism and accounting research: towards a trans-disciplinary perspective, *Accounting, Auditing and Accountability Journal 13* (2), 130-55.

Quattrone, P. & Hopper, T. (2001). What does organizational change mean? Speculations on a taken for granted category, *Management Accounting Research*, *12* (4), 403-35.

Quattrone, P. & Hopper, T. (2005). A time–space odyssey: management control systems in two multinational organizations. *Accounting, Organizations and Society*, *30* (7–8), 735-764.

Quattrone, P., & Hopper, T. (2006). What is IT? SAP, accounting, and visibility in a multinational organisation. *Information & Organization, 16* (3), 212–250.

Quattrone, P. (2006) The Possibility of Testimony: A case for case study research. *Organization, 13* (1), 143 - 157

Quattrone, P. (2009). Books to be practiced: memory, the power of the visual, and the success of accounting. *Accounting, Organizations and Society*, *34* (1), 85-118.

Quattrone, P. (2015a). Governing social orders, unfolding rationality, and Jesuit accounting practices: a procedural approach to institutional logics. *Administrative Science Quarterly, 60* (3), 411-445.

Quattrone, P. (2015b). Value in the Age of Doubt: Accounting as a Maieutic Machine. In Kornberger M., Jusesen L., Mouritsen J. and Koed Madsen A. (Eds.) *Making things valuable* (pp. 38 – 61). Oxford University Press.

Ragin, C. C. (1992). "Casing" and the process of social inquiry. In C. C. Ragin & H. S. Becker (Eds.), *What is a case? Exploring the foundations of social inquiry* (pp. 217-226). Cambridge, UK: Cambridge University Press.

Ramalingam, B. (2013). *Aid at the Edge of Chaos: Rethinking International Cooperation in a Complex World.* Oxford: Oxford University Press.

Reinnecke, J. (2015). The Politics of Value: A Case Study of 'Conflict-free' Gold. In Kornberger M., Jusesen L., Mouritsen J. and Koed Madsen A. (Eds.) *Making things valuable* (pp. 209 – 231). Oxford: Oxford University Press.

Reinnecke, J. & Ansari, S. (2015). Taming wicked problems: The role of framing in the construction of corporate social responsibility. *Journal of Management Studies, 53* (3). 299-329.

Revellino, S. & Mouritsen, J. (2015). *The maieutic force of mediating instruments*, Paper presented at the IPA Conference, Stockholm, 2015.

Rieff, D. (2002). *A Bed for the Night: Humanitarianism in Crisis.* London: Vintage Publishing.

Robson, K. (1991). On the arenas of accounting change: the process of translation. *Accounting, Organizations and Society*, *16* (5-6), 547-70.

Robson, K. (1992). Accounting numbers as 'inscriptions': action at a distance and the development of accounting. *Accounting, Organizations and Society*, *17* (7), 685-708.

Rona-Tas, A. (2011). Review symposium on David Stark's the Sense of Dissonance. *Socio-Economic Review, 9* (3), 597 - 612

Rottenburg, R. (2009). *Far-Fetched Facts: A Parable of Development Aid.* United States of America: MIT Press.

Sandel, M. (2012). *What Money Can't Buy: The Moral Limits of Markets.* New York: Straus & Giroux

Sargiacomo, M. (2014). Accounting for natural disasters and humanitarian interventions. *Critical Perspectives on Accounting*, *25* (7), 576-578.

Sargiacomo, M. (2015). Earthquakes, exceptional government and extraordinary accounting. *Accounting, Organizations and Society*, *42* (2) 67-89.

Sargiacomo, M., Ianni, L. & Everett, J. (2014). Accounting for suffering: calculative practices in the field of disaster relief. *Critical Perspectives on Accounting, 25* (7), 652-669.

Schacter, D.L., Wang, P. L., Tulvig, E. & Freedman, M. (1982). Functional retrograde amnesia: A quantitative case study. *Neuropsychologia, 20* (5), 523 - 532.

Schumpeter, J. (1950). *Capitalism, Socialism and Democracy.* New York: Harper Torchbooks.

Silverman, D. (1993). *Interpreting Qualitative Data: A Practical Handbook.* London: Sage.

SCHR (1997). The Sphere Project: Minimum Standards in humanitarian Response: A Programme of the Steering Committee for Humanitarian Response and InterAction. Sphere Project Proposal

Shapin S. 2012. The sciences of subjectivity. *Social Studies of Science, 42* (2), 170–84.

Sombart, W. (1967). *The Quintessence of Capitalism.* New York: Howard Fertig.

Sphere Handbook (2011). *The Sphere Handbook.* Available at: http://www.sphereproject.org/resources/downloadpublications/search=1&aid=nne Accessed on 27 January 2014.

Sphere Project (2013). *HAP, People in Aid and the sphere Project on the Harmonisation of Humanitarian Standards.* Available at: http://www.sphereproject.org/news/statement-from-hap-people-in-aid-and-the-sphere-project-on-the-harmonisation-of-humanitarian-standards/ Accessed 20 March 2014.

Stark, D. (2009). *The Sense of Dissonance: Accounts of Worth in Economic Life.* Princeton: Princeton University Press.

Stockton, N. (2005). The Accountable Humanitarian. Available at: http://www.hapinternational.org/pool/files/the-accountable-humanitarian-2-12-05.pdf Accessed on March 20th, 2014.

Strathern M. (2000). The tyranny of transparency. *British Educational Research Journal*, 26 (3), 309–321.

Taylor, D., Tharapos, M. & Sideway, S. (2014). Downward accountability for a natural disaster recovery effort: evidence and issues from Australia's black Sunday. *Critical Perspectives on Accounting, 25* (7), 579-603.

Tinker, A. M., Merino, B. D. & Neimark, M. D. (1982). The normative origins of positive theories: ideology and accounting thought. *Accounting, Organizations and Society*, 7 (2), 167-200.

Tinker, T. (1991). The Accountant as Partisan. *Accounting, Organizations and Society*, 16 (3), 297-310.

UNHCR (2012). Global Refugee Crisis Unmatched in Recent History. Available at: http://www.undispatch.com/global-refugee-crisis-unmatched-in-recent-history Accessed on 14th November 2012.

UNHCR (2013a). Global Trends. Available at: http://www.unhcr.org/statistics/country/5399a14f9/unhcr-global-trends-2013.html Accessed on 25th November 2015.

UNHCR (2013b). Syrian Humanitarian Aid in the Technology Age. Available at:http://www.unhcr.org/cgibin/texis/vtx/refdaily?pass=463ef21123&id=528b0ab55 Accessed 25th November 2015.

UNHCR (2014a). Syria Jordan Response Plan. Available at: http://www.unhcr.org/syriarrp6/docs/syria-rrp6-jordan-response-plan.pdf Accessed on 25th November 2015.

UNHCR (2014b). Zaatari Water Network Technical Working Group. Available at: https://data.unhcr.org/syrianrefugees/download.php?id=5536 Accessed on 12th December 2015.

UNHCR (2015). Worldwide Displacement Hits All-time High as War and Persecution Increase. Available at: http://www.unhcr.org/558193896.html Accessed on 12 November 2015.

Unerman, J. & O'Dwyer, B. (2006a). On James Bond and the importance of NGO accountability. *Accounting, Auditing & Accountability Journal*, *19* (3), 305–318.

Unerman, J., & O'Dwyer, B. (2006b). Theorising accountability for NGO advocacy. *Accounting, Auditing & Accountability Journal*, *19* (3), 349–376.

Vollmer, H. (2013). *The Sociology of Disruption, Disaster and Social Change.* Cambridge, UK: Cambridge University Press.

Walker, S. P. (2000). Editorial. *Accounting History, 5* (2), 5-12.

Walker, S. P. (2004). Expense, social and moral control. Accounting and the administration of the old poor law in England and Wales. *Journal of Accounting and Public Policy, 23* (2), 85-127.

Walker, S. P. (2008). Accounting, Paper Shadows and the Stigmatized Poor. *Accounting, Organizations and Society*, *33* (4-5), 453-487.

Walker, S. P. (2014). Drought, Resettlement and Accounting. *Critical Perspectives on Accounting, 25* (7), 604-619.

Walker, S. P. (2016). Revisiting the Roles of Accounting in Society. *Accounting, Organizations and Society*, *49*, (2). 41-50.

Weber, M. (1956). *The Protestant Ethic and the Spirit of Capitalism*. New York: Charles Scribner.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *1*, 1-19.

Weick, K. E. (1979). *The Social Psychology of Organizing* (2nd ed.). Reading, NJ: Addison-Wesley.

Weick, K. E. (1988). Enacted sensemaking in crisis situations. *Journal of Management Studies*, *25* (4), 305-317.

Weick, K. E., Sutcliffe, A. (2007). *Managing the Unexpected: Resilient Performance in an Age of Uncertainty.* (2nd ed.). Reading, MA: Addison-Wesley.

Weick, K. E., Sutcliffe, A. (2015). *Managing the Unexpected: Sustained Performance in an Age of Uncertainty.* (3nd ed.). Reading, MA: Addison-Wesley.

Willmott, H. (1983). Paradigms for accounting research: critical reflections on Tomkins and Groves' Everyday accountant and researching his reality', *Accounting, Organizations and Society*, *8* (4), 389-405.

Woolgar, S., & Pawluch, D. (1985). Ontological gerrymandering: The anatomy of social problems explanations. *Social problems*, *32*(3), 214-227.

Woolgar, S. (2015). *It Could Be Otherwise,* Keynote speech presented at the IPA Conference, Stockholm, 2015.

Woozley, A.D. (1973). *Theory of Knowledge.* Great Britain: Hutchinson &Co.

Wouters, M., & Wilderom, C. (2008). Developing performance measurement systems as enabling formalization: A longitudinal field study of a logistics department. *Accounting, Organizations and Society*, *33* (4–5), 488–516.

Appendices

Appendix 1: Interview Catalogue Humanitarian Field Officer

Introduction

Can you give me a general outline of your job as…?

Can you describe the structure of your operations?

How many people directly report to you as a manager?

What is the nature of your objectives and how are they defined?

Technical Aspects of Performance Measurement System

How do you measure the performance of your operations? Which system was chosen? Why?

How was the system designed? Why like this?

Which criteria do you use (outcome, output, process, input, qualitative, quantitative)? Why? Challenges?

What kinds of devices or technology do you use to measure each criterion? What kind of data do they provide?

Could you provide an example of an evaluation process? Visuals?

How do you analyse the data? How frequently do you look at them? What do you do with the data once you have analysed them (communication?)

Do the data you receive provide a sufficient overview of your operational performance?

What is the impact of distance on your work relationship between you and your staff?

Do other organizations become involved in the performance measurement process? If so, is their involvement visible through the performance measurement technologies you told me about?

Do you trace performance to particular individuals or teams? If yes, how?

Do you rely on what operational staff have self-reported in the performance measurement of activities?

How is bogus or false information detected?

Is there any work on how reporting/performance technologies might be misused?

What happens when you fail to meet a measure?

In policy documents, humanitarian accountability is very often defined as: "the responsible use of power." Does this emphasis on 'power' make sense for you within the context of humanitarianism? Does it play any role in your day-to-day work?

Coordination

How do you coordinate activities with other field managers and organizations involved in the response operation? Which devices do you use?

How do you plan your response operations? What are the factors involved?

How do you detect challenges?

How do you adjust to challenges?

Rationale for using Sphere (if relevant).

Why are you using Sphere?

What do you like about Sphere and what do you find challenging, if anything?

At which stage of operations is Sphere used? From the beginning, half way through, for final evaluation only? Why like this? Challenges?

What is the role of Sphere standards in the coordination of the cluster system?

Note: If, for example, the health cluster realizes that crude mortality rate might be problematic, how do they coordinate with water/sanitation, nutrition or shelter?

How did you assess your performance before Sphere?

How and why does a humanitarian performance standard become accepted in your view? What are the factors that contribute to its acceptance?

Did the use of the Sphere standards change the way you conduct your operations? If yes, how?

Notes for this question: How did it define your perception of accountability? Community engagement? Coordination with other humanitarian organizations?

In your opinion, did humanitarian workers' responses and behaviour towards performance and control systems evolve over time? If yes, how, why and towards what?

Do you train your workers how to use Sphere? If yes, how?

Sphere also makes specific reference to a 'rights-based approach' towards humanitarian performance. Is this endorsement of a 'rights-based approach' of any relevance in your work?

Resistance to Humanitarian Performance evaluation

In your view, what are the dangers of using performance evaluation systems in your work?

Notes for this question: separates moral and technical aspects of performance? Places focus of humanitarian performance away from donors, root causes of conflict; hides that humanitarian organizations are not isolated actors?

Do you agree with arguments against humanitarian standards? See e.g. Medecins Sans Frontieres. If, yes how and why?

Conclusion

Finally, is there anything important regarding humanitarian performance in your view, which this interview has not touched upon?

Appendix 2: Question Catalogue Sphere

Background

What is your current role? What are your responsibilities?

What did you do before Sphere?

How did you become affiliated with Sphere? Why?

Sphere design and development

Can you describe the basic idea behind Sphere's evaluation system?

What are the crucial elements of Sphere in your view? Why were they designed like this?

How does Sphere conceptualize humanitarian performance?

What is Sphere's approach towards performance measurement?

Who was Sphere designed for? Who should use it? Why?

Do you feel like this humanitarian performance has become more calculable and measurable through Sphere?

How does Sphere address the variations between each different humanitarian crisis?

Is unpredictability an issue for Sphere's approach? If so, how does it address it?

Can you explain the role of the Handbook in the design of Sphere? Why is there such an emphasis on handbook as such?

Were there any important elements that did not make it into the Sphere handbook? Why were they left out?

How do you see Sphere's core ethical principles, such as the humanitarian charter, being related to the concept of effectiveness when it is used in the field? How does it relate to other important humanitarian concepts, such as impartiality, dignity?

In humanitarian performance measurement a concern is often voiced about how abstract ethical principles can be related to specific technical indicators? Did you have these debates? What were the controversies in these debates? How were they settled, if at all?

How is the process of updating and further developing Sphere organized? Who is involved? Why are they involved? How are they involved?

Are there any core building blocks that you would not touch when updating Sphere?

How do you organize trainings for Sphere? What are the aims of the trainings? What is their approach?

Why does Sphere not pursue certification? What is the future of humanitarian certification in your view?

Why is it not a membership organization?

What are Sphere's biggest achievements? What are its shortcomings?

Challenges

In your view, what are the dangers of using Sphere, if any?

Notes for this question: separates moral and technical aspects of performance? May place focus of humanitarian performance away from donors, root causes of conflict. May hide that humanitarian organizations are not isolated actors?
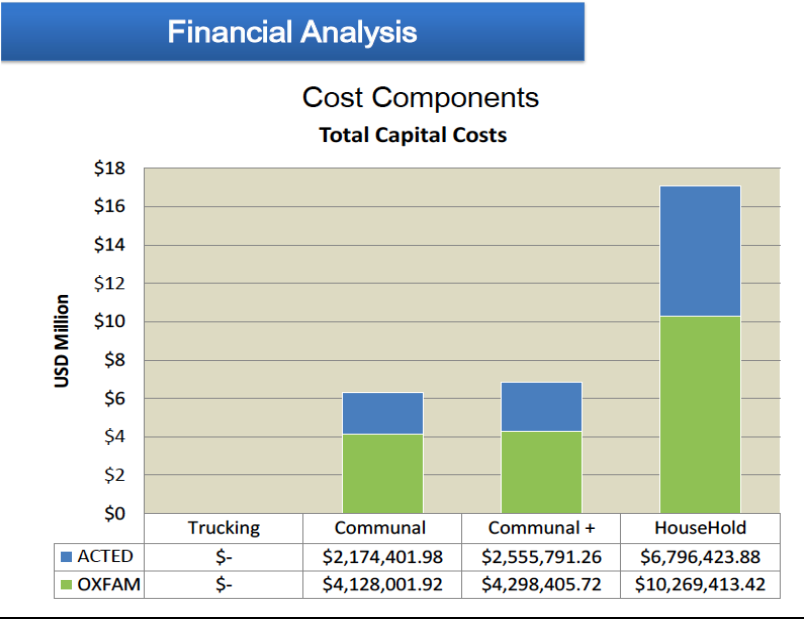
From your experience has Sphere ever been misused? Do you have an example?

How do you ensure that Sphere is used in the way that you would like it to be used?

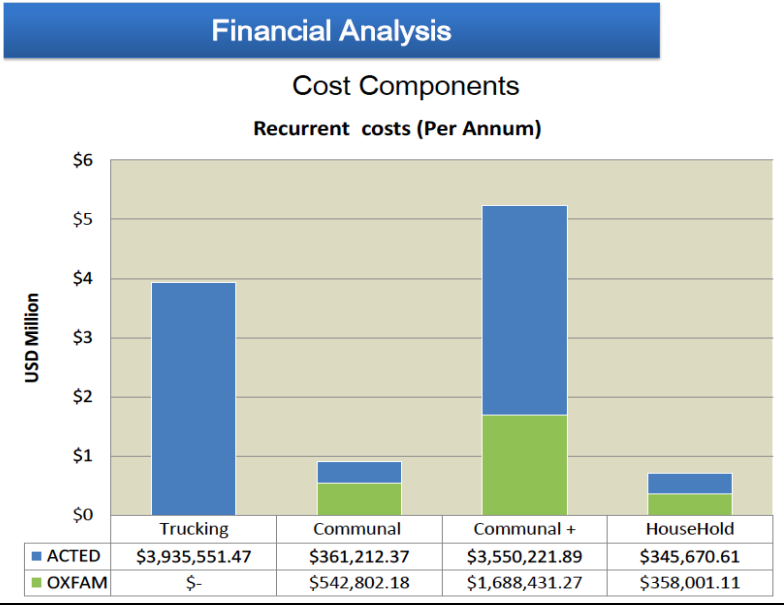*Other Possible Issues in Relation to Humanitarian Performance*

Finally, is there anything important regarding humanitarian performance in your view, which this interview has not touched upon?

Appendix 3: Projection of Total Capital Costs

Financial Analysis

Cost Components
Total Capital Costs

	Trucking	Communal	Communal +	HouseHold
■ ACTED	$-	$2,174,401.98	$2,555,791.26	$6,796,423.88
■ OXFAM	$-	$4,128,001.92	$4,298,405.72	$10,269,413.42

(UNHCR, 2014b, p.70)

Appendix 4: Projection of Recurrent Costs

Financial Analysis

Cost Components
Recurrent costs (Per Annum)

	Trucking	Communal	Communal +	HouseHold
■ ACTED	$3,935,551.47	$361,212.37	$3,550,221.89	$345,670.61
■ OXFAM	$-	$542,802.18	$1,688,431.27	$358,001.11

(UNHCR, 2014b, p.71)