IFHV
Institut für Friedenssicherungsrecht und Humanitäres Völkerrecht
Institute for International Law of Peace and Armed Conflict

A Balanced Scorecard for the Humanitarian Sector? Adaptability of the Balanced Scorecard Model to Sector-Wide Performance Management in Humanitarian Aid

Jan Wulf

IFHV Working Paper, Vol. 2, No. 1

Bibliographic information:

Title:	A Balanced Scorecard for the Humanitarian Sector? Adaptability of the Balanced Scorecard Model to Sector-Wide Performance Management in Humanitarian Aid. Feasibility and Implications
Author(s):	Jan Wulf
Source:	IFHV Working Papers, Vol. 2, No. 1
Date:	May 2012
DOI:	https://doi.org/10.17176/20220622-125602-0
ISSN:	2199-1367

Suggested citation:

Wulf, J. (2012). A Balanced Scorecard for the Humanitarian Sector? Adaptability of the Balanced Scorecard Model to Sector-Wide Performance Management in Humanitarian Aid. Feasibility and Implications. IFHV Working Paper, 2(1).

ISSN: 2199-1367

RUB

RUHR-UNIVERSITÄT BOCHUM

IFHV Working Paper Vol. 2, No. 1, May 2012

A Balanced Scorecard for the Humanitarian Sector?

Adaptability of the Balanced Scorecard Model to Sector-Wide Performance Management in Humanitarian Aid

Feasibility and Implications

Jan Wulf

IFHV Working Paper Vol. 2, No. 1, May 2012

URL: http://www.ruhr-uni-bochum.de/ifhv/documents/workingpapers/wp2_1.pdf

A Balanced Scorecard for the Humanitarian Sector?

Adaptability of the Balanced Scorecard Model to Sector-Wide Performance Management in Humanitarian Aid

Feasibility and Implications

Jan Wulf*

MA in International Humanitarian Action (NOHA) and BA in Business Administration j.wulf73@gmail.com

Abstract

Background: Today, performance management in the humanitarian sector comprises a multitude of different initiatives and approaches that seek to improve efficiency of project management, foster cooperation, or establish accountability through standards, tools, frameworks and participative approaches. This multitude has led scholars to call for comprehensive, integrated, and system-wide views on, approaches to, and frameworks for performance management. This paper ascertains the applicability of the Balanced Scorecard model to the system-wide management of performance for the humanitarian sector. It combines the concepts of performance management, the generic Balanced Scorecard for businesses and further approaches to extend the Balanced Scorecard for cross-company networks as well as to apply the Balanced Scorecard to the non-profit and public sector. These concepts are discussed against the background of existing performance initiatives and approaches in the humanitarian sector.

Results: The Balanced Scorecard works on the implicit assumption that all aspects of performance are measurable to a certain extent. The humanitarian sector shows characteristics that make it nearly impossible to isolate the contribution of humanitarian interventions to the overall mission to save lives, alleviate suffering, and maintain human dignity, rendering the Balanced Scorecard inapplicable to system-wide performance management in humanitarian aid.

* This paper is a revised version of a master thesis originally submitted at the Joint European Master's Programme in International Humanitarian Action (NOHA) at the Institute for International Law of Peace and Armed Conflict (IFHV) at the Ruhr University Bochum.

URL: http://www.ruhr-uni-bochum.de/ifhv/documents/workingpapers/wp2_1.pdf

A Balanced Scorecard for the Humanitarian Sector?

Adaptability of the Balanced Scorecard Model to Sector-Wide Performance Management in Humanitarian Aid

Feasibility and Implications

Content

List of Acronyms

AIDS	Acquired Immune Deficiency Syndrome
ALNAP	Active Learning Network for Accountability and Performance
BSC	Balanced Scorecard
CARE	Cooperative for Assistance and Relief Everywhere
CE DAT	Complex Emergency Database
CERF	Central Emergency Response Fund
CHAP	Common Humanitarian Action Plan
DAC	Development Assistance Committee
ECHO	European Commission's Directorate General for Humanitarian Aid and Civil Protection
FAO	United Nations Food and Agriculture Organization
GHD	Good Humanitarian Donorship
HAP	Humanitarian Accountability Partnership
HNTS	Health and Nutrition Tracking Service
HRI	Humanitarian Response Index
HIV	Human Immunodeficiency Virus
IASC	Inter-Agency Standing Committee
ICRC	International Committee of the Red Cross
IFRC	International Federation of Red Cross and Red Crescent Societies
INGO	International Non-Governmental Organization
LFA	Logical Framework Approach
KPI	Key Performance Indicator
NGO	Non-Governmental Organization
NPM	New Public Management
OCHA	Office for the Coordination of Humanitarian Affairs
OECD	Organisation for Economic Co-operation and Development
SMART	Standardized Monitoring and Assessment of Relief and Transitions
UN	United Nations
UNDP	United Nations Development Programme
UN-HABITAT	United Nations Human Settlements Programme
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
USD	United States Dollars
WFP	World Food Programme

1. Introduction

Since the beginning of the 1990s, the humanitarian sector has changed tremendously. The number and scale of emergencies has picked up and humanitarian funding has increased significantly. Established organizations have grown and expanded their operations, and new organizations have entered the stage. The early phase of these developments coincided with the Rwandan Genocide that led to a wakeup call for the humanitarian sector. For long, the notion that humanitarian aid inherently "does good" and good intentions are not to be scrutinized was prominent among humanitarian agencies. The findings of the joint evaluation of emergency assistance to Rwanda, carried out in 1996 caused strong repercussions to this notion. Under the recognition that aid can do harm and that new ways to effectively deliver aid have to be found, numerous initiatives and approaches emerged from various directions. The growth of professionalization initiatives, standards, codes, think tanks, accountability approaches, project management tools and methods, and cooperation initiatives was equally rapid as the growth of the sector as such.

> I would like to see a system where agencies collaborated more with one another; I would like to see a system where there was more cooperation and I would like to see a system where agencies had a collective and shared goal and recognized that there are properties that emerge from the interaction of their activities that are greater than just their own particular program. A national health service is greater than just one hospital. An education system is more than just about individual schools.
>
> *John Mitchell* (2010), Director of ALNAP

In today's humanitarian sector, not even *"the largest organization [can] launch an effective response on its own"* (Ramalingam et al. 2009, p. 3). At the same time, in comparison to the for-profit sector, humanitarian aid lacks a natural accountability mechanism. In the for-profit sector, the recipient of services or products possesses bargaining power that allows him to make purchasing decisions and thus to directly or indirectly influence the quality of the provided service or product. The quality of humanitarian services is not strictly regulated through such a competition mechanism. In fact, the bargaining power mostly lies with the donor agencies that dictate conditions onto international non-governmental organizations (INGOs), non-governmental organizations (NGOs), the United Nations (UN) system, and the Red Cross movement. The quality of aid is, therefore, a result of conscious efforts of humanitarian agencies to provide effective, efficient, and suitable aid, in ways that are based on a broader perspective than the single-agency perspective. However, it is not entirely clear how and to what extent the numerous performance initiatives contribute to improved quality of aid. There is

reasonable concern as to whether these multiple approaches to performance show overlap and gaps, and might be based on false judgment on priorities. Among scholars there is growing consensus that *"the actors and institutions that collectively undertake humanitarian action do not form a coherent and integrated system with shared principles, policies, and modus operandi."* (Stockton 2000 cited in Griekspoor/Sondorp 2001, p. 211). The Organisation for Economic Co-operation and Development (OECD) stresses that *"strategic frameworks need to be established which set the overall objectives for the international community's response"* (1999, p. 13). But while the problem has been phrased, and collective scholarly vows call for sector-wide approaches to performance (Griekspoor/Sondorp 2001, p. 209; Hofmann et al. 2004, pp. 3, 6; Mitchell 2008; Ramalingam et al. 2009), little has been proposed on the solution side. System-wide performance management is expected to:

- "Synthesize existing material into a comprehensive representation;
- Integrate different dimensions of performance;
- Generate indicators of performance assessment;
- Report on performance against agreed criteria, and on the evidence base for the judgments being made;
- Provide a platform for debate on performance throughout the system;
- Allow for adaptations and adjustments, both incremental and radical, to the humanitarian business model" (Ramalingam et al. 2009, p. 75).

But it has not been clearly expressed, what a sector-wide approach to performance could constitute of and how it would incorporate or relate to existing approaches to performance.

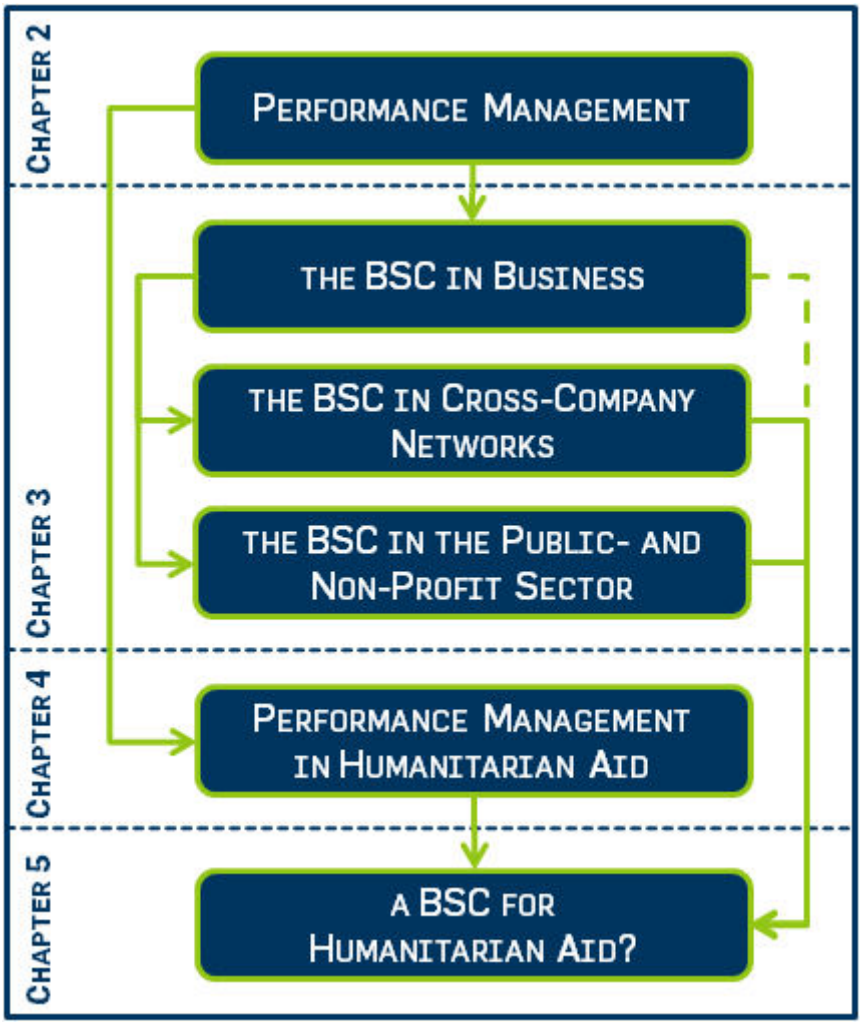
The Balanced Scorecard (BSC) is a strategic performance management tool that has been diffusing into companies worldwide since its emergence in 1997. Its success has led scholars to extend the research on the BSC. The current state of research discusses BSC models for the public and non-profit sector, and the extension of the model to cross-company networks in business. These two dimensions could provide a basis for the development of a system-wide approach for performance in humanitarian aid. Ramalingam et al. recently discussed a *"balanced approach to system-wide performance"* (ibid., p. 76), loosely based on the idea of the BSC. However, their proposal falls short of discussing the applicability of the BSC methodology to the various levels of performance in the humanitarian sector. This paper seeks to fill this gap, by finding an answer to the following research question:

> Under which conditions and through which adaptations could the BSC serve as a model for sector-wide, multi-level performance management in international humanitarian aid?

To approach this research question, first the theoretical background on performance management in general will be discussed. Then, the BSC and its methodology will be explained and put in relation to performance management.

Moreover, the current state of research on cross-company BSC models in business and the BSC in the public and non-profit sector will be explored. In a next step, the current state of performance management in humanitarian aid will be analyzed before the background of the concepts of performance management. Finally, the characteristics of the humanitarian sector and the identified challenges in humanitarian performance management will be used as a basis for the deduction of a potential model for sector-wide multi-level performance management in humanitarian aid from the cross-company BSC model and the BSC model for the public and non-profit sector. This paper concludes with the answer to the research question, important lessons drawn from the research and a future outlook on performance management in humanitarian aid. Figure 1 illustrates the structure of the main chapters of this paper.

Figure 1: Structure of the main chapters



2. State of the art in performance management

This section provides a definition of what the concepts of performance, performance measurement, and performance management compose of and explores the state of research on performance measurement and management by

expounding the most important concepts. Thus, it prepares the ground for understanding the theoretical frame for the BSC model and the appraisal of the state of the art of performance management in humanitarian aid that are discussed in the subsequent chapters.

2.1 Definition and delineation of performance management

To approach the concept of performance management, it appears to be beneficial to consider the two terms that it composes of. Amongst other definitions, the Oxford English Dictionary defines performance as *"the action or process of performing a task or function"* (2011). Management is defined as *"the process of dealing with or controlling things or people"* (ibid.). Drawing together these two definitions, performance management may be understood as the process of dealing with or controlling the process of performing a task or function. As unspecific as this definition might appear, as simplistic are the roots of the notion of performance. Accordingly, Dubnick finds that

"outside of any specific context, performance can be associated with a range of actions from the simple and mundane act of opening a car door, to the staging of the Broadway musical 'Chicago'. In all these forms, performance stands in the distinction from mere 'behaviour' in implying some degree of intent" (2005, p. 391).

Out of this natural and intuitive human drive to reflect on one's activities in order to improve their results, to minimize the effort spent on a set of activities, or to find alternative activities that lead to the same results, the management of performance has been taken to a more rational and institutional level for the dealing with far more complex issues than the opening of a car door.

Regardless of deep and broad research efforts conducted on performance management, there appears to be no such thing as a homogenous scientific field with a certain frame for understanding, or a common ground for what the essence of performance management is. Bouckaert and Halligan proclaim that *"performance management lacks a coherent treatment that explicates its significance, analyses its several dimensions as a working system and challenges its shortcomings"* (2006, p. 1). Partly this may be due to the fact that research on performance management as such has a variety of points of departure: various scientific or occupational fields have their own rationale for understanding the dynamics of performance within their principal field of research. In this sense, the notion of performance management can be seen as an additional layer on top of the capital subject of research for nearly *any* field. The result of this diverse approach leads to a state of knowledge that composes of scattered ideas, concepts, and unconsolidated convictions on the *how to* of performance management. Along these lines, Ramalingam et al. point out that

"performance is frequently presented as an umbrella for a host of other ideas – including effectiveness, productivity, quality, transparency and accountability – each of which leads to yet more frameworks and extensive literatures" (2009, p. 10).

In opposition to some authors who use the term of performance management in a sense, judging on the level of maturity or professionalism of an examined performance management approach or system (Ramalingam et al. 2009; Van Dooren et al. 2010; Bouckaert/Halligan 2008)[1], in the course of this research, performance management is used as a rather neutral term, regardless of the maturity of the respective performance management approach or system that is being dealt with.

2.2 The history of performance management

The roots of institutionalized performance management lie well beyond our present times. Throughout many centuries,[2] its evolvement has taken different, parallel paths. Besides the separate pursuits for better performance in the military sector and in the public sector, the assumedly most influential streaming has been the one of the for-profit sector. Due to the inherent nature of this sector, its actors always had to deal with the dynamic pressures of competition and continuously sought to optimize their efficiency in order to stay competitive. The striving for the optimizing of operations is a logical result of these pressures.

A notable thinker contributing largely to the evolvement of management techniques as such and especially the precursors of what performance management is today was Frederick Winslow Taylor. Being mechanical engineer by profession, he worked in the field of industrial manufacturing around the beginning of the 20th century. Taylor developed a new way of approaching the challenges of running industrial manufacturing businesses. An essential element to his philosophy of management was the introduction of a scientific approach to managerial decisions (Taylor 1970; Sheldon 1924; Locke 1982, p. 14). His scientific management concept intended to replace tradition, rule of thumb, guesswork, precedent, personal opinion, or hearsay, which were the predominant management practices at that time, by decisions based on proven facts (Locke 1982, p. 14). Through this fundamental paradigm shift and the associated call for standardized procedures, management by setting of defined tasks, structured learning and feedback cycles, and output-based incentives (Taylor 1994; Ramalingam et al. 2009, p. 12f.), Taylor set the stage for further development of management techniques that today are commonly used in companies around the world.

Since then, academically and in practice, the field of performance management has been undergoing changes stimulated from various directions. Internal or external triggers have caused the redefinition or reinvention of parts of or the whole model of thought behind performance management (Kennerley/Neely

[1] According to these authors, in contrast to their definition of performance management, the generic way of dealing with performance, whether it is in an un-reflected manner (or *not* dealing with performance) or more reflected or mature manner is merely referred to as performance or managing performance.

[2] Lawrie et al. trace the roots of performance management back to at least the times of ancient Egypt (2006, p. 2).

2002, p. 1241). Around the beginning of the 1990s, management accounting systems, commonly used in businesses to monitor the company's success, increasingly became subject to scrutiny: they were considered to be falling short of grasping the company's operations in a way that would comprehend all factors relevant to success. During times of economical continuity, most companies exclusively relied on financial key figures as indicators of success, neglecting that *"financial figures are retrospective of nature and provide feedback too late to be a management tool for improvement"* (Bedrup 1995a, p. 65). In this context, Bedrup mentions a quite grim example of an overly excessive financial view on a company's performance:

"a Russian director of a large agricultural collective received the productivity award for production of meat for subsequent three years. The fourth year he shot himself. His successor soon discovered that the collective had no breeders anymore" (Bedrup 1995a, p. 85).

The insights that accompanied this paradigm shift, which Eccles and Neely call a 'performance measurement revolution' (Eccles 1991; Neely 1999), initially made larger enterprises change their priorities on how to measure their performance. A company's objectives and its environment have shifted into the core of what performance measurement and -management should encompass. The once uniform financial view on performance that fit the majority of all companies was replaced by a rather individualized process of developing and administering multidimensional performance management systems. Frameworks evolved to allow companies transform their objectives and the features of the environment they operate in, into an appropriate system for performance management. For instance, Kennerly and Neely list the following prominent frameworks (2002, p. 1224):

- the Balanced Scorecard (Kaplan/Norton 1992),
- the Performance Prism (Kennerley et al. 2007),
- the Performance Measurement Matrix (Keegan et al. 1989),
- the Results and Determinants Framework (Fitzgerald et al. 1991), and
- the SMART pyramid (Lynch/Cross 1992).

The academic novelty appears to have reached the practitioner's sphere rapidly: *"survey data suggests that between 40 and 60 per cent of companies significantly changed their measurement systems between 1995 and 2000"* (Kennerley/Neely 2002, p. 1222).

The introduction of performance management concepts in the public sector had seen their point of departure around the same time as the private sector. Researchers conclude that the 1990s constituted an 'age of performance' (Talbot 1999; Bouckaert 1996; Ramalingam et al. 2009) for governments, increasingly implementing documentations and stressing the need for explicit outcomes of government action (Radin 2000, p. 168 cited in G. Bouckaert/J. Halligan 2006, p. 2). Ramalingam et al. find that *"the New Public Management movement arose in response to a number of pressures"* (2009, p. 21):

- *"economic pressures, including budget deficits, structural problems, growing competitiveness and globalisation, and*
- *political and social pressures, including a lack of public confidence in government, growing demand for better and more responsive services, and better accountability for achieving results with taxpayer money"* (see also: Van Dooren et al. 2010, p. 4).

Recent publications on performance in public management (including new 'movements' such as *evidence-based policy*; Van Dooren et al. 2010, p. 44) suggest that the philosopher's stone has not yet been found in the scientific debate on performance management in the public sector.[3] Bouckaert et al. state that attempts of implementing the ideas of new public management (NPM) have led to different levels of success in different countries (2010), leaving much room for realization of state-of-the-art concepts, which in turn are subject to an ongoing debate. The crisis of the international financial system of 2008 will certainly add an additional layer of complexity to this already multidimensional discussion.

Performance management in development aid has seen lesser consolidated implementation efforts than in public management as a whole. This may be attributed to a lower importance of the development agenda, compared to other segments and services in the public sector that stand in a more direct connection with the citizens' trust in their government and are therefore a higher priority for governments' strategies.[4] Moreover, compared to the 'classical public policy' the development sector is an international field with a multitude of actors in different functions. Potential efforts to introduce and establish performance management concepts require much more extensive coordination. However, one notable aspiration to introduce systematic and consolidated means of performance management is currently ongoing: the Millennium Project represents a one-time program seeking to align priorities and strategies, to harmonize operational procedures, and to make the delivery of aid traceable (Ramalingam et al. 2009, p. 29). The accompanying 'aid effectiveness agenda' adds operational principles to the strategy and stipulates performance assessment laid out in the Managing for Development Results Sourcebook. Even though the methodology and structure is a good example of results-based management, the full-blown benefits of this aspiring initiative remain to be assessed by the deadline of the Millennium Project in 2015 the earliest.

Reflecting upon the development of performance management, in the following sections, parallels are drawn and distinctions are made based on the identified characteristics of the private sector, the public sector, and the development sector. In the further course of exploring today's performance management

[3] For recent publications on performance management in the public sector, see e.g.: Van Dooren 2009; Bouckaert/Halligan 2008; Halligan 2006; Lawrie et al. 2006; Bouckaert/Halligan 2006.

[4] In simple terms, the running of a national public health system would be of greater importance to a government than fostering agricultural micro-credits in South-Eastern Asia, since the former increases the trust of citizens in its government, more than the latter.

approaches in humanitarian aid, as a basis for the actual *managing* of performance, the structural defining, gathering, and interpreting of information, the concept of *performance measurement* serves as a point of entry, before the characteristics of *performance management* as such are further expounded.

2.3 Characteristics of performance management

The fundamental motif behind measuring institutional performance is the intentional directing of attention towards those aspects of an organization's performance in operations that are most vital to its success. Put differently, "*what you measure is what you get"* (Kaplan/Norton 1992, p. 71). This bold statement assumes that if an organizational unit (e.g. an individual, a team, an organization, a network of organizations, or even a whole system of organizations) is able to identify and measure the factors that are the key drivers to achieving its formulated strategy, objectives and goals, the achievement of these will be the effect. Surely, reality is more complex and involves a number of aspects that need to be taken into consideration to allow this suggestion to become reality. Nevertheless, it represents a good point of departure for exploring the provisions of the concept of performance measurement.

The baseline for performance measurement and at the same time the most commonly used logic of performance measurement is built on the basis of a production model (Van Dooren et al. 2010, p. 17). Not surprisingly, as the main influence to performance measurement in most sectors originate from the for-profit sector, the production model consists of a linear chain of inputs, activities, and outputs, adequate to grasp manufacturing processes in their entirety (see also .

Figure 2).

Along this chain, performance can be measured from three logical angles:

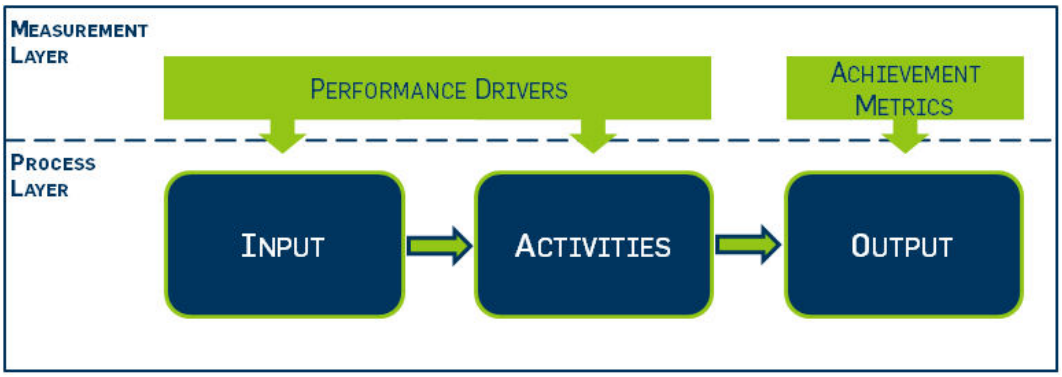
- Effectiveness: *"To what extent are previously formulated goals met?"* (Bedrup 1995a, p. 85)
- Efficiency: *"How economically are the resources of the organizational unit utilized in order to meet the formulated goals?"* (ibid.)
- Changeability: *"To what extent is the organizational unit able to react to changing short-term, mid-term, and long-term requirements from the outside?"* (ibid.).

Notably, the reference to *previously formulated goals* in the effectiveness and efficiency angles acknowledges the importance of the linkage of performance measurement to an organization's objectives and the changeability angle stresses the relevance of an organization's environment for performance measurement (see section 2.2).

Based on these ideas of the production model and the three angles to performance, a set of performance indicators (also 'key performance indicators' – KPIs) can be developed. *"KPIs can be described as metrics of an organization, an*

organizational unit or a process, which display factors that are vitally relevant for the present or the future success of the organization" (Hoffmann 1999, p. 103). KPIs may be quantitative or qualitative metrics, financial or operational metrics, achievement metrics or performance drivers. As parts of a hierarchically organized set, they may be aggregated, correlated, subordinated, and prioritized, in order to reflect the comparative importance of different metrics for the intended purpose of performance measurement, but they do not necessarily stand in a straight mathematical connection to each other. Figure 2 positions achievement metrics and performance drivers along the production model.

Figure 2: Production model and KPI types



Source: Own composition

Richert confirms that it is advisable to employ a variety of different KPIs: he demands a balanced system of performance drivers and achievement metrics (2006, pp. 31ff.). Furthermore, he refers to Klingebiel, in postulating the "*development of specific influenceable areas of performance for all business units instead of overall monetary metrics"* and

- "the connection of long-term, and short-term metrics,
- the connection of monetary and non-monetary metrics and
- the selection of early indicators as a focal point" (2000, pp. 1ff.).

According to Gladen (2005, p. 11), KPIs furthermore have to:

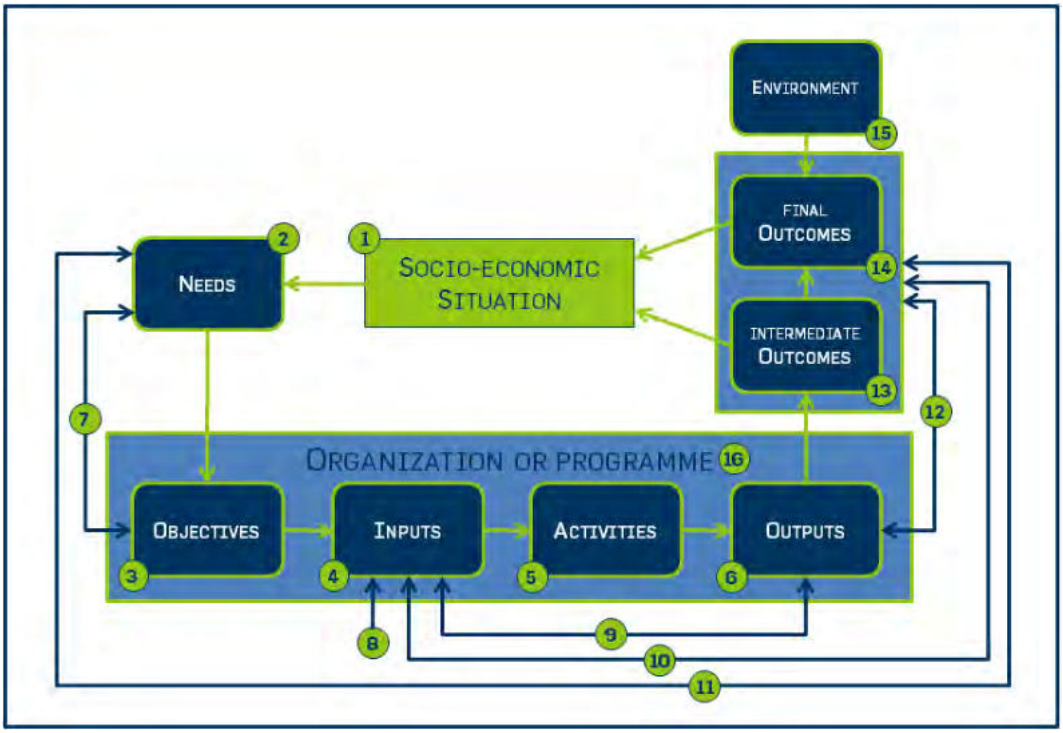
- display complex operational issues, structures and processes in a simple kind of way in order to ensure a quick but capacious overview,
- provide managing authorities with a basis for (occasional or continuous) analysis, and
- serve managing authorities in continuous planning, enforcement and monitoring by eliminating irrelevant data.

The *production model* falls short of serving the purposes of public and non-profit organizations, as it neglects the extended necessity within these two sectors to incorporate socio-economic needs into the equation. Bouckaert and Halligan have extended what they refer to as the *span of performance* beyond the conventional scope of the production model (2008, pp. 15–18). Figure 3

illustrates this extended process model, also hinting at KPIs that establish relations between several of the incorporated parameters.

Like in the public sector, actors within the private sector base their strategy on the socio-economic situation (1 in Figure 3) of a defined market. What differentiates these two sectors from each other, however, is the fact that the feedback mechanism about the value or acceptance of an output (6 - produced services or goods) is measurable more clearly in the private sector: the sum of accumulated number of individual purchasing decisions is the direct indication for a company, to what extent their goods or services meet the needs of society. In the public and non-profit sector, the linkage between output and outcome (13 and 14 - the added value of a good or service) is comparably looser. The feedback mechanism about the value and acceptance of public goods and services consist of a rather vague elections-based mechanism (and this only in democratic societies). Public debate between civil society, the media and public actors may give some indication on the quality of some specific goods and services provided, but the correlation of election successes and failures with specific societal issues is nearly impossible to identify. An additional layer of complexity lies in the constituency of goods and services (Bouckaert/Halligan 2008, p. 19): only parts of those outputs created by public organizations are directly consumable (e.g. subsidy, information, etc.). Other types of outputs are increases in availability of goods (e.g. infrastructure, water purification), or undividable public goods (e.g. legal provisions), all of which pose serious challenges to the measurability of the associated outcome.

Figure 3: The production model of performance; numbers are elaborated below



Source: Adapted from Van Dooren et al. 2010, fig.2.1

Accordingly, the needs for action (2) that arise out of a socio-economic situation are an *interpretation* of the *actual societal needs*. They are defined by

politicians, under influence of other actors such as the media, civil society, and other interest groups. Based on these identified needs, the objectives (3) for an organization or a program are formulated. The conversion of these objectives into outputs (6) requires inputs (4 - material or financial and human resources) and activities (5 - the utilization of inputs over time). Through the already described 'black box' between outputs and outcomes (Bouckaert and Halligan call this a 'grand canyon' (2008, p. 17)), goods and services generate (vaguely measureable) value within society. Van Dooren et al. distinguish between intermediate outcomes and final outcomes (2010, pp. 18–21). Intermediate outcomes are usually generated in the short-term, whereas final outcomes are long-term effects that are under influence of the environment (15) which represents *"socio-economic or ecological trends, but also policy measures from other governments"* (Van Dooren et al. 2010, p. 21).

All these parameters are the theoretical background for what performance measurement could and should encompass in humanitarian aid. If they are systematically put in relation to each other, concrete KPIs can be derived, according to Van Dooren et al. (2010, pp. 18–21):

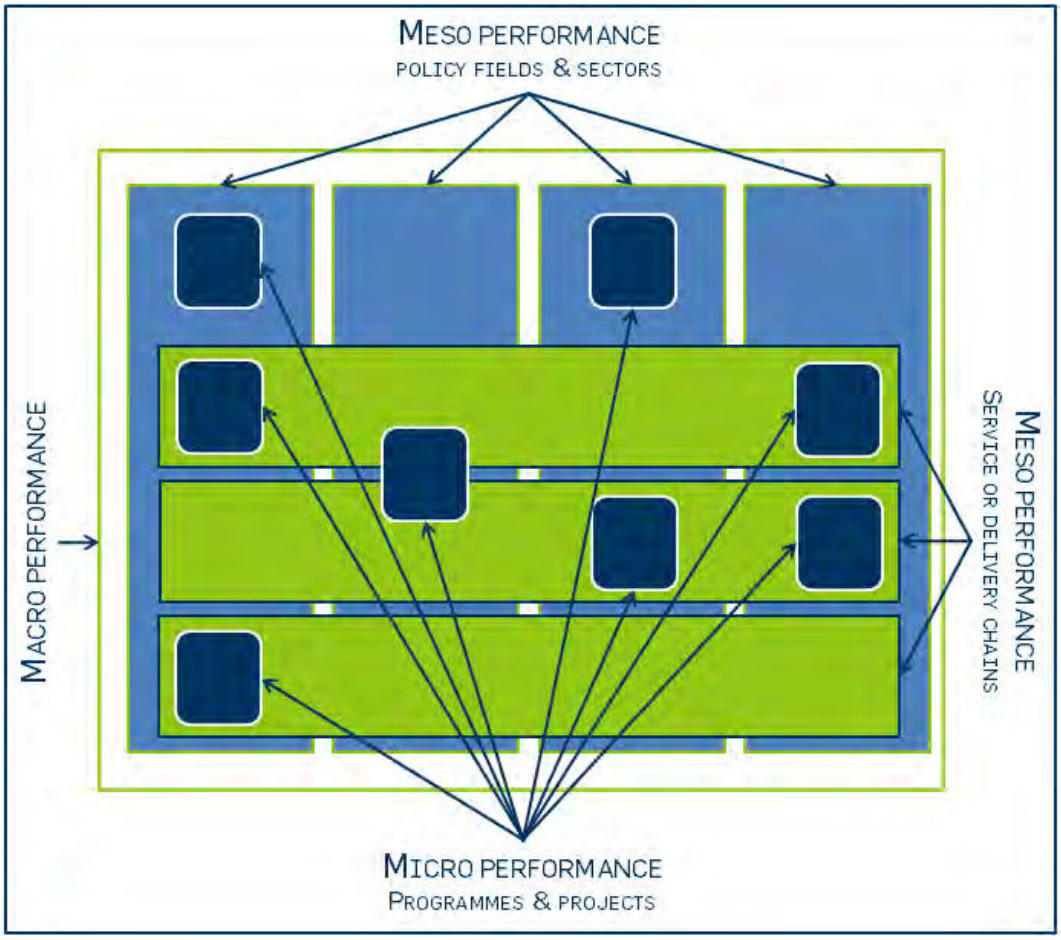
- "(7) Relevance: objectives over needs; to what extent does a set objective correspond with an identified need?
- (8) Economy: input over alternative input; how well selected are the inputs for an organization or program in terms of their cost?
- (9) Efficiency: outputs over inputs; how successfully are the given inputs utilized to achieve a maximum output (output efficiency)/ how successfully are the inputs minimized to achieve a required output (input efficiency)?
- (10) Cost-effectiveness: outcomes over inputs; how successfully are the given inputs utilized to achieve a maximum outcome/how successfully are the inputs minimized to achieve a required outcome?
- (11) Utility and sustainability: outcomes over needs; how successfully were the identified needs served by an organization or program?
- (12) Effectiveness: outcomes over outputs: how successfully do the outputs of an organization or program generate usable outcomes in society?"

When replacing the *production model* with this comparably complex model of performance in the public sector, the three logical angles of performance measurement remain. Like the production model for the private sector, the extended model of Van Dooren et al. applies the logic of measuring effectiveness and efficiency. The aspect of changeability is not explicitly mentioned in relation to this model, but is mentioned in another instance: "*good indicators are sensitive to change"* (2010, p. 60).

A further dimension to performance measurement is the concept of the *depth of performance*, as defined by Bouckaert and Halligan (2008, pp. 18–26). It is the vertical counterpart to the *span of performance*, if the latter is understood as the

horizontal extent of the delivery process of goods and services. The *depth of performance* encompasses three levels of performance: the micro-level refers to single organizations, the meso-level encompasses *"consistent policies or public enterprises in specific policy fields"* (ibid., p. 18), and the macro-level is the one of "*government-wide or governance wide policies"* (ibid.).

Figure 4: Macro-, meso-, and micro performance in a complex nested configuration



Source: Adapted from Van Dooren et al. 2010, fig.2.4

Bouckaert and Halligan's model is related to the public sector alone. Unlike in the public sector, operations in the non-profit sector are more commonly based on program-based and project-based work (especially in the humanitarian sector). An adaptation of the depth of performance that fits both the public and the non-profit sector at the same time would define the micro-level as the one of projects and programs, the meso-level as those projects and programs within *"particular policy fields, specific sectors, or specific service or delivery chains"* (Ramalingam et al. 2009, p. 23), and the macro-level as the level of *"entire governance systems or cross-sector-wide approaches"* (ibid.). Figure 4 further illustrates the levels according to the concept of the depth of performance.

While this section on *performance measurement* provided answers to the question: *what is the content of performance management;* the following section investigates on the *how* and *when*. It provides insight into the operationalization of the performance measurement features into a self-contained system that

allows learning and improvement as well as accountability towards external stakeholders, based on comprehensive performance measures.

2.4 Characteristics of performance management

Performance measurement alone would be static and singular. Incorporation into information and management systems is necessary in order to base decision making on the information contained in a performance measurement system. Otherwise, performance measurement would not go beyond a one-time inspection of the status quo. The formal goals of performance management follow a twin purpose; management improvement (learning and innovation) and performance reporting (accountability and transparency) (Ramalingam et al. 2009, p. 25). Thus, the priorities of the outcomes of a performance management system, depending on the context it operates in, may have primarily internal (management improvement) or external (accountability and transparency) purposes for a specific organizational unit (program/project, organization, sector, etc.). If a sector or organization is, for instance, largely reliant on external funding, the accompanying need for external reporting will dominate the priorities of its performance measurement system. The internal perspective is further described by Van Dooren et al., who put forward three levels of incorporation of performance information (2010, p. 7):

- *single loop learning* is described as the usage of benchmarking and bench learning to upgrade systems to specific standards. Single-loop learning seems to be present when goals, values, frameworks and, to a significant extent, strategies are taken for granted,
- *double loop learning*, in contrast, represents the contemplative questioning of these goals, values, frameworks, and strategies, and
- *meta learning* means to *"adjust systems constantly and learning how to learn"* (ibid.) and represents the most advanced of these three levels.

Whilst Van Dooren et al. themselves describe the performance management as the sequence of "*collecting, integration of data into the management systems […] and putting information at work"* (ibid., p. 6)[5], starting with the *collecting* of data, Klingebiel identifies the following three elements of performance management (1998, pp. 1ff.; 2000, pp. 36ff.), pointing out the importance of a targeting process as a starting point for performance management:

- the process of setting targets,
- monitoring of the development of performance (performance measurement),

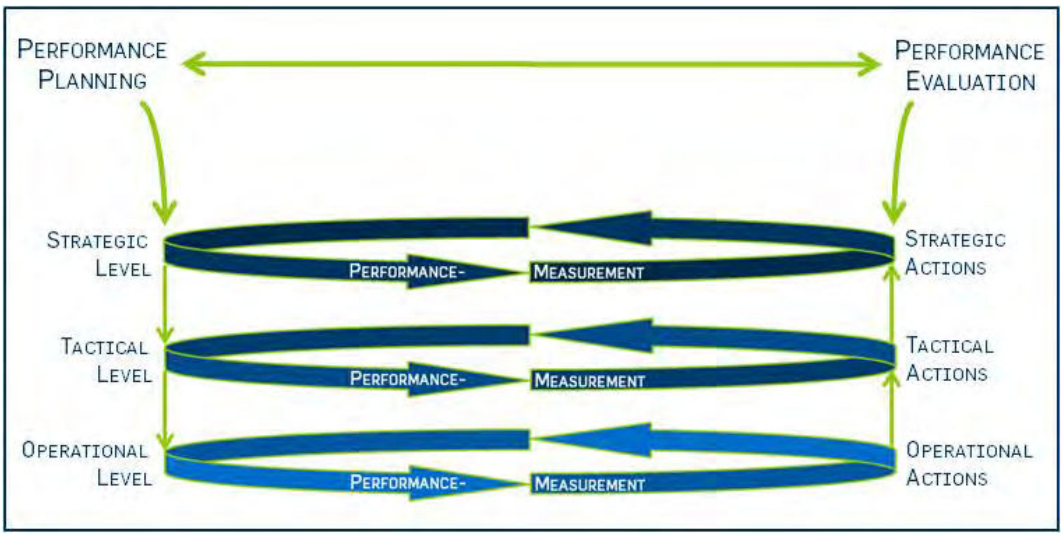
[5] Even though the process of designing a set of indicators and compiling relevant reports from it is, in opposition to the logic applied in this paper, considered a part of performance measurement by Van Dooren et al., they implicitly confirm that the baseline for designing a performance management system lie within a mapping of what is to be achieved and therefore to be measured (2010, pp. 56–60).

- evaluation of reactions (definition of measures for the case of falling short of the set target).[6]

In accordance with the concept of the *depth of performance* (see section 2.3 Characteristics of performance management), the objectives can be anchored in the micro, meso, and macro-level of performance. Depending on the intended level of depth, the following indications can help with the outlining of a performance management system: amongst others, organizational charts, trees of objectives, stakeholder analyses, and program logics can serve as a frame for targeting performance measures (Van Dooren et al. 2010, p. 57).

It becomes clear that the connection of performance management to a strategic process has to be in place before data is collected. This will help ensuring that a set of indicators actually produces meaningful information on the cornerstones of the underlying objectives of the organization, the program or project, the sector, etc. Moreover, performance has to be managed at different levels. These levels are further illustrated by Bedrup (see Figure 5), stipulating a *closed loop process*, encompassing the performance measures strategic, tactical, and operational levels and linking performance planning, performance evaluation, and according actions.

Figure 5: Closed loop of performance management



Source: Own composition, based on Bedrup 1995b, fig.8.3

Bedrup's model is to be understood as a cyclical model, in which performance planning and performance evaluation occur continuously, taking into account performance measures at all levels, and deriving necessary action, both for change in operations and change in the way performance is measured. This model can be understood as a representation of meta learning. It must be kept in mind that this model is rather generic: the concrete levels of performance measurement vary, depending on whether the point of departure for performance management is on the macro, the meso, or the micro-level. If, for instance, a performance management system would be designed for an entire

[6] Taking into account the idea of meta-learning, the evaluation of reactions would incorporate the option of altering goals, values, frameworks and strategies.

policy sector (meso-level), the operational level of performance for the sector level would be corresponding with the strategic level of performance for organizations within the sector.

3. The Balanced Scorecard

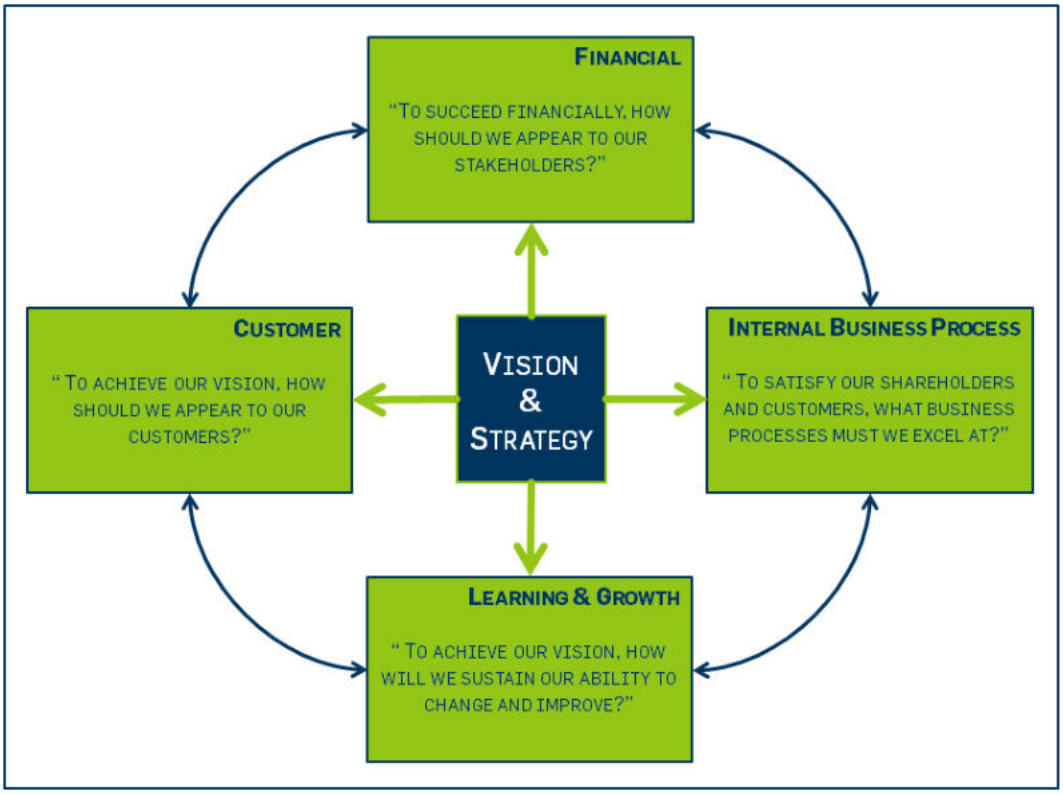
This chapter explores the basic concept of the BSC and the accompanying methodology that extends the BSC from a performance measurement tool to a strategic performance management system that helps turning company strategies into operational success. Subsequently, its diffusion into business practice is expounded and further developments of this concept, which was originally designed to serve single legal entities within the for-profit sector, are discussed: the adaptation to cross-company networks in the for-profit sectors and the adaptation to the public and non-profit sector.

3.1 The Balanced Scorecard in the for-profit sector

Trying to live up to the growing need of management models that provide a wider angle beyond solely financial figures as the ultimate measure of success, companies welcomed models such as the BSC (see also section 2.2). Along with the increasing awareness to perceive a company's objectives and its environment as conditions that determine performance, the rationale behind this model was the notion that a company's ability to mobilize and utilize its fixed and intangible assets has become increasingly more important than the investment in fixed assets and the their coordination (Kaplan/Norton 1997, pp. 2f.). Amongst other characteristics, intangible assets enable companies to "*introduce innovative products and services that are expected by customers"*, and to *"mobilize the abilities and the motivation of employees for a continuous improvement of business processes, quality and reaction times"* (Kaplan/Norton 1997, p. 3). Incorporating several of the characteristics of the performance management process described in section 2.4, the demand for a fairly balanced set of KPIs was taken up by Kaplan and Norton and transferred into an instrument that is both a performance measurement system and a management system for the implementation of a company's strategy (Weber/Schäffer 1998, p. 7).

More specifically, Kaplan and Norton's generic BSC stipulates four "perspectives" on a company's vision and strategy (see Figure 6). As a compound, these perspectives should cover both the short-term adherence of targets as well as the outcome of corporate activities (through the financial perspective), and the long-term-oriented observation of performance drivers as well as the development of intangible assets (through the other perspectives). These perspectives of the BSC have to be considered as a template rather than an imperative model. Therefore, the substance of the perspectives may be changed, depending on a company's or an organization's strategic priorities. Perspectives may be added or removed. Stakeholders that are vital to a company's success may be reflected in a separate perspective (Kaplan/Norton 1997, p. 33).

Figure 6: The Balanced Scorecard as a framework to translate strategy into action



Source: Own composition, based on Kaplan/Norton 1996, fig.1-1

Along with the fanned out view on a company's strategy, a procedure to put the BSC into practice is explained by Kaplan and Norton: in an integrated process that complies with the idea of a closed loop process (see section 2.4) the BSC urges a company to find a consensus about the corporate vision and strategy as a first step. As a next step, specific goals, formulated in a way, explicitly expressing quality, time, and service related aspects, are assigned to each of the perspectives of the BSC. The derived goals have to be translated into KPIs, communicated to the divisions and departments, and connected to incentives. As a further step, benchmarks must be set, strategic measures defined, resources disposed, and milestones fixed. Lastly, strategic feedback has to be provided in order to evaluate the success of the previously formulated strategy and, if necessary, to re-modulate it (ibid., p. 10).

The BSC requires the conjunction of goals with metrics, especially the conjunction of goals of one perspective and metrics, with those of the other perspectives. The goals and metrics are interrelated through cause-and-effect-chains (ibid., pp. 28ff.), also referred to as a strategy map. Through the BSC's cause-and-effect methodology, one observes the interrelations between its goals and metrics, so that these can be monitored and controlled (ibid., p. 144). Hereby, the traceability of reasons for success or failure is possible. Likewise, an early warning system can be provided by the incorporation of early indicators (ibid., p. 28). According to Horváth and Kaufmann, the cause-and-effect-chains are to be aligned through qualitative and subjective appraisal due to little

cognition in scientific research concerning methods for setting up cause-and-effect-chains (1998, p. 48).

Since the introduction of the concept in the early 1990s, a clear trend can be ascertained. Observing the levels acceptance and success of the BSC in businesses, one finds that *"its diffusion was so rapid that as early as 1997, it was labeled as one of the most influential management instruments of the 20th century"* (Sibbet 1997, p. 12 cited in De Geuser et al. 2009, p. 93). Studies that examine its diffusion in quantitative figures over time found that it still appears to follow an expansive course. Bain and Company found that in 2005, out of a sample of 960 international companies, 57 per cent had had a BSC in place (Rigby/Bilodeau 2005, p. 13), while in 2007, this figure picked up to 66 per cent (out of a sample of 1,221 companies) (Rigby/Bilodeau 2007, p. 14). Moreover, De Geuser et al. carried out empirical research on the level of positive impact on organizational performance that could be attributed to the use of the BSC (2009). Based on a sample of 76 questionnaires from 24 companies, collected in 2000, they concluded that the Balanced Scorecard does 'add value' to organizational performance management (ibid., pp. 116f.). Although the diffusion is certainly easier to evaluate than to draw conclusion on the benefits, the BSCs very quickly had an important impact on strategic performance management in the for-profit sector and thus led to subsequent efforts of extending, modifying and adapting it for broader or other purposes.

3.2 The Balanced Scorecard in cross-company networks

Since the 1990s, companies have increasingly begun focusing their core businesses. Functions that were considered inefficient were outsourced to other companies to narrow business processes and to scrap support functions. These outsourcing activities led to more vertically expanded, geographically scattered supply chains, the coordination of which was no longer the task of a single company and its suppliers, but rather of polycentric and legally separate entities. The objectives of these multi-tier supply chains nevertheless remained the same. Products and services of good and consistent quality were to be delivered on time and at the lowest possible cost. But the new challenge was to plan, administer, and supply under the constraint of lead times and uncertain customer demands, and the accompanying danger of the bullwhip-effect[7] across an increased number of organizational boundaries along a supply chain. Considering that one company seldom is part of a supply 'chain', but much rather of a supply network, with one or more 'partners' up and down the 'chain', companies understood that, primarily with selected business partners, closer strategic collaboration would be worthwile in order to live up to the new requirements of rising complexity in the coordination of their supply networks.

[7] Supplier lead times for components of a specific end-product demand of companies to forecast an uncertain customer demand of the end-product in the future. As this process takes place at every tier of a supply chain, uncertainty about end-customer demand procreates and multiplies backwards along the supply chain, leading to excess stock or unfulfilled extra demands. For more information, see Lee et al. 1997.

Complying with the notion that *"competition will no longer take place between single companies but rather between competing supply chains"* (Zimmermann 2003, p. 1), cooperation initiatives such as Collaborative Planning, Forecasting and Replenishment as well as Vendor Managed Inventory became new recipes. To embed these collaboration initiatives into a broader frame of strategic alignment, scholars extended the concept of the BSC for it to be suitable to cross-company scenarios.[8] They propose various means to modify Kaplan and Norton's generic BSC. The overall BSC methodology remained roughly the same, but naturally became more complex, due to the increased number of companies whose strategies would have to be reflected in a cross-company BSC system. The methodology would require of the involved companies to initiate a strategic alignment process on the supply network level and to cascade the agreed, common strategy from supply network level down into the companies.

Two important challenges were taken up. First, it was necessary to structurally align the multitude of individual companies' objectives and KPIs towards a homogenous system. Secondly, as cooperation between the companies naturally involves (at least) two actors, the BSC model system would need to be modified into something different than a company scorecard.

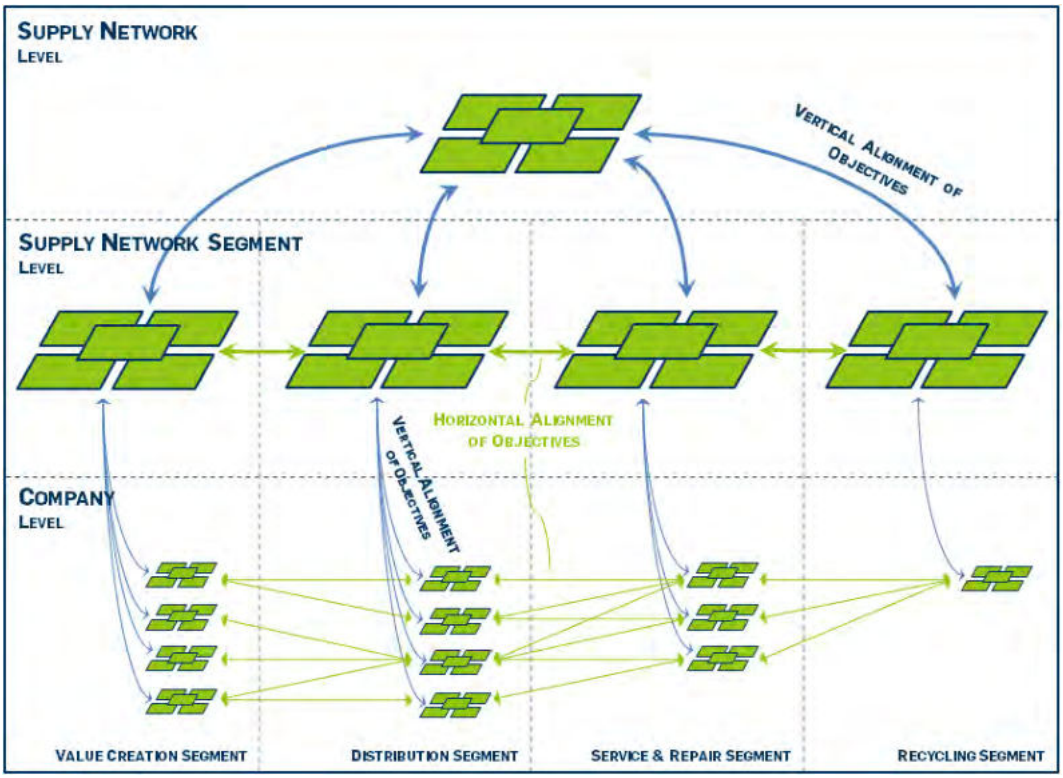
Erdmann meets these challenges by introducing a network scorecard that reflects the common objectives of the supply chain partners beyond the company scorecards. It reflects individual companies' objectives, while proposing vertical and horizontal alignment of objectives between and within these layers (2003, p. 181). It is notable that Erdmann inserts segment scorecards in between the supply network scorecard and the company scorecards (ibid., p. 179). He substantiates this idea by stating that segment scorecards allow benchmarks of supply chain functions with competing supply chains. Furthermore, he proposes the insertion of cooperation perspectives into both the network scorecard and the company scorecards.

Richert follows a similar approach as Erdmann. He supports adding a cooperation perspective and also proposes an inter-organizational and intra-organizational scorecard (2006, p. 78). Moreover, he proposes prerequisites to implementing a network-wide BSC:

- "A strategy aligned between the SC partners is existent and accepted;
- The management personnel is able to understand the cause-and-effect-interrelations of the key metrics and strategic goals;
- All supply chain partners have already implemented a company BSC to ensure that all companies have their own strategies and goals (thus, conflicts between goals can be identified at an early stage) and have an appropriate skill-level in dealing with the complex dynamics of the BSC" (ibid.).

[8] Amongst others, the following scholars developed concepts that apply the balanced scorecard to cross-company scenarios: Brewer/Speh 2000; Erdmann 2003; Lange et al. 2001; Richert 2006; Stölzle et al. 2001; J. Weber et al. 2002; Werner 2000a; Werner 2000b; Zimmermann 2003.

Figure 7: Concept for a Balanced Scorecard system for supply networks



Source: Own composition, partly based on Erdmann 2003 and Richert 2006

Figure 7 summarizes the ideas of Erdmann and Richert and illustrates the depth dimension[9] of a network-wide BSC system. Even though it could be interpreted in a way that a BSC system for a supply network is thinkable including a variable number of companies from each of the segments, it only displays the *theoretical* interconnections much rather thank the *probable* ones. In practice, this full-blown case, in which *all* companies within a business sector would join a sector-wide BSC initiative out of their own interest, is very improbable. As companies seek to gain competitive advantages over within their business, it is highly improbable that they would cooperate with competing companies *from the same segment* out of free will. Factually, BSC systems would, depending on actual market shares, comparative bargaining power, and the structure of the business sector, be developed around one focal company and its suppliers and customers up and down its various supply chains. In that sense, companies of the same segment would indeed be part of the same BSC system, but only as members of the supply network of *another* company that has the bargaining power to coerce its partners into a certain mode of business relations.

3.3 The Balanced Scorecard in the public and non-profit sector

Following the rapid diffusion amongst business companies, attempts emerged to modify what seemed to be successful in the for-profit sector for other sectors. While, according to Johnsen, in 2001 implementations of the BSC were already

[9] For the definition of the depth of performance measurement, see section 2.3.

"widely diffused in business" (2001, p. 319), he also points out that is has arrived in public management *"probably also to some extent"* (ibid.). His article on the applicability of the BSC model to public management, together with the research of Kloot and Martin on the BSC's applicability for management issues in local government (2000), Kaplan's ideas on strategic performance measurement and management in non-profit organizations (2001), Moore's public value scorecard (2003), and Niven's BSC for government and non-profit agencies (2008) constitute some of the first approaches towards adapting the BSC to the public sector and the non-profit sector. To transpose the provisions of the BSC model onto other sectors than the for-profit sector is certainly not done by replacing the word *company* by the word *organization*. In fact, the first noticeable approximation between the BSC and the non-profit sector can be attributed to Kaplan himself. His 'research agenda' and the related practical efforts to introduce the model to a number of non-profit organizations, date back to as early as 1996 (2001, p. 357). While upholding the overall BSC methodology, he proposes major adaptations for the structure of the non-profit model.

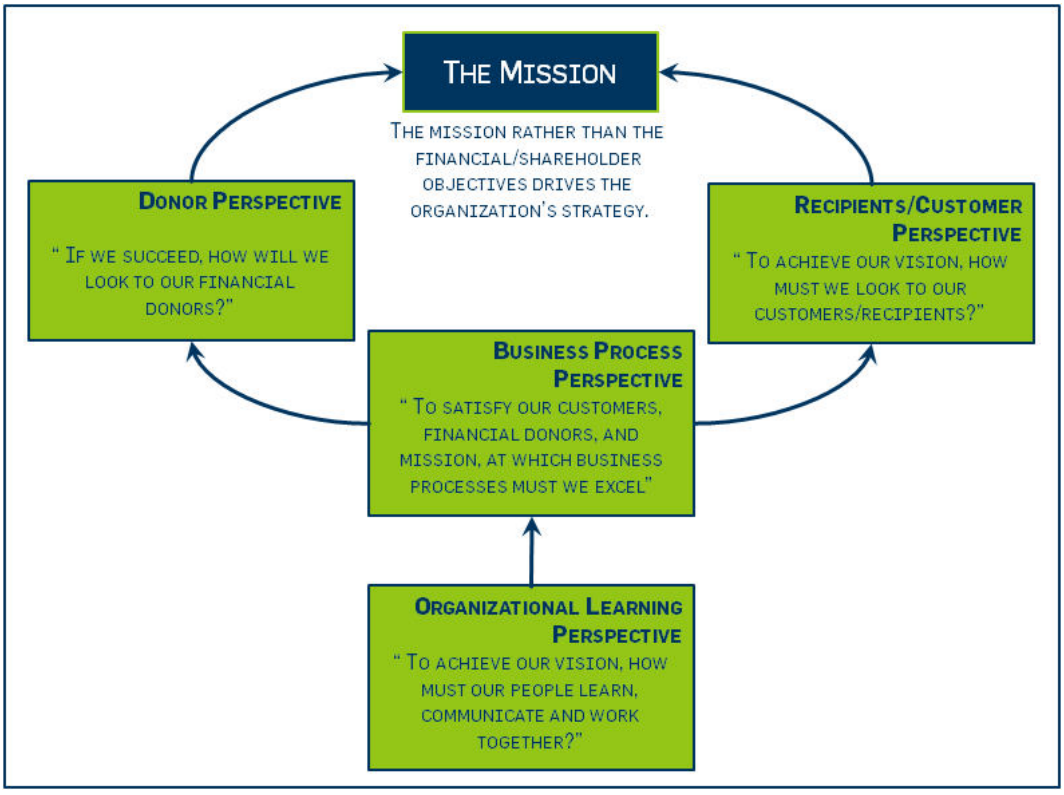
First, what is the ultimate purpose for a private company certainly cannot be applied to a non-profit organization. Businesses strive for the long-term maximization of shareholder value, represented by the financial perspective being placed at the top of the hierarchy of the BSC. In contrast, non-profits seek the fulfillment of a chosen social objective, in this sense the maximization of the social value. Therefore, an organization's mission *"represents the accountability between it and society – the rationale for its existence"* (Kaplan 2001, p. 360). Consequently, the organization's mission is placed, together with the customer perspective, at the top of the hierarchy. All other perspectives are subordinated under the mission as they are meant to contribute to its achievement.

Secondly, Kaplan proposes a modification of the definition of the customer (ibid., pp. 360f.). Whilst company-customer transactions inherently feature the customer as the one who pays *and* receives the service or product, a comparable transaction in the non-profit world is fundamentally different. By providing financial resources, donors take over a part of what is the customer's role in the for-profit world. Complying with the observation that *"stakeholders that are vital to a company's success may be reflected in a separate perspective"* (see section 3.1), the simple transformation of this scenario into the non-profit BSC could be achieved by using two separate perspectives; the donor perspective and the customer/recipient perspective. Again, even more than the for-profit model, Kaplan's non-profit BSC is not an imperative model. Kaplan even abstains from naming the perspectives that he describes (2001, fig. 2). Figure 8 illustrates Kaplan's suggestion for an adapted BSC, complemented with names for the provided perspectives.

The ideas of Niven (2008) are more elaborate than Kaplan's and they fill an important gap that was left by Kaplan, because they explicitly encompass the public sector. Like Kaplan, Niven elevates the mission and the customer perspective to the top of the hierarchy, pointing out the challenge to identify the customer. He illustrates this through a very colorful example of a law

enforcement agency that might consider the criminals they arrest its customers (ibid., p. 34). While governmental agencies represent a part of the customer perspective in Kaplan's non-profit model, the very same question arises in the public model: who are the customers of governmental agencies? One could argue that a state's citizens might take that role, or that in case of services addressed to recipient outside of a government's national boundaries, the beneficiaries of these are the customers. The latter would again pose the challenge of a split customer role. External services received by another group than the citizens of a country are not identical with the group of people financing these; the tax payers.

Figure 8: Adapting the Balanced Scorecard framework to non-profit organizations



Source: Own composistion, adapted from Kaplan 2001, fig. 2, complemented with names for Kaplan's perspectives

Instead of delivering a formulated recommendation for this difficulty, Niven provides two guiding questions that should help narrowing down the list of stakeholder to one identifiable customer. First, a customer could be *"a person or a group that directly benefits from […] products or services"* (ibid., p. 167) and secondly, a customer could be, whose disappearance would cause an organization *"to be irrelevant or unnecessary"* (ibid.).

While Kaplan excludes an explicit financial perspective from his generic non-profit model, in the implementations of the BSC in non-profit organizations, the financial perspective remains relevant (2001, pp. 361–368). Nevertheless, the measures he associates with the financial perspective are inherently no financial measures in the sense of the for-profit BSC, as they are concerned with fundraising, budget disposable for projects, etc. Therefore they might as well be

integrated into the process perspective. In contrast, Niven claims that *"no balanced scorecard is complete without a financial perspective"* (Niven 2008, p. 34). Taking a closer look at the argumentation behind this imperative claim, one finds, however, that he derives the necessity for a financial perspective from the importance of financial means as the *"enablers of customer success"* or *"constraints within which the group must operate"* (ibid.). These two points are likely to refer to efficient use of funds, or the amount of raised funds. Again, these measures could be modeled into the (business) process perspective, as they are much rather performance drivers that enable the actual overarching objective (to create social value) than financial measures that reflect the level of achievement of the overarching financial objective, as they do in a for-profit BSC. Despite different interpretations on the customer question (that is in parts due to the broader scope of Niven's model, including the public sector) and inclusion or exclusion of an explicit financial perspective, Kaplan and Niven's concepts show a high level of congruency.

3.4 The remaining research gap

Through chapters 2 and 3, which provided the background on the concepts of performance, performance measurement and performance management, comprehensive theoretical reference has been established, to investigate on and assess the state of performance management in humanitarian aid in the course of the next chapter. Furthermore, the for-profit BSC model and especially the research on its extendibility across organizational boundaries and adaptability to the public non-profit sector are the potential baseline for conceptualizing a concrete BSC model for the humanitarian sector. In essence, what distinguishes the BSC from an overall approach to performance management (as presented in chapter 2) is its multidimensional view on performance that can reflect the essence of a company's core business and its *"notion of cause-and-effect – how the measures link together to tell a strategic story"* (Niven 2008, pp. 25f.). The generic for-profit model has been subject to further research by two separate schools. The previous chapter shows how:

1. the perspectives of the BSC can be exchanged, modified or added to suit the purpose of sectors with fundamentally different frame conditions, while the accompanying methodology remains fairly stable. The adaptability to the public and non-profit sector has been discussed explicitly. It appears that, based on the elucidated concepts, the BSC concept could provide a basis for the application to humanitarian organizations, programs, and projects;
2. by exploring research on the extension of the BSC model beyond organizational boundaries in the for-profit sector, a concept has been introduced that might be applicable to the humanitarian sector.

Based on these preliminary results, the further course of this research assesses the special characteristics of humanitarian aid, identifies issues in humanitarian performance, and derives preconditions for effective performance management. Following this, the paper seeks to validate in how far a sector-

wide BSC is transformable into a system-wide humanitarian performance management model, and under which precondition this might be the case.

4. The state of performance in humanitarian practice

This chapter develops an understanding about the rise of initiatives within humanitarian aid that can be attributed to growing awareness about the necessity of performance management. As this chapter shows, different reasons caused the emergence of a variety of tools, frameworks, and concepts. Therefore, the following sections first describe important performance initiatives and put them into context before the background of performance management theory that was discussed in chapter 2. Secondly, existing deficiencies within the conglomerate of performance management approaches are pointed out, leading over to the following chapter in which the requirements for a sector-wide humanitarian performance management system are discussed.

4.1 Performance management approaches in humanitarian aid

As opposed to the for-profit sector that coerces companies to constantly improve their organizational capacity through its inherent competitive mechanism, striving for improving and managing performance in humanitarian aid has not been and still is not self-evident. Referring to both the public and the non-profit sector, Niven points out that *"what has been lacking is the answer to the seemingly simple question: Is what we are doing making a difference? Is anyone better off as a result of our efforts?"* (Niven 2008, p. 31).

As pointed out in sections 2.3 and 3.3, the link between public agencies and the recipients of their services is already a loose one. Elections, media coverage, and articulated public interest are a few of the feedback mechanisms that make up the accountability link between the 'customers' of a public service and its provider. This link is even looser in humanitarian aid. No structural accountability demands of public or non-profit agencies to improve their performance. Public interest and media coverage fill this gap to some extent. However, the core of humanitarian aid is still factually based on the foundation of charity and voluntary self-commitment. For long, the notion that 'doing good' is good enough was prominent for aid agencies, public, and private donors. As Harrel-Bond points out:

"Humanitarian work […] is thought to be selfless, motivated by compassion, and by its very definition suggests good work. Most voluntary agencies place as great an emphasis upon the motivations of their employees as upon their technical expertise… As relief is a gift, it is not expected that anyone (most especially the recipients) should examine the quality or quantity of what is given" (1986, p. XII).

But eventually, *"the motivation to 'do better' has come both from within the humanitarian agencies as well as from pressure exerted by donors and the media"* (Griekspoor/Sondorp 2001, p. 209). The 1990s constituted a major turning point for the fundamental assumption that humanitarian aid was

naturally doing good. Confronted with the changing nature of conflicts after the Cold War and the resulting lack of legal measures to access and help victims of internal conflicts, the humanitarian sector, facing the dead faint of the Rwandan genocide, successively realized that *"once political failure led to the crisis, many more lives could have been saved had humanitarian organizations better coordinated and acted more professionally"* (Hilhorst 2002, p. 194). In its staggering report, mentioning at least 100,000 casualties, the Joint Evaluation of Emergency Assistance to Rwanda concluded that *"the current mechanisms for ensuring that NGOs adhere to certain professional standards are inadequate"* (Eriksson et al. 1996, p. 29) and that *"the present accountability mechanisms within the humanitarian aid system are quite inadequate"* (ibid., p. 30).

During the same decade, from 1990 to 2000 the official budgets for humanitarian aid increased from 2.1 billion United States Dollars (USD) to 5.9 billion USD (Macrae et al. 2002, p. 3), in parts to respond to numerous major humanitarian crises, for instance in the Democratic Republic of the Congo, the Former Republic of Yugoslavia, the Kosovo, Somalia, Timor-Leste, Ethiopia, and Eritrea. These growing global needs for humanitarian aid, along with the recognition that humanitarian aid can be used as a political tool, *"donors increasingly asked for value for money and evidence-based interventions"* (Griekspoor/Sondorp 2001, p. 210). They decided that the delivery of humanitarian aid through implementing organizations should become more transparent.

This recognition coincided with widespread reforms of the public sectors in the West. From the 1980s on, NPM introduced management practice ideas, borrowed from the for-profit sector. Before the background of realigning inefficient public structures around concepts such as market-orientation, lean management, customer orientation, and state-of-the-art controlling, NPM demanded a results-based view on public spending. These reforms did not spare the humanitarian sector. Donors and several UN agencies started developing results-based management systems. To support their goal of making transparent how much 'value for money' they receive from implanting agencies, appeals for government funds were tied to features of results-based management. Although it is certainly difficult to draw a line between internal and external motivation, it may be said that not only conditions for funding set by donor agencies made NGOs revise their mode of operations. Initiatives such as the sphere standards, launched in 1997, hint at collective NGO acknowledgement about the need to professionalize.

In consequence, the past fifteen to twenty years have seen a multitude of initiatives that approach the challenge of improving the humanitarian performance from different angles *(types of initiative: what is their purpose?)*, at different scales *(level of depth: which organizational level does the initiative target)*, and with differing scopes *(extent of the span: which parts of the performance process are targeted by the initiative)*. While there are some organizations that oppose the use of standards, as they fear that they *"would undermine necessary flexibility to adjust responses to the local context, or be a*

threat to their independence" (Griekspoor/Sondorp 2001, p. 209), the majority of agencies in the humanitarian sector support the endeavor for improving performance. Performance initiatives can be categorized according to the types of *"conceptual and operational initiatives"*, proposed by Ramalingam et al.:

- *"establishment of accountability through standards and codes*
- *improvement of performance through information and management techniques such as needs assessment, project-cycle management, and evaluations*
- *emphasis of local capacities, linking relief to development, and disaster-risk reduction*
- *focus on the rights of affected people and protection"* (2009, p. 49).

Further typologies are, for instance, provided by Niven and Hilhorst. Niven proposes financial accountability, program products or outputs, adherence to standards of quality in service delivery, participant related measures, KPIs, and client satisfaction as types of performance-related priorities (2008, p. 29). Hilhorst, categorizes initiatives who on a more abstract level, according to the organizational management approach, the rights approach, the contingency approach, and the ownership approach, adding the topics of standards, accountability, evaluation and monitoring (2002). While Niven's typology certainly encompasses most aspects relevant to humanitarian aid, it does not suffice in explicitly referring to issues such as the rights-based approach or disaster-risk reduction. Hilhorst's typology is more complete in this sense, but it does not have a clear enough structure, compared to the typology of Ramalingam et al. Their typology will guide the identification and contextualization of performance management approaches and initiatives in the following.

In the further course of this chapter, important initiatives will be outlined and analyzed before the background of the theoretical aspects of performance management provided in chapter 2. They are grouped along the micro, meso, macro-logic, provided by Bouckaert and Halligan (2008, pp. 18–26). The grouping into the levels is conducted according to the *application level* of each initiative: even if standards or management techniques are for instance developed by a consortium of NGOs, as long as the application of the performance management initiative is under the coordination and responsibility of each of the NGOs individually, it is still to be considered a means for performance management at organization-level, and not cross-organizational level.

4.1.1 Performance management at the micro-level

Since the insights of the 1990s, the most self-evident efforts for performance improvement have concerned the core of humanitarian work: the way in which humanitarian aid is delivered. The methodology of the project cycle as a fundamental approach to program and project management has widely diffused into donor agencies' funding application processes and consequently into the

procedural guidelines of NGOs. In an analysis of quality management tools used by the partners of the Europeans Commission's Directorate General for Humanitarian Aid and Civil Protection (ECHO), in 2002 Bugnion found that "between 88.4% and 90.7% of partners use at least one tool for the different stages of the Project Cycle" (2002, p. 55). The project cycle seeks to structurally link together the phases of project management, comprising of situation analysis, response analysis, response planning, response implementation, and monitoring and evaluation (Ramalingam et al. 2009, fig. 55). Since performance management approaches and -initiatives are not equally present and important in each of these phases, it makes sense to distinguish between the phases and the related initiatives and approaches in the following.

Needs Assessment

Needs assessment is the practical dimension of what the humanitarian principle of impartiality formulates as follows:

"Aid is given regardless of the race, creed or nationality of the recipients and without adverse distinction of any kind. Aid priorities are calculated on the basis of need alone" (IFRC/ICRC 2010, p. 3).

Needs assessment takes places whenever a new disaster situation has arisen. Based on needs assessments that seek to grasp situations *"in terms of reasonable consistency"* (Darcy/Hofmann 2003, p. 1) and inform decision making related to the humanitarian response, interventions are planned. In this sense, it can be considered the pre-stage of planning. In contrast to the structure and procedures in public management in general, it is carried out in a more fragmented way, regarding its location, its duration, and the responsible assessors of needs: while public management has a relatively continuous information basis over social needs in a limited geographical context, humanitarian needs assessment is mostly non-recurring, ad-hoc, geographically scattered, and only in parts carried out by the public institution that seeks to finance the fulfillment of the identified needs. Much rather, NGOs that intend to respond to an identified disaster, carry out needs assessments as part of their response planning and as part of their applications for funding.

This fragmented landscape of various agencies conducting needs assessment leads to many truths about the factual underlying needs, making it difficult to ensure the relevance (the extent to which set objectives correspond with actual needs, see section 2.3) of planned interventions in the first place. Secondly, and taking into account that common performance measurement concepts use the identified needs as a baseline for measuring the utility and sustainability (the extent to which the outcomes of a program or project served the identified needs, also see section 2.3) of a program or project,[10] an unnoticed gap possibly remains throughout the whole project cycle, the magnitude of which is mainly and highly dependent on the blur between reality (the socio-economic context) and the

[10] ECHO's uses the criteria of sustainability and impact (equaling outcome) to evaluate project success (ECHO 2005, p.17).

interpretation of reality (needs assessments). This flaw is certainly an inherent dilemma of any project or program that seeks to achieve sustainable outcomes in a complex reality. Nevertheless, especially the humanitarian sector seems to be facing the problem, that *"assessment typically is subsumed within a process of resource mobilisation, with assessments being conducted by agencies in order to substantiate funding proposals to donors"* (Darcy/Hofmann 2003, p. 2). Needs assessment and program- and project planning are not separated clearly enough, implying the danger that perceptions of humanitarian needs are partly the result of a wish list of needs of implementing agencies that in turn compile this wish list according to their competences.

The necessity to approximate the optimum scenario, in which the response analysis is not dictating parts of the situation analysis, has been recognized. In 2004, the World Food Programme (WFP) has launched the Strengthening Emergency Needs Assessment Capacity project in cooperation with the European Commission, an initiative to develop *"improved analytical methods, tools and guidance materials"* and to improve *"the availability and management of pre-crisis information in countries"* (WFP 2006).

Such projects possibly contribute to the precision of assessment information. However, the structural problem described above should much rather be tackled at its core. Truly needs-based assessments can only be achieved if *one* accurate truth is generated, serving implementing agencies as the frame for response planning. Such an overarching effort can be seen in the common humanitarian action plan (CHAP) of the inter-agency standing committee (IASC) that seeks to join efforts in context analysis of humanitarian emergencies. This concept holds the potential to take the needs assessment stage to a cross-organizational level and to narrow the gap between actual needs and identified needs. But, as Ramalingam et al. point out, the CHAP has not (yet) reached the critical mass in participating donors and NGOs (2009, p. 57).

Another frequently discussed point concerns the contextualization of what is sufficient and enough, in the sense of quantitative and qualitative aspects of services, such as food, shelter, water and sanitation, etc. An additional dimension of these considerations is added by changing levels of risk that influences needs in various ways (Darcy/Hofmann 2003, p. 2). This will be further discussed in relation to standards and frameworks below.

Planning

The consecutive step of converting identified needs into objectives and to break these down into manageable activities is the planning stage of a project or program. The management tool used most commonly for this purpose is the logical framework approach (LFA). Following the logic of results-based management, its advantages are seen in thorough methodology, using measurable indicators that help quantify and qualify results and verify whether the set objectives are achieved (Griekspoor/Sondorp 2001, p. 210). Moreover, the systematic procedure that the LFA asks of its users holds the potential to alleviate misinterpretations about project work: if used as a central reference for

communication throughout a project implementation, its comprehensive quantitative and qualitative description of objectives, results, and activities leaves little room for a project team to run off into different directions (Dearden 2001, p. 9). In this sense, it *should* encompass more than solely the planning stage of programs and projects.

In practice, however, it becomes clear that the LFA is indeed often treated as a bureaucratic necessity towards donors, instead of being embedded into the reality of planning, monitoring and evaluation. *"Although using the LFA is theoretically a way of planning interventions and monitoring and evaluating their progress, in practice in most cases the LFA is only explicitly used at the planning stage"* (Bakewell/Garbutt 2005, p. 7). Practitioners report that *"the logframe matrix has been used like a Procrustean bed to force a pre-existing design to conform to a bureaucratic requirement"* (Dearden 2001, p. 5). Other deficiencies lie in the actual (non-)use of the framework's possibilities for monitoring and evaluation. This essentially has two reasons: first, because the factual priority lies in the LFA being a piece of documentation at the beginning of a project that is rarely updated or revisited (ibid.). Secondly, even though the LFA demands data collection and indication of sources for the different levels of the logframe, it leaves no room for documentation of unexpected factors, events, and the incorporation of the two into a revisited project plan. Hereby, its structure is not as flexible as the dynamic environment of humanitarian aid would require it to be. One should not forget, however, that in a dynamic environment, planning inherently opposes retaining flexibility to some extent; an argument used by the opponents of performance management through standardization.

In essence, the LFA may be seen as a good step towards holistic project cycle management. As a template or format it needs room to trace unplanned factors (beyond the already included assumptions and risks). As an approach, the focus should be shifted away from its character of a plain fundraising tool towards a tool that is continuously used for the entire process of project management. Nevertheless, any planning can only be as good as its basic parameters that in turn are defined by the assessment of needs, as discussed above. The CHAP (see above) could provide a basis for the most important link between needs assessment and planning, given that it gains more prominence within the sector.

Monitoring

In theory, the monitoring of programs and projects is the continuous tracking of progress, identification of deficiencies, and improving or adjusting the activities (and if necessary and applicable, the results or objective of the whole program or project). If the planning phase has been carried out with a clear focus on the subsequent monitoring and evaluation activities, monitoring is the consecutive tracing of the program- or project implementation based on pre-defined KPIs.

Along with the rather intrinsic motivation of monitoring for the internal purposes, external monitoring requirements are often asked by funding agencies

to ensure the adherence of agreed procedures, intermediate results, and milestones. Monitoring is distinguished from evaluation by its selective character. It focuses on singular aspects, previously identified and defined in the planning phase, as opposed to putting different reporting aspects together into a bigger picture. Presuming that, compared to planning and evaluation, monitoring is a relatively straight forward exercise, one might think that common understanding on monitoring requirements amongst donors and implementing agencies would make the issue of monitoring an easy one to deal with. Again, reality is very far away from theory. *"Humanitarian agencies seeking to maximise their funding, and reduce dependence on a particular donor, will frequently have to report to ten or more different donors - a very heavy burden for many agencies"* (Ramalingam et al. 2009, p. 58), leading to a reporting overload that eats up valuable time at the cost of internal reporting requirements and learning processes.

It would be up to the donors to either simplify their reporting requirements as such, to structure them in a way allowing implementing agencies to integrate their internal reporting requirements, or to develop a harmonized framework for monitoring and reporting that would disburden humanitarian agencies to live up to numerous different requirements, at best aligned with the internal need for learning and improvement.

Evaluation

As the stumbling block that initiated numerous performance initiatives was the Joint Evaluation of Emergency Assistance to Rwanda (see section 4.1), interestingly, the most commonly used tool of performance management nowadays are predominantly variants of evaluation. This observation might be attributed to the very simple notion that it is more comfortable to judge on outcomes, once 'the deed is done'. But this brings with it the dilemma of ex-post learning: lessons learnt from evaluations can most frequently not be applied to the exact scenario they derive from. They inherently lay the focus on improving the response to the 'next' disaster. And it might not even be certain that lessons drawn from an ex-post evaluation actually inform institutional learning.

In response to what the OECD Development Assistance Committee (DAC) calls 'methodological anarchy' at the end of the 1990s (1999, p. 2), several approaches have arisen to improve the results of evaluations. The OECD DAC established guidelines for the evaluation of humanitarian assistance in complex emergencies, reinforcing the importance of *"explicit and clear statements of objectives"* that will *"require concerted effort over a period of time at several different levels of the humanitarian system"* (ibid., p. 13). Only through an accurate baseline of needs, and corresponding accurate objectives, can meaningful evaluations be conducted. As Hallam points out, the effectiveness of the evaluation process is in many cases significantly limited by the lack of *"clearly stated objectives for the overall programme and its various sectoral and project components, and monitoring information necessary for assessing the performance of projects and thus of the overall programme"* (1998, p. 37).

The common single agency ex-post evaluations have been complemented by additional types of evaluations, such as joint evaluations and real-time evaluations (Ramalingam et al. 2009, p. 59). Evaluations are highly individualized undertakings to gain insight about different parameters of the span of performance (see section 2.3). Like planning and monitoring requirements, they can be motivated by an agencies' own interest for self-improvement or be part of donor conditions.

While in 2000, Crisp pointed out that evaluations too easily turn into technocratic assessment in which the main question is whether an organization has met the formulated objectives, neglecting whether these objectives stand in any connection to needs on the ground (2000, p. 7), over the past ten years the spectrum has broadened. Evaluations now cover all phases of the project cycle by analyzing how successfully the design and implementation contributed to valuable results and what the deficiencies and shortfalls were along this process. The communication distance between headquarters and field office staff as well as the significantly high staff turnover rates in humanitarian aid make the building up of institutional memory that draws together the different perspectives on an intervention probably as crucial as in probably no other sector. Depending on the intended purpose of the respective evaluation, the focus may be set on (OECD DAC 1999, p. 22):

- efficiency (see also: section 2.3),
- effectiveness (see also: section 2.3),
- impact,[11]
- relevance, and/or
- sustainability.

While the evaluation of efficiency mainly depends on the availability and quality of internal documentation about planning and monitoring, evaluating the other parameters, which all demand of the evaluator to interpret the relation between an intervention, program, or project, and the environment it is set in, poses a challenge very similar to the one of the needs assessment process. The mode of impact assessment is highly dependent on the context and the availability of necessary data. Depending on the restrictions on the ground, one or a combination of the following three approaches, provided by Hallam, is employed (1998, pp. 82–85):

The scientific approach is favored by donors who like to see reliable and comparable data on the impact of an intervention. It relies on experimental techniques such as multiple regression and the use of control groups. Besides

[11] As the OECD defines impact as looking at "the wider effects of the project – social, economic, technical, environmental – on individuals, gender and age-groups, communities, and institutions. Impacts can be immediate and long range, intended and unintended, positive and negative, macro (sector) and micro (household). Impact studies address the question: what real difference has the activity made to the beneficiaries? How many have been affected?" (1999, p. 22), this definition can be equated with the definition of utility (see section 2.3) of Bouckaert et al.

immense efforts that are required before, during and after the evaluated intervention, program or project, other practical problems limit the usability of the scientific approach. In an emergency setting it is hardly possible to create control groups from which the delivery of aid services is retained. Moreover, if before and after comparisons are applied, environments of rapid change make it difficult to isolate the effect of humanitarian aid from other factors influencing the situation on the ground.

The deductive/inductive approach is an acknowledgement of the shortfalls of the scientific approach that seeks to gain narrative insight into the effects of humanitarian aid by using anthropological and socio-economic interpretation methods. Through triangulation that uses the expertise of different evaluators and key informants, validity of an evaluation is sought to be established. Consequently, it relies on the ability to judge of those who evaluate.

The participatory approach seeks to evaluate the impact of humanitarian interventions through the viewpoint of the recipients of aid. It is the logical consequence of the notion of accountability towards beneficiaries. Arguing that reality is too complex to distill scientific reasoning about the effects of aid, proponents of this approach suggest that the impact of aid depends on the subjective value added to the lives of beneficiaries. It is questionable whether all aspects of aid are even noticeable by those who receive them. Means of prevention, such as vaccinations would, for instance, certainly be of less subjective value for beneficiary groups than tangible goods such as housing, food, and protection.

Acknowledging the many pros and cons of these approaches, in reality, evaluators rely on a mixture of methods, the deductive/inductive approach being the most predominant (Hallam 1998, p. 84).

Despite widespread consensus about the importance of beneficiary participation and initiatives that stress or advocate for the rights of beneficiaries (e.g. the humanitarian accountability partnership (HAP), the active learning network for accountability and performance (ALNAP), the good humanitarian donorship (GHD) initiative, sphere standards, etc.), it seems that beneficiary involvement still has not gained high prominence in evaluations. While the argument against large-scale involvement of beneficiaries during the needs assessment phase is that it is time consuming, especially in sudden onset disasters, ex-post evaluations would provide the frame for thorough insight into the views of beneficiaries without time pressure. Nevertheless, a recent ALNAP survey found that *"a majority of the respondents to the question of beneficiary engagement [...] stated that they do involve beneficiaries in their impact assessments to some degree"* (Proudlock et al. 2009, p. 41).

Reflecting on performance management along the project cycle as a whole, it has to be noted that setting the focus on evaluations to some extent derives from the deficiencies of needs assessment and planning: if these two stages of the project cycle through tools that lived up to the requirements for humanitarian program- and project management undoubtedly led to sustainable results, extensive ex-post evaluations would become obsolete. In this sense, the comparative

importance of evaluations as a tool of performance management hints at big room for improvement in needs assessment and planning.

Standards and networks

Alongside the approaches to performance that seek to professionalize and integrate the project cycle as such, various initiatives, networks, and think tanks have come about, which focus on specific aspects of the project cycle, supporting selected goals of humanitarian aid or advocating for improvement of performance in relation to these goals. Moreover, developing standards that can guide the design and implementation of interventions has been a topic of attention. These initiatives may be understood as support processes to the actual delivery of humanitarian aid. Their existence is mainly based on voluntary participation of actors within the humanitarian system. As already indicated in several instances in the course of this paper, accountability towards beneficiaries and participatory approaches dominate the field of these initiatives.

The importance of accountability towards beneficiaries stems from the notion that too little attention is given to the voices of the beneficiaries, who are in fact not less than the reason for the existence of the humanitarian sector (see section 3.3), but are underrepresented in the shaping of the way humanitarian aid is delivered. Taking into account the definition of accountability according to which *"accountability is the means by which individuals and organizations report to a recognized authority, or authorities, and are held responsible for their actions"* (ReliefWeb 2008), 'downwards accountability' concerns mechanisms that allow to establish the accountability link between beneficiaries and all aid agencies that is not inherently present within the humanitarian system (see section 4.1).

HAP is the most important singular initiative solely dedicated to creating this link. "*Through self-regulation, compliance verification and quality assurance certification"* (HAP 2011) of its member organizations, HAP employs different methods to ensure downwards accountability. The 2010 HAP standard in accountability and quality management promotes principles of accountability and benchmarks that make up the basis for certification criteria of the 77 current and potential future HAP members (HAP 2010; HAP 2011). A unique feature of the HAP approach is that it offers different levels of certification that in turn are connected to a complaint mechanisms about non-adherence of quality standards. Nevertheless, until now, only 11 organizations have been fully certified through an independent audit (HAP 2011). The inherent shortfall of such voluntary self-commitment is what happens if an organization breaks its self-commitment. The certification mechanism of HAP is a remarkable step forward beyond common practice of self-commitment. Nevertheless, its diffusion is not significant enough to influence accountability on a large scale.

Further think tanks complement the work of HAP in promoting accountability. ALNAP has a broader focus on performance management as such, but regularly

publishes reports on the state of humanitarian performance and evaluates organizational performance, giving consideration to downwards accountability.

The oldest set of standards and at the same time probably the lowest common denominator for the humanitarian sector as a whole is probably the Red Cross' code of conduct. Especially the humanitarian principles are somewhat of a widespread consensus on the nature of humanitarian service provision. The fact that numerous organizations have adopted the humanitarian principles into their internal codes of conduct or their statutes shows that delivering aid is based on a quite homogenous idea of *what should be done*.[12]

What is done, due to the existence of the principles is a different side of the coin. The reason for the widespread diffusion into the humanitarian sector is to some extent attributed to the fact, that they are *only* principles. They are formulated in a rather abstract way, not demanding specific operational rules, meaning that many ways of organizational action can be interpreted as adherence to the principles. The humanitarian principles are undoubtedly a main pillar of humanitarian aid as such, but without more detailed provisions that shape behavior more specifically, they do not hold the potential of improving humanitarian performance management.

The introduction of the sphere standards into humanitarian practice (see section 4.1) somehow represents an alternative approach to the very goal that accountability initiatives pursue. Both seek to provide guidelines for the way in which humanitarian aid should be delivered. But, while accountability initiatives try to tailor responses to specific contexts, inevitably making the delivery of humanitarian aid more complex, the sphere standards seek to find universal patterns applicable to all humanitarian emergencies and thus reduce complexity. Even though standards stress the importance of beneficiary participation and a rights based approach in several instances (Greaney et al. 2011), they essentially replace the process of identifying individual needs on a case by case basis with assumed need levels. Standards provide technical recommendations to each of the key areas of humanitarian aid. Besides the question of *"what is enough?"* which is an apparent criticism of an approach that seeks to establish general rules, the use of minimum standards implicates the risk that the minimum of what would be applicable becomes the baseline for planning and therefore the provision of aid. The question has to be raised, whether it is the aspiration to provide minimal services, just at the verge of enabling people to sustain their lives.

[12] Despite widespread consensus on the humanitarian principles, there still is disagreement about the principle of independence. It is contested by major institutional donors. The Treaty of Lisbon clearly leaves out 'independence', stating that "humanitarian aid operations shall be conducted in compliance with the principles of international law and with the principles of impartiality, neutrality and non-discrimination" (European Union 2007, chap. 3, art. 188J,2.). This may be interpreted as an implicit acknowledgement of the factual dependence of humanitarian aid policies of political interests.

4.1.2 Performance management at the meso- and macro-level

As the meso-level has been defined as *"particular policy fields, specific sectors, or specific service or delivery chains"* (Ramalingam et al. 2009, p. 23), it concerns either cluster-wide performance approaches (specific service or service delivery chain), humanitarian organization as such (specific sector), or the combination of both through performance management in entire cross-organizational interventions. The benefits of approaching performance management at a higher level than the micro-level are apparent. Micro-level performance management inherently focuses on what a project or program does. First, organizational learning can be improved if the internal structure and procedures are scrutinized on the basis of all of its programs and projects put in relation with each other. Secondly, as a program or project by its very nature only covers certain aspects in a given environment, it is difficult to judge on its value in a wider context. Aspects such as overlaps or shortfalls of the sum of various programs and projects only become visible when taking a step back, and taking a look at the bigger picture of how a cluster or a whole intervention responded to needs on a larger scale. The culmination of widening the scope of performance management would be a global perspective on international humanitarian aid that would allow analyzing and improving performance in relation to global humanitarian needs, combining the views on projects and programs, organizations, and whole interventions into an all-encompassing picture of the global humanitarian performance.

Organizational performance management

Beyond program and project-related performance management, organizational performance management means the interconnection of an organization's strategy, mission, or policy with its programs and projects to ensure their congruence. Keeping in mind that by performance management one seeks to employ a closed loop process, by integrating the projects-view and the program-view with the organizational view, management turns from managing multiple programs and projects into managing a homogenous whole.

The complexity of managing organizational performance varies drastically in the humanitarian sector, bearing in mind the differing sizes of humanitarian organizations. It would be relatively simple to establish a coherent organizational performance management system in an organization of fifteen employees and four projects per year. But the necessity for performance management grows with the size of an organization. While the described 'mini organization' would possibly be able to manage its activities in a coherent, efficient, and effective way without formal management procedures, humanitarian 'heavyweights' such as the UN organizations, Cooperative for Assistance and Relief Everywhere (CARE), the International Committee of the Red Cross (ICRC), and the International Federation of Red Cross and Red Crescent Societies (IFRC), etc. are simply by the dimensions of their operations in greater need of structured approach to performance management.

Evaluations of whole organizations (as already discussed in section 4.1.1) play an important role in assessing the state of individual organizational performance management and can give important indications, to what extent an organization employs appropriate performance management methods. But they can only play the role of an enabler to a full-fledged performance management process.

The European Framework for Quality Management offers a comprehensive model through which business can assess their proficiency in performance management. It has been re-interpreted by the quality standards task group of the national council for voluntary organizations into a self-assessment workbook (Griekspoor/Sondorp 2001, p. 212). The impact of this easy-to-access performance management framework for smaller organizations has been quite significant in the United Kingdom (NCVO 2011), including some humanitarian organizations, but its focus is too small, geographically and too broad, by its sector scope, to be considered a significant performance management approach within the humanitarian sector.

Ramalingam et al. discuss two representative results-based management approaches; the WFP's Indicator Compendium and the Planning for Results approach of the Red Cross (2009, pp. 68f.). They refer to a review by the national audit office, according to which the WFP's exhibits major shortfalls (NAO 2007):

- the formulated strategic objectives were too aspirational and the WFP was unable to measure performance against them,
- the selection of KPIs did not suffice in measuring all factors required to achieve an the indented outputs and outcomes, and
- the employed indicators were not frequently enough revisited.

As opposed to this incomplete realization of results-based management, they find the Red Cross' planning for results to be the most extensive organizational performance approach. It encompasses "*planning, budget construction and appeals, implementation, logistics, financial and human-resource management and evaluation"* and *"is an integrated system used at every level of the institution, and at all phases of the project cycle"* (Ramalingam et al. 2009, p. 69). In this sense it appears to be the closest approximation between theory and practice in humanitarian performance management. But it remains a commendable exception.

Noticeably, a BSC approach has been implemented by the Kenya Red Cross Society. The organization has recently been consulted by Balanced Scorecard Eastern Africa, an affiliate of the Balanced Scorecard Institute, in undergoing the process of setting up a BSC system. Their BSC encompasses a defined mission, vision and according core values, a strategy map laying down the four perspectives of beneficiaries/stakeholders, financial stewardship, business processes, and organizational capacities. Each of the perspectives' objectives is complemented with performance measures that, in turn are linked with targets and initiatives. Even though the implementation is too young to judge on its sustainable impact on the operation of Kenya Red Cross Society, their head of

the financial department finds a significant impact on the society's strategic management: *"with the Balanced Scorecard, we can now understand how a project contributes to the achievement of a Strategic Objective which in turn contributes to the results of the strategic focus areas…"* (Omolo 2010, p. 5). This example indicates that the BSC could help consolidating performance management through interlinking the strategic and organizational dimension of performance with the project-level.

Cross-organizational and intervention-wide performance management

Cross-organizational and intervention-wide approaches to performance seek to improve the extent to which responses are needs-based, the capacity to respond, the transparency of responses and thus the coordination and effectiveness of a response. Numerous initiatives have evolved in order to achieve this. These initiatives are less a concerted effort than the sum of individual interests and motives of different groups of actors that lead to a scattered picture of cross-organizational performance management. Accordingly, cross-organizational and intervention-wide performance management approaches cover different areas of the span of performance. While some focus on the assessment of impact and the involvement of beneficiaries into this process, others concern the establishment of information systems to build up a common truth about background facts for interventions, made available to all actors involved, or the collaborative coordination of responses involving the whole project management cycle. Cross-organizational approaches require the aggregation of performance information from programs and projects of single agencies onto the meso-level and into thematic areas (health, food, shelter, etc.) or common themes (beneficiary participation, etc.) (Ramalingam et al. 2009, p. 61).

The UN system has been a precursor in this quest to harmonize responses across organizational boundaries. As the UN system's organizational structure leads to a variety of UN agencies scrimmaging on the humanitarian stage[13], the need to harmonize the humanitarian work of the UN as a whole has been of interest already before the humanitarian sector became commonly aware of the need to improve performance (see section 4.1). The creation of the Office for the Coordination of Humanitarian Affairs (OCHA) was one of the first steps to anchor cross-organizational coordination in a specifically designed coordination body. Initially focusing on the coordination of UN agencies, now OCHA has taken over a broader role on facilitating concerted humanitarian response of NGOs, the UN system, and the Red Cross. The cluster approach, launched in 2005 by the IASC links up humanitarian organizations involved in a specific emergency. On the basis of voluntary participation, the cluster approach

[13] WFP, the United Nations High Commissioner for Refugees (UNHCR), the United Nations Development Programme (UNDP), the United Nations Children's Fund (UNICEF), the UN Food and Agriculture Organization (FAO), the United Nations Human Settlements Programme (UN-HABITAT) are the main UN agencies involved in humanitarian aid. They work based on clearly defined mandates, but in practice their operations show overlap, double-work and inefficiencies.

provides a semi-formal platform for communication and coordination among agencies to facilitate a multi-sectoral response. In practice, OCHA is the focal point of collecting, collating, and disseminating monthly reports of agencies on the ground of a specific operation. Through the 'who does what where database', this information is made available to the public in order to increase the transparency of humanitarian operations (Ramalingam et al. 2009, p. 65).

The first major application of the cluster approach during the humanitarian response to the 2005 Pakistan earthquake allowed an insight into its practical strengths and weaknesses. A review by the humanitarian practice network, based on interviews with representatives of organizations on the ground found that (Street/Parihar 2007):

- The number of clusters grew 'exponentially' in the initial phase of the disaster response;
- Assigning an agency in charge per cluster was appreciated by the organizations;
- Communication was improved among organizations due to the regular forum for exchange;
- The application of the cluster approach was perceived as too mechanic, leading to some organizations becoming entangled in the handling of the cluster approach that they neglected the actual work on the ground;
- The level of success varied from cluster to cluster, also depending on the quality of the cluster lead;
- The cluster approach lacks a strategy for involvement of local NGOs; practically (also due to language barriers), the cluster approach reinforces exclusion of local capacities;
- Inter-cluster coordination was a major concern: *"participants found the increased demands due to the need for inter-cluster coordination to be a burden rather than a help in ensuring adequate and timely responses, in managing information and avoiding both gaps and duplication"* (ibid., p. 34).

A lot of the confusion around the cluster approach can be attributed to the fact that it was still in the development stage, while being used in Pakistan. However it is doubtful in general, whether an approach that dissects the humanitarian response into clusters to then find one of the greatest challenges to be the coordination among those clusters (because of the many factual interdependencies), is the right mode of coordination.

Relating the cluster approach to the models of the depth and the span of performance, it is unclear whether it may be considered an actual meso-level approach. It certainly covers at least the stages of situation analysis, response analysis, response planning, and response implementation, and therefore encompasses most of the span of performance. But its voluntary nature and therefore loose coherence leaves room for doubt whether it is not more of an informal forum for exchange on individual organizational performance management, in which no mechanism ensure the actual adjustment of one organization's activities to those of all others.

Another UN induced performance approach is the central emergency response fund (CERF). It was initiated in 2005 by a resolution of the UN general assembly in order to pre-position funding to allow earlier responses and to circumvent regular appeals for funding that were found to be too slow to allow rapid response. It specifically focuses on under-funded emergencies (comparable to the EU's focus on forgotten crises). As long as being based on appropriate global needs assessment, generating a birds-eye perspective through centralization of funds is a commendable step towards improving strategic decisions on where to intervene. It should be an elementary, part of sector-wide performance management, together with common humanitarian planning processes (see also: section 4.1.1 on the CHAP).

While comprehensive quantitative data to support decision making is often made available only after an intervention, several cross-initiatives try to improve the real-time availability of mainly health- and nutrition-related indicators to support monitoring and evaluation efforts. The SMART initiative, the Complex Emergency Database (CE DAT) and the Health and Nutrition Tracking Service (HNTS) are examples of these efforts, all seeking to collect relevant data and to administer databases (SMART 2011; CE DAT 2009; HNTS 2011). Their relevance is currently growing, as more agencies start being involved in the surveillance processes and more countries are targeted incrementally (Ramalingam et al. 2009, p. 63). A shortfall of these initiatives is that they deliver quantitative information that is largely independent of the emergency context. Therefore, they do not yet represent a solid baseline for the evaluation of outcomes attributed to humanitarian interventions.

Joint evaluations are probably the currently most powerful tool in humanitarian performance management, analyzing outcomes of entire interventions. Even though they only focus on the last stage of the project cycle, their cross-organizational character allows understanding the potential of common approaches to respond to humanitarian emergencies. Only through the perspective of the humanitarian intervention as such may comprehensive conclusions be drawn on the impact of humanitarian aid on local contexts. Despite the benefits of these evaluations, due to the immense effort involved, it took a long time before they seemingly have become common practice in the sector.

Joint evaluations increasingly work with means of beneficiary participation. Besides this, inter-agency initiatives have come about that seek to transform the call for beneficiary accountability into tangible tools that should allow beneficiaries to become a part of the aid provision process. The participatory method of evaluation (see section 4.1.1) is the core of an important approach to performance that displays the feasibility of beneficiary involvement on a large-scale. Over the past years, the Fritz Institute has carried out beneficiary studies on the sustainable impact of the aid they received. These surveys concerned the Java earthquake in 2006, the 2005 Pakistan earthquake, Hurricane Katrina in 2005 and the 2004 Indian Ocean Tsunami. These approaches hold the potential of complementing evaluations with a wider view on humanitarian services and

their outcomes. However, they are face the same restriction as traditional evaluations; the ex-post learning dilemma (see section 4.1.1).

The cross-organizational and intervention-wide approaches described above, tend to based on links too loose between the organizations involved. Whether the approaches are focusing on coordination, data collection, or joint evaluations, no effective mechanisms are in place to ensure compliance. Apart from the discussed approaches and initiatives, cross-organizational and intervention-wide means for performance management are still too rare.

System-wide performance management

As meso-level performance approaches have only made an appearance in the past decade and are still given too little importance, it is not surprising that no real system-wide approaches to humanitarian performance exist. Hallam's observation that *"there is currently no capacity in the system for regularly promoting and organising 'system-wide' evaluations"* is still valid thirteen years later (1998, p. 28). Efforts such as the Millennium Project in development aid have no counterpart in humanitarian aid. The least common denominator for concerted performance management consists of the humanitarian principles and legal frameworks. However, some initiatives recently appeared that take an overarching perspective on humanitarian aid and advocate for ideas following the notion that a system-wide view on humanitarian performance is valuable and necessary.

ALNAP is a unique forum founded in 1997 that seeks to *"foster a culture of active learning and accountability to improve humanitarian performance"* (ALNAP 2007, p. 4). The network of currently seventy three full members encompasses important donor agencies, the ICRC and IFCR, national Red Cross Societies, the UN's humanitarian agencies, and major NGOs (ALNAP 2011). The work of ALNAP is more of a platform for research and exchange than an actual performance tool as such. Their promising activities of investigating on and promoting performance approaches make them an important protagonist on the meta-level of performance management. A core feature of ALNAP's work is the Humanitarian Performance Project and the connected Annual Review of Humanitarian Action that investigates on current trends and challenges in humanitarian performance and proposes approaches to enduring and new problems of humanitarian performance management.

Another cornerstone of heading towards system-wide performance management might be the GHD initiative. In 2003, the sixteen donor governments of the OECD DAC consolidated their view on what should constitute principles of good donorship. Acknowledging that *"as financiers and increasingly as strategic actors in their own right, official donor governments exert a significant influence over the outcome of humanitarian action"* (Harmer et al. 2004, p. 1), the GHD initiative holds the potential to impact the way which humanitarian performance is carried out. Donors have the power to coerce their conditions on implementing agencies (leaving aside those NGOs with a strong foundation of private donors). Thus, the GHD initiative should be considered as a 'vehicle' to

target performance on a system-wide level. Factually, though, the GHD initiative is little more than a lip service embedded in a reality of politicization of aid, with no mechanism that would ensure the adherence of donor commitments (ibid., pp. 7f.). Out of this shortfall, the humanitarian response index (HRI) was born. In 2007, it was created by DARA, in order to assist "*the humanitarian community […] and in particular OECD/DAC donor governments to meet the objectives to which they themselves have signed up"* (DARA 2010, p. 10). The recent HRI is based on field interviews with almost 500 implementing agencies and donor representatives, assessing the adherence of the principles and objectives of the GHD initiative through quantitative and qualitative assessment (ibid.).

4.2 The status-quo: putting together the pieces

The landscape of performance management initiatives in humanitarian aid is a heterogeneous one. The diagnosis that *"compared to the for-profit sector, the nonprofit world is back in the late1970s and early 1980s, when Japan was beating up on American businesses"* (Bradach 2002 cited in Byrne 2002, n.p.) appears to be particularly applicable to humanitarian performance management. Even though an implicit common goal underlies the humanitarian endeavor, performance is still mainly considered in terms of program-based and project-based views on performance.

The introduction of project cycle management methodology has been a major step towards professionalizing humanitarian operations. But up until today, the use of project cycle management tools, such as the LFA, still show significant shortfalls, both in terms of their structure and their application. There is still no clear separation of situation analysis, response analysis, and planning. Instead, planning of humanitarian projects and programs is guided more by internal considerations than external needs, leading to a certain detach between factual needs of affected populations and identified 'needs' that become the baseline for project implementation and subsequent monitoring and evaluation.

As needs assessment is the point of entry for humanitarian operations, it is crucial to find solutions to actual 'needs-based needs assessment'. Inaccuracies in indentified needs are inevitably carried on through the implementation of a program or project and will not even be identified in the evaluation phase that is, in turn, measuring the achievement of objectives derived from inaccurately identified needs. Too little has been done so far to ameliorate this fundamental deficiency. The CHAP provides a good first step towards better assessment, but has to gain importance to make a noticeable difference.

Furthermore, planning tends not to lay down a clear enough frame for monitoring and evaluation through clear establishment of objectives and measures to track the achievement of these to the necessary extent. On top of this, external reporting requirements in all project phases cause a reporting overload that keeps organizations from developing their own project management methods that would support organizational learning.

Evaluations have been elevated into being the 'magical bullet of performance management', both for external accountability and learning. This seems to be a necessary consequence of the inadequacies of needs assessment, planning, and monitoring. If these three were conducted in a more coherent way, humanitarian agencies would be enabled to have a better idea about the appropriateness of their project activities and the impact of their operations. Ex-post perspectives are valuable opportunities for learning, but are to some extent the replacement for omissions in project management.

While the LFA is still subject to scrutiny, few mature project cycle management approaches stand out. The ICRC's planning for results approach, described in sub-section 4.1.2, has been found to be a good example of coherent project cycle management and closest to the theory of performance management. But unfortunately the ICRC remains a mentionable exception.

As extensive involvement of beneficiaries during the onset of an emergency is difficult, evaluations have become the dobbin for living up to the idea of a participatory approach. But this cannot be enough to deliver humanitarian services tailored to the local context. If evaluations find that specific beneficiary needs have not been satisfied, this might give hints for organizational improvement and help the 'next emergency', but another chance was missed to fulfill the humanitarian mission. In fact, no reliable solution has been found yet to bridge the gap of a lacking structural accountability towards beneficiaries. It is not even a consensus, which role beneficiaries can play in the sphere of humanitarian performance management. The emergence of initiatives solely dedicated to the theme of accountability gives indication that accountability has not yet become common practice in humanitarian programming. ALNAP is currently on the forefront of combining these two aspects.

Understanding the challenges at the micro-level of performance helps understanding those on the meso- and macro-level. Without functioning recipes to performance management of programs and projects, the quest to take performance beyond this level seems like a Sisyphean task. At present, *"the separate pieces [...] do not [...] add up to even the sum of their parts"* (Mitchell 2008, p. 18). What lacks in the first place, is a commonly defined objective, which would encompass all international actors in humanitarian aid. The implicit common ground consists of the humanitarian principles and the frameworks, and the humanitarian principles are not even in place without ongoing debate about them. Without a clear strategic alignment (see section 2.4), the embodiment of cross-organizational performance management inevitably becomes what it is in humanitarian aid today. There are "*no baselines, no agreed definitions of performance, and an absence of any kind of mechanisms able to track performance"* (ibid., p. 17), and the multitude of frameworks, tools, and initiatives inevitably leads to overlaps and parallel efforts.

Few initiatives target only fragments of the span of performance on the meso-level. The cluster approach is a valuable tool to harmonize project activities of organizations on the ground along the entire project cycle, but its loose mode of

cooperation causes unnecessary and unclear coordination tasks for the involved organizations, diverting resources from actual project work. It is worth asking whether coordination through cluster segmentation adds too much of an additional (and artificial segmentation) of humanitarian operations. This approach would require some revisiting and clearer procedures to become a role model for inter-agency coordination.

Joint evaluations are appreciated as being able to deliver comprehensive, cross-organizational insight into impact and sustainability of interventions. But due to their specific and highly customized nature they fail to inform system-wide and comparative learning (Ramalingam et al. 2009, p. 6). Only few means of data collection and -dispersion fill this gap by supplying real-time quantitative indicators (and that only about health and nutrition data) to enable monitoring and evaluation in a comparable manner.

Bearing in mind that donors have been, by their inherently influential role in the humanitarian sector, the driving force behind structural changes of the sector in the past (as discussed in section 4.1), realistically, only with their commitment may changes towards system-wide performance management be realized. The GHD initiative, as long as becoming more than a lip service may be the pivotal point for effective changes hauled into the implementing arena of humanitarian aid. The emerging of the HRI and its recent findings reveal, however, that there is much room for improvement for donor policies and conditions with regards to humanitarian performance management. Analyzing donor practice through five perspectives, the humanitarian response index finds that donors still do not separate their political agenda from humanitarian policies and

"at the crisis level, donors' actions to support greater accountability towards beneficiaries are limited. Most humanitarian organisation interviewed stated that such initiatives were [...] developed at their own initiative or as part of their own internal procedures and commitments" (DARA 2010, p. 47).[14]

Furthermore, DARA identifies a *"slow progress in reforming the humanitarian system"* which *"means that aid efforts are not as efficient or effective as they should be"* (ibid., p. 20).

It seems as if the bottom-up process of performance management of the past fifteen to twenty years that has led to the state of performance management today would need consolidation, induced by a top-down process. Only through donor coercion may the scattered landscape of performance management fragments be turned into a comprehensive, integrated, and efficient system to

[14] The HRI assesses (DARA 2010, p. 19): 1. Are donor responses based on the needs of the affected populations and not subordinated to political, strategic or other interests? 2. Do donors support strengthening local capacity, prevention of future crises and long-term recovery? 3. Do donor policies and practices effectively support the work of humanitarian organizations? 4. Do donors respect and promote international humanitarian law, and actively promote humanitarian access to enable access to enable protection of civilians affected by crises? 5. Do donors contribute to accountability and learning in humanitarian action?

handle performance. But considering current donor policies and practice, a large-scale paradigm shift, as caused by the post-Rwanda discussion generated might be needed to bring the necessary changes towards consolidated humanitarian aid about.

5. Adaptability of the Balanced Scorecard to system-wide performance management in humanitarian aid

In this chapter, based on the preliminary conclusions of sections 3.4 and 4.2, the adaptability of the BSC model to a sector wide model for performance management will be discussed. Requirements for the design of a system-wide performance BSC system broadly derive from the following areas:

- the overall methodology and provisions of performance management,
- the general BSC methodology,
- considerations regarding the BSC in cross-organizational contexts,
- considerations regarding the BSC in the public and non-profit sector,
- the interaction of the humanitarian sector with the environment that it is set in, and the implications on the manageability of performance derived thereof,
- the internal structure and dynamics of the humanitarian sector and its implications on the manageability of performance, and
- considerations regarding the current status quo in humanitarian performance management.

These requirements will guide the discussion in the course of this chapter: as a first step, an analysis of the common ground for consensus on an overarching mission for humanitarian aid will be carried out, which is the precondition for the further design of a system-level BSC (as discussed in chapter 3). Following this, the internal structure and the dynamics of the humanitarian sector, as well as the implications of the environment on requirements for a sector-wide performance management are discussed against the background of the theory of performance management and more specifically the BSC methodology (discussed in chapters 2 and 3). Special consideration is given to the state of the art of research on the developed concepts for a BSC in cross-organizational contexts (see section 3.2) and the BSC in the public and non-profit sector (see section 3.3).

As discussed in section 3.4, the preliminary result on both scenarios show that the theoretical baseline for a BSC could provide an approach to sector-wide performance management in humanitarian aid. However, further consideration has to be given to those characteristics of the humanitarian sector, which scholars have not considered yet. Whereas the discussed concepts of performance management show how to extend the view of performance management onto impacts and outcomes of humanitarian aid, and the ideas of

Kaplan, Niven, Erdmann and Richert show ways to incorporate the different stakeholders' views (donors and customers), and the building up of organizational capacities into performance management, key questions that have not yet been answered will be discussed below. At the end of this chapter, a conclusion on the adaptability of the BSC to sector-wide humanitarian performance management is drawn.

5.1 A common mission for multi-level performance management

The starting point for the design of any BSC has to be the formulation of a strategy. Consequently, a system-wide BSC model for the humanitarian sector would require clarity on, and consequently the definition of a common mission amongst all actors of the humanitarian system. Without high validity, the acceptance of a common definition will be low. And any subsequent steps of designing the scorecard will be erroneous, leading to incongruence between formulated mission and objectives, decisions based thereon, and the actual behavior of humanitarian actors.

As discussed in sub-section 4.1.2, there is no formally accepted common mission in humanitarian aid. What comes closest, are the humanitarian principles. At the very top of the list of humanitarian principles stands the principle of humanity. Slim calls it *"the least controversial"* of the principles (1997, p. 345), and Gnaedinger considers to be, next to the principle of impartiality, *"the principle that most, if not all, humanitarian actors adhere to"* (2007). Therefore, the principle of humanity represents a good point of departure for identifying common ground for the 'humanitarian mission'.

According to the principle, the main purpose of humanitarian aid is to *"prevent and alleviate human suffering wherever it may be found"*, and more specifically to *"protect life and health"* while *"ensuring respect for the human being"* (IFRC 2011). As this definition of *humanity* leaves room for interpretation, the means to pursue this purpose may vary broadly.

While Harmer et al.'s definition stresses that the time-frame for humanitarian response is during and in the aftermath of an emergency (2004, p. 2), other means of aid have shifted into the focus of humanitarian agencies: for instance, they include disaster risk reduction, advocacy, development tasks or peace building.

In defining a common humanitarian mission, it might seem advisable to choose a broadly formulated mission to cover the whole range of means to an end. On the other hand, a rather generic mission must not necessarily exclude disaster risk reduction and other 'non-classical' means of aid. But in order to judge on the benefit of capacity building measures (as somewhat of an alternative to direct response) they should be explicitly included in the system-level mission for humanitarian aid.

Thus, the proposed mission for humanitarian aid on a system-level is the following, which is also used by the GHD initiative (2003):

"To save lives, alleviate suffering and maintain human dignity during and in the aftermath of man-made crises and natural disasters, as well as to prevent and strengthen preparedness for the occurrence of such situations."

This definition will be the basis for the further considerations in the subsequent sections.

5.2 Adaptability to the system and organizational level

Based on the formulated mission of humanitarian aid, this section discusses the implications of a potential sector-wide BSC model to the system-level (macro-level) and the organizational level (meso-level) of performance.

A system-level scorecard would represent a view on the performance of the humanitarian endeavor as a whole, in relation to global humanitarian emergencies. Based on a broadly formulated mission, the system-level scorecard could provide different angles to those aspects of performance that are most vital to humanitarian performance. The perspectives of the system-level scorecard reflect these angles. Sections 3.2 and 3.3 discussed proposals for cross-organizational use of the single-agency BSC model on the one hand, and its adaptation to public and non-profit organizations. In the following, these two approaches are sought to be combined and applied to a humanitarian system-level scorecard that is the point of departure for further disaggregation into the subjacent levels of performance (meso-level and micro-level).

5.2.1 Adapting the perspective at system-level

As the defined mission of humanitarian is to

"save lives, alleviate suffering and maintain human dignity during and in the aftermath of man-made crises and natural disasters, as well as to prevent and strengthen preparedness for the occurrence of such situations,"

it obviously seems commendable to keep the customer perspective at the top of the system-level scorecard. This is in line with scholar's opinions, who unanimously vow for the customer perspective (together with the formulated mission) to reflect an organizations (or a non-profit sector's) ultimate purpose (see section 3.3).

Taking into account that a system-level view on humanitarian aid requires a logical aggregation of different types of actors, Kaplan's proposal to separate the customer perspective into beneficiaries and donors cannot be upheld for the system-level scorecard. At the overarching level, donors cannot logically be considered a customer of aid, as donors can simply not seek to serve donors. Therefore, beneficiaries would remain below the mission as the single type of customer. Potential objectives of the beneficiary perspective will be discussed in detail below.

The process perspective would encompass all means to enable, support, and improve humanitarian services for beneficiaries. Processes are recurring procedures within the humanitarian sector that influence the operations of

donors and implementing agencies. On a system-level they may be common processes, such as the CHAP, the CERF, or aggregated single-agency processes.[15] The process perspective could comprise of different aspects of primary, supportive, and management processes:

- needs assessment processes,
- planning processes,
- monitoring and evaluation techniques and standards,
- logistical processes,
- fundraising processes, and
- financial management processes.

Following the ideas of Erdmann and Richert (as discussed in section 3.2), cooperation is a key element of extending the BSC beyond organizational boundaries. In the course of this paper, it has become evident that cooperation among humanitarian organizations is a key enabler to improving outcomes of aid (see especially: section 4.2). Moreover, cooperation with beneficiaries is an important key aspect for improving the suitability of aid. Thus, a cooperation perspective, which reflects these two dimensions of cooperation, should be included in the system-level scorecard. It would be subordinated under the process perspective. On the meta-level, cooperation helps to improve meta-processes that one agency alone would not be able to manage effectively and efficiently on its own. While the process perspective would measure the effectiveness and efficiency of common approaches such as the CHAP and the CERF, the cooperation perspective would appraise, to what extent the sector uses these common approaches (as opposed to single-agency approaches). The cooperation perspective could provide objectives and KPIs for:

- inter-agency cooperation with regards to:
    - appeals for funding,
    - collective needs assessment,
    - collaborative planning, monitoring, and evaluation efforts;
- the level of beneficiary involvement throughout the project cycle;
- beneficiary surveys.

Finally, organizational capacity and learning represents the long-term angle concerning humanitarian aid, which would seek to identify and improve means of sustaining organizational knowledge and efficiently employ this knowledge for the improvement of processes and the way of cooperation. Concrete objectives for the organizational capacity and learning perspective could be:

- entry requirements for humanitarian staff,

[15] To aggregate the single-agency view, accumulation or averaging of single-agency fulfillment of objectives would be used, depending on the type of objective.

- training, education, and personal development,
- quality of research initiatives, and
- linkage of research and practice.

Figure 9: System-level Balanced Scorecard

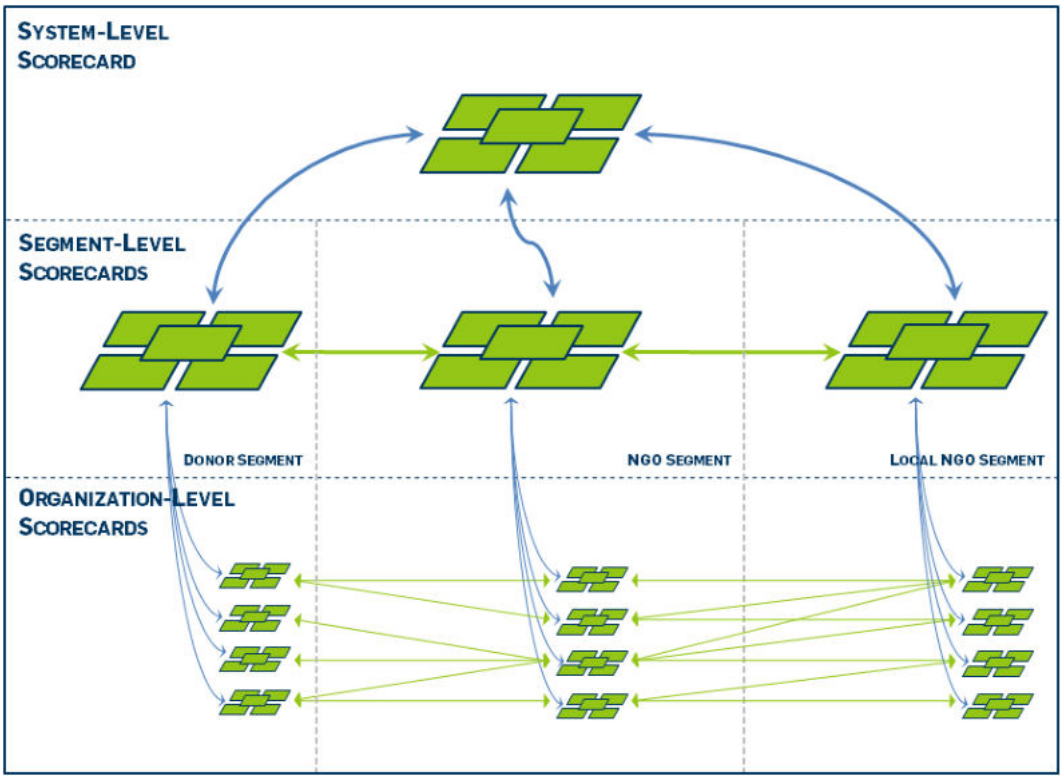
Source: Own composition

Figure 9 illustrates the hierarchical interrelation of perspectives of a potential system-level BSC. The perspectives underlie the cause-and-effect relation described in section 3.1. The beneficiary perspective would be the substantiated formulation of the mission. It would break down the mission into its 'components' through identified objectives and KPIs. It represents the outcome-view on humanitarian aid. The process perspective would be the enabling perspective (in the sense of performance drivers, see section 2.3) to the achievement of objectives formulated in the beneficiary perspective. It depicts all those aspects of performance (as described above) that the humanitarian sector undertakes to succeed in its mission. The cooperation perspective would represent all cross-organizational efforts (including beneficiary involvement) to improve humanitarian processes (appeals for funding, needs assessment, beneficiary surveys, etc.) through consolidated efforts. Finally, the organizational capacity perspective would foster process improvement and cooperation. Through research, better ways of cooperation are identified and may be put into practice and through focusing on human resources and their selection, training, and personal development the quality of all humanitarian processes can be improved.

5.2.2 Integrating the levels of performance

The BSC asks the conjunction of the system-level perspectives and objectives (vertical alignment of objectives; see section 3.2) amongst each other and across the levels of performance. As the humanitarian system consists of distinguishable types of organizations that contribute in different ways to the fulfillment of the 'humanitarian mission', it could be advisable to employ segment scorecards (as discussed in section 3.2). These segment scorecards would represent consolidated views on performance management for the group of donors, the group of NGOs, and local NGOs.[16] The two criteria for this three-fold distinction are the level to which actual project work is carried out by an agency, and the role of the specific agency in the flow of funds, and consequently the upwards accountability deriving from this. Donors are usually not actively involved in the implementation of aid. They provide funds linked with specific conditions regarding the kind of aid, and the way needs are assessed and planning, monitoring, and evaluation are carried out. NGOs (INGOs, Red Cross agencies, UN agencies) conduct humanitarian work themselves, but also pass on funding, while local NGOs are at the receiving end of funds.

Figure 10: Disaggregation of objectives from the system-level scorecard



Source: Own composition

This depiction is certainly a simplification of reality, but suffices to distinguish the types of organizations for performance management. This distinction is important, as the different types of organizations would have to employ different

[16] For performance management purposes, the legal status of UN agencies and elements of the Red Cross movement is not relevant. The UN system's humanitarian agencies and the elements of the Red Cross movement are therefore included in the NGOs' engagement.

objectives for the perspectives of the segment and organization scorecards. For instance, the cooperation perspective of donors would focus more on the effective allocation of funds, whereas cooperation between NGOs and local NGOs would refer more to the operational cooperation in projects. The aggregation of performance from organization-level to segment-level, and from segment-level to system-level would be achieved through accumulation or averaging of the level of fulfillment of an objective by all organizations relevant to this objective. Figure 10 illustrates the disaggregation logic from the system-level down to segment and organization-level.

5.2.3 Identifying humanitarian objectives

According to the BSC methodology, objectives for each of the perspectives have to be identified. Starting from the beneficiary perspective that represents a concrete subdivision of the mission into its enabling components, in a downwards process, objectives for the subjacent perspectives would have to be defined.

The beneficiary perspective should break down the mission into its structural components. It would have to include an exhaustive set of objectives that, if met, help achieving the formulated mission. Objectives associated with the beneficiary perspective represent outcomes in the sense of the span of performance. Therefore, the theme of accountability would integrate into the beneficiary perspective. For humanitarian aid, the mission to save lives and alleviate suffering and the different means to achieve this would have to be translated into concrete objectives.

There is a consensus on the core sectors of humanitarian aid. It comprises of the four *"life-saving sectors*" (Greaney et al. 2011, p. 4):

- *"water supply, sanitation and hygiene promotion,*
- *food security and nutrition,*
- *shelter, settlement and non-food items, and*
- *health action"* (ibid.).

Relating these areas to the time-frame of the formulated mission, the life-saving sectors represent means that take place *and* have effect during or in the aftermath of man-made crises and natural disasters. But humanitarian operations have been extended well beyond this traditional scope of humanitarian action (as discussed in section 4.1). The means to save lives and alleviate suffering are not necessarily immediate and situated in the aftermath of emergencies. These means have complemented the life-saving sectors (also see The beneficiary perspective would exhibit the extent to which objectives associated with it are met on a global level. To ascertain this, needs would have to be identified on a global level on a permanent basis and compared with outcomes of humanitarian aid.

Figure 11) and either

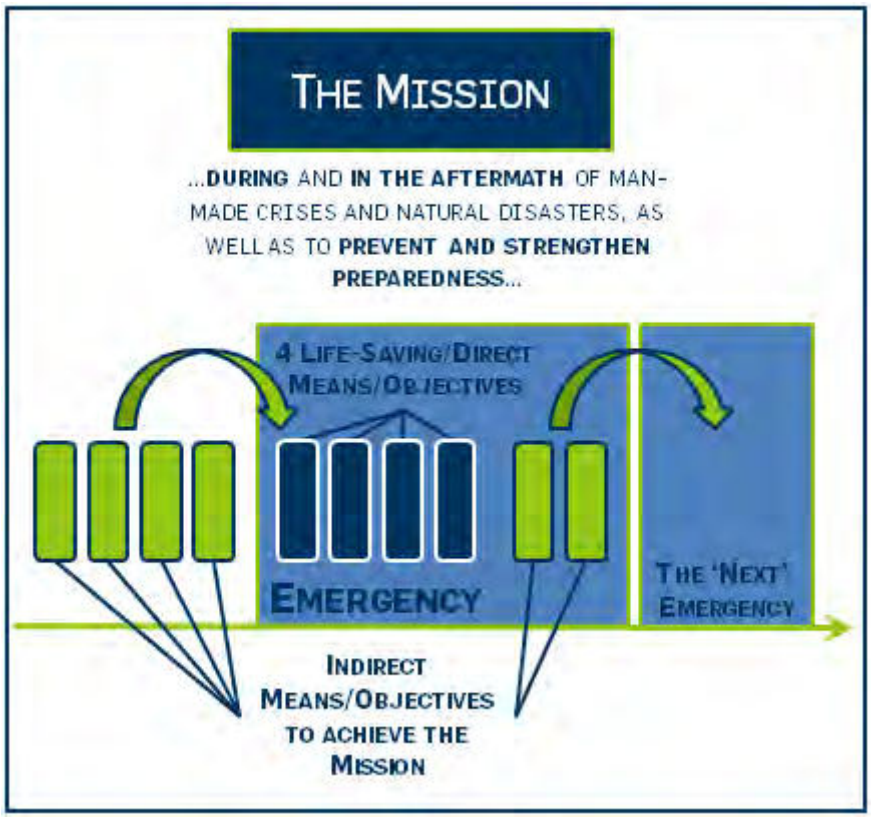
- take place before an emergency in order to have effect through strengthened capacities during or in the aftermath of the emergency, or
- take place during or in the aftermath of an emergency to strengthen capacities in the long-term (the next disaster).

The list of 'indirect' means contributing to preventing and strengthening preparedness of emergencies and consist of, but is not restricted to:

- disaster risk reduction,
- protection,
- gender issues,
- education,
- income generation,
- infrastructure,
- human rights,
- the rule of law,
- HIV/AIDS, and
- environmental issues.

The beneficiary perspective would exhibit the extent to which objectives associated with it are met on a global level. To ascertain this, needs would have to be identified on a global level on a permanent basis and compared with outcomes of humanitarian aid.

Figure 11: Means to achieve the mission of humanitarian aid



Source: Own composition

An important aspect to take into account when trying to define objectives for the beneficiary perspective is that the life-saving means of humanitarian are inherently reactive (as initially discussed in several parts of chapter 4). The scholarly proposals discussed in section 3.3 have not given indications, how to treat this fundamental characteristic of aid. While development aid and for-profit businesses are rather active, in a sense that they add value to a comparably stable and predictable socioeconomic environment, humanitarian aid seeks to 'fill gaps' that interruptions of a presumably stable environment have caused (see Figure 12).

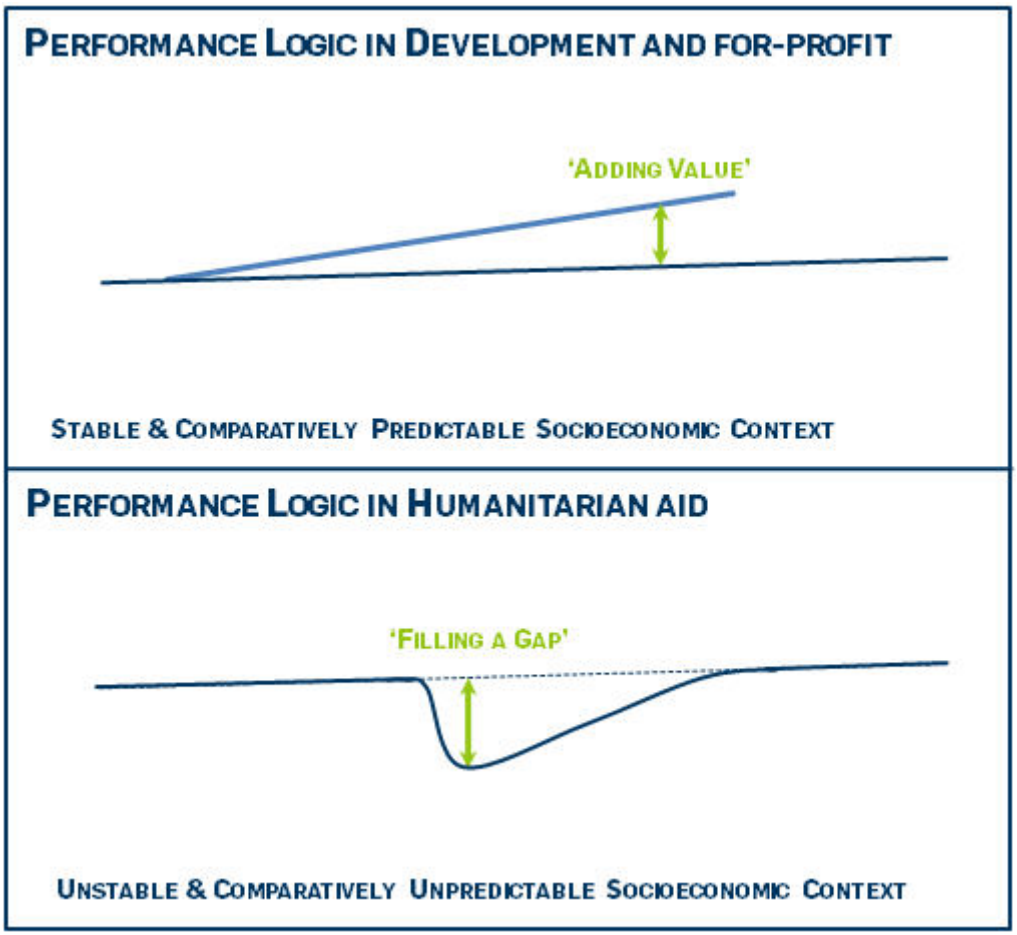
The key cause for planning difficulties, unexpected outcomes and at the worst, failing interventions, has been the difficulty to assess humanitarian needs. Taking into account the reactive nature of humanitarian aid, the general difficulties of assessing needs (as discussed in section 2.3) are complemented by an additional layer of difficulty: performance measurement in humanitarian aid requires that the measurement of outcomes, and consequently the extent to which objectives are met, has to be conducted against the scale of negative developments that humanitarian aid seeks to counteract, and that are largely unknown, until the actual onset of an emergency. "*In humanitarian aid the aim is often to avert negative change (for example to prevent famine), rather than bring about a positive change"* (Hofmann et al. 2004, p. 1). Consequently those interruptions, that cause the very necessity for humanitarian action, make performance planning and before and after-comparison of KPIs associated with objectives of the beneficiary perspective a major challenge for performance

measurement, even at lower levels than the system-wide level. In the global context, humanitarian needs vary from day to day, from year to year rendering planning and reliable absolute measurement of outcomes very difficult. Outcomes are always dependent on and therefore relative to the extent of the negative impact of emergencies.

The variety of means brings with it another big challenge: bearing in mind the reactive nature of humanitarian aid and the according difficulty to assess its effects, the various dimensions of means listed above make a valid and comprehensive appraisal of the attribution of the various means of humanitarian aid to the fulfillment of the mission immensely difficult. A complex web of interrelations connects the life-saving sectors and the indirect means of humanitarian aid. Depending on the specific context, they might be complementary, interchangeable or rivaling means to the same end. Clear cause-and-effect relationship can barely be established. Health-related objectives (e.g. number of treated cases of malnutrition) might stand in a relatively clear connection to the mission, in terms of how they contribute to its fulfillment. In contrast, the outcomes of shelter and non-food items-related activities and even more, the outcomes of 'indirect means' are hardly traceable in terms of their contribution to the mission.

This phenomenon becomes more difficult, the broader the perspective on performance is. As initially discussed in the context of evaluation (see section 4.1.1), when assessing outcomes on a project level, the contribution of humanitarian aid to the saving of lives and alleviation of suffering might still be fairly identifiable. But on a system-level, ascertaining the global contribution of aid through a variety of means employed by a large number of different organizations throughout different emergencies is nearly impossible.

Figure 12: The performance logic of humanitarian aid



Source: Own composition

In fact, activities in humanitarian aid are based on lessons of practitioners on 'what works', but are hardly measurable or comparable in terms of their outcomes. The sphere standards circumvent this difficulty by providing imperative indicators. But even by counting, for instance, the calories taken in per person per day in a given emergency and by ensuring that a sufficient calorie count is provided, there would still be no evidence on the impact of this on the contribution to the mission to save lives.

Non-achievement of an objective clearly causes failing the mission to save lives and alleviate suffering: any death is a clear indicator for the failure of aid. But to reverse this causal connection is comparably difficult. While it is *probable* that food interventions, health interventions, and other means of aid have an impact on the survival of recipients of aid, a complex web of other factors that might 'make the difference' come into play. This makes measuring the effect of aid, in the sense of attributing the achievement of objectives to a humanitarian intervention (e.g. to what extent the supply of food helped ensuring sufficient calorie intake) and, in turn, attributing the achievement of the mission to the achievement of an objective: even if a clear causal link between providing food and sufficient calorie intake could be established, again, there are numerous other factors that can impact positively or negatively on lives saved.

As discussed, the causal connection of the mission to save lives and alleviate suffering with the objectives through which one seeks to achieve this is loose. The component of the mission, *"to maintain human dignity"* seems even more incompatible with the concept of performance management on a global level. To maintain human dignity depends on factors that differ from individual to individual. Clearly assessing to what extent dignity has been maintained would require immense efforts. Realistically, aside from beneficiary surveys, only standards, common sense, and intercultural sensitivity are usable tools to ensure human dignity in the delivery of aid.

Thus, using the beneficiary perspective for performance management on the system-level is a major stumbling block that brings up the question whether the systematic and quantitative methodology of the BSC is applicable. Setting clear objectives is certainly valuable in itself. But the variety of means of humanitarian aid is too broad to be forced into a frame asking clearly interlinked objectives that are to represent the outcomes of aid for beneficiaries. The current proficiency of needs and impact assessment techniques, that even fail to connect the phases of the project cycle on project level (as discussed in section 4.2) would not allow attributing global outcomes of humanitarian aid to the achievement of the humanitarian mission. Consequently, the subsequent steps of identifying objectives for the process perspective that foster the quality of outcomes for beneficiaries and objectives for the subjacent perspectives as well as the further disaggregation of the system-level scorecard would be similarly fallible.

5.3 Adaptability to interventions and projects

As discussed above, on the macro-level the BSC's systematic methodology of setting and disaggregating objectives across the levels of performance is incompatible with the characteristics of humanitarian needs and impact assessment. Consequently, the application of the disaggregation method down to the micro-level could only be fallible. However, for further research on sector-wide performance management it is important to point out that the research approaches discussed in chapter 3 do not yet provide suggestions for another characteristic of aid that would, had the BSC been applicable to the macro and meso-level of humanitarian performance, pose difficulties to the applicability on micro-level: beyond the ongoing debate about the actual continuity of the 'relief to development continuum' that challenges the view that humanitarian aid is a short-term intervention in the face of protracted crises, the mode of humanitarian operation is predominantly project-based. This would pose a challenge for applying the BSC model that is based on the assumption of rather continuous operations. Not even in research on the usability of the BSC for project management in corporate businesses, much has been said about the interrelation of the BSC with the characteristics of project management. When connecting the BSC provisions with project management, the main feature is the closer interlinkage of the strategy with priorities of project management (Norrie/Walker 2004).

In terms of humanitarian aid, this would mean that cross-organizational, but rather *institutional* performance management frameworks such as the BSC could help aligning humanitarian objectives in the long-term-perspective could be a strategic reference, when an intervention program or project is initiated. Sectoral alignment on humanitarian objectives could help identifying shortfalls in program objectives against the background of organizational core competencies, formulated objectives, and the defined priorities of learning and accountability. Nevertheless, it is important to stress that instruments such as the BSC can only guide and complement, but not replace project cycle management methods.

5.4 Results

While it seems advisable for the humanitarian sector to approach performance in a coherent manner by putting into context the various dimensions of performance described in chapter 4 through a balanced view that is provided by the BSC, the full-fledged methodology that accompanies the balanced scorecard is not applicable to a system-wide angle onto humanitarian performance.

The balanced view that the BSC provides, would allow a comprehensive view on many identified aspects and challenges related to humanitarian performance. The beneficiary perspective could take a look at outcomes of aid and would represent the appraisal of the extent to which the humanitarian sector is accountable towards its beneficiaries. The process perspective would represent the view on the internal efficiency and effectiveness with regards to the primary processes along the project cycle and support processes, such as funding procedures, logistics, and financial management. The cooperation perspective would examine the progress in cooperative approaches that are needed to undertake humanitarian aid as a homogenous endeavor. The organizational capacity and learning perspective would assess the means for long-term improvement of methods and the generating and maintenance of organizational learning. The perspectives deducted from the BSC models discussed in chapter 3 could be an important starting point for the humanitarian sector to evolve towards more advanced performance management across all levels of the depth (see section 2.3) of its performance.

But the systematic methodology that is required to put the BSC into practice as a cross-organizational performance management can neither be applied to the system-level, since attribution of objectives of the beneficiary perspective to the mission is nearly impossible (see section 5.2.3) nor does it allow the vertical alignment of objectives from the system-level down to the segment and organizational level, as the challenge to attribute achievements of organizations' objectives to the system-level objectives and consequently the 'humanitarian mission' is the same as on the system-level.

Nevertheless, some important initiatives towards consolidated performance are in place already. The CHAP and the CERF represent two important holistic ways to take needs assessment and planning to the cross-organizational level, through which consolidated allocation of funds and better harmonization of 'who

does what where' is possible. These initiatives seek to tackle the probably most important challenge in humanitarian performance management: still, better needs baselines are required to allow project cycle management that is based on realistic appraisal of what is required on the ground. Only through accurate needs baselines is measurement of humanitarian outcomes possible. Furthermore, an organizational BSC is in place already at the Kenyan Red Cross. Taking into account the findings in section 5.3, further research would be required, to examine how the BSC of the Kenyan Red Cross could be a template to link performance management on the organizational level (meso-level) with project performance management (micro-level) and how to extrapolate this model to other organizations.

As discussed in sub-section 5.2.3, the challenge to attribute outcomes to humanitarian interventions becomes more difficult, the higher the level of performance through which performance is looked at. Therefore, it would be wise to 'aim low' first; in finding ways to assess needs more reliably, not 'blurring' needs assessment by focusing on organizational competencies instead of needs (see sub-section 4.1.1) and assess outcomes on the micro-level, before taking needs and impact assessment to the system-wide level.

6. Conclusion and future outlook

The challenge to improve humanitarian performance is ongoing. It has to be acknowledged that, compared to fields that are set in a more stable and predictable environment, such as the public sector, development aid and for-profit business, the humanitarian sector inherently faces the various difficulties of a highly diverse field in terms of its geographical scattering, its ad-hoc dependence on external factors and the heterogeneous composition of actors. Therefore, it should not be surprising that in today's humanitarian performance management *"the actors and institutions that collectively undertake humanitarian action do not form a coherent and integrated system with shared principles, policies, and modus operandi."* (Stockton 2000 cited in Griekspoor/Sondorp 2001, p. 211) and that performance frameworks and integrated approaches to performance management are called for. But these calls have thus far not been underpinned with solid concepts and models that would allow putting the idea of consolidated performance management into practice. Along these lines, this paper has shown, the BSC cannot be more than a balanced view on performance and therefore is not the cure for the diversity in humanitarian performance management. The uncertainty of humanitarian impact due to the low level of measurability of aid does not allow transposing this business performance management model into a performance management concept for humanitarian aid, not even with seemingly applicable adaptations to its original structure. Some key points may, however be derived from the findings of this paper.

First, humanitarian performance management needs consolidation and professionalization at the lower levels of performance. Along the lines of Richert's finding that for a cross-company BSC, the involved companies require

organizational BSCs, including formulated strategies and goals, prior to the extension of performance management beyond organizational boundaries (Richert 2006, p. 78, see also section 3.2), it is recommendable to overcome the known deficiencies of micro-level performance management, which mainly lie in the structural and procedural flaws in humanitarian project cycle management (as discussed in section 4.1.1), before system-wide performance management is approached.

Secondly, there is a structural gap between performance management on system and organizational level on the one hand, and performance management on project level on the other. This limits the applicability of balanced approaches such as the BSC to the system and organizational levels and requires different means of performance management at the project-level (as discussed in section 5.3). Nevertheless, with clarity and consensus on the mission and objectives of an organization or even the humanitarian sector as a whole, humanitarian project management can benefit from balanced views on organizational- and sectoral performance. The example of the BSC of the Kenyan Red Cross has given a first indication that a balanced view onto organizational performance can help simplifying the planning of humanitarian projects without being overly simplistic (see section 4.1.2).

Third, the role of donors in humanitarian action and consequently humanitarian performance management is essential. In being the customer who 'pays the bill' (as discussed in section 3.3), donors are (probably the only) group of actors who are in a position to bring about change in an obligatory way. Donor commitment to changing approaches in humanitarian performance management is a necessary precondition to a potential system-wide performance management. But their agenda is two-fold. As the HRI has found, donor policies and practice are still dominated by political motives rather than accountability to beneficiaries (see section 4.2). Considering that system-wide performance management would seek to establish 'one truth' on humanitarian needs, the question must be asked whether it is in the interest of donors to help generating this truth, as 'many truths' keep the blending of political agendas with supposed commitment to the impartiality of aid much simpler. In fact, accountability of humanitarian aid is much rather the result of clear donor commitment and according practice than of formal performance management approaches.

It remains to be seen whether the humanitarian sector, with the help of advocacy for donor accountability towards beneficiaries by organizations such as ALNAP, HAP, and DARA will be brought about, and whether balanced approaches can guide the endeavor for system-wide performance management.

7. References

ALNAP, 2011. ALNAP Members. *ALNAP*. Available at: http://www.alnap.org/members.aspx [Accessed November 7, 2011].

ALNAP, 2007. ALNAP Strategy 2008-2013. Available at: http://www.alnap.org/pdfs/alnap_strategy_2008-2013.pdf [Accessed November 7, 2011].

Bakewell, O./Garbutt, A., 2005. The use and abuse of the logical framework approach. Available at: http://www.intrac.org/data/files/resources/518/The-Use-and-Abuse-of-the-Logical-Framework-Approach.pdf [Accessed November 1, 2011].

Bedrup, H., 1995a. Background for Performance Management. In A. Rolstadas, ed. *Performance Management: A Business Process Benchmarking Approach*. London: Chapman / Hall, pp. 61–88.

Bedrup, H., 1995b. Performance Measurement. In A. Rolstadas, ed. *Performance Management: A Business Process Benchmarking Approach*. London: Chapman & Hall, pp. 169–190.

Bouckaert, G., 1996. Measurement of Public Sector Performance: Some European Perspectives. In A. Halachmi/G. Bouckaert, eds. *Organizational Performance and Measurement in the Public Sector: Toward Service, Effort and Accomplishment reporting*. Westport: Greenwood Publishing Group, pp. 223–237.

Bouckaert, G./Halligan, J., 2008. *Managing performance: International comparisons*, London/New York: Routledge.

Bouckaert, G./Halligan, J., 2006. A Framework for Comparative Analysis of Performance Management. In *Study Group on Productivity and Quality in the Public Sector*. Conference of European Group of Public Administration. Milan, Italy 6–9 September 2006. Milan: Università Bocconi [Accessed November 16, 2011].Bouckaert, G./Peters, B.G./Verhoest, K., 2010. *The Coordination of Public Sector Organizations: Shifting Patterns of Public Management*, Basingstoke: Palgrave Macmillan.

Bradach, J.L., 2002. J.L. Bradach on Organizational Infrastructure and Staff Development in Nonprofit Agencies. Available at: http://www.businessweek.com/magazine/content/02_48/b3810001.htm [Accessed November 8, 2011].

Brewer, P.C./Speh, T.W., 2000. Using the Balanced Scorecard to Measure Supply Chain Performance. *Journal of Business Logistics*, 21(1), pp. 75–94.

Bugnion, C., 2002. Analysis of "quality management" tools in the humanitarian sector and their application by the NGOs. Available at: http://portals.wdi.wur.nl/files/docs/ppme/quality.pdf [Accessed November 1, 2011].

Byrne, J.A., 2002. The New Face of Philantropy. *Business Week* no. 2, pp. 82–86.

CE DAT, 2009. Complex Emergency Database. *CE DAT*. Available at: http://www.cedat.be/ [Accessed November 6, 2011].

Crisp, J., 2000. Thinking outside the box: evaluation and humanitarian action. *Forced Migration Review* no. 8, pp. 1–7.

DARA, 2010. *The Humanitarian Response Index 2010 - The problems of politicisation*, Madrid: DARA. Available at: http://daraint.org/humanitarian-response-index/humanitarian-response-index-2010/ [Accessed November 1, 2011].

Darcy, J./Hofmann, C.A., 2003. According to Need? Needs Assessment and Decision-Making in the Humanitarian Sector. *HPG Report* no. 13.

Dearden, P.N., 2001. Programme and Project Cycle Management (PPCM): Lessons from DFID and other organisations. In *Symposium conducted at the meeting of the Foundation for Advanced Studies for International Development (FASID).* Tokyo, Japan February 2001. Tokyo: Centre for International Development and Training (CIDT), University of Wolverhampton.

Dubnick, M., 2005. Accountability and the Promise of Performance: In Search of the Mechanisms. *Public Performance & Management Review*, 28(3), pp. 376–417.

Eccles, R.G., 1991. The Performance Measurement Manifesto. *Harvard Business Review*, 69(1), pp. 131–137.

ECHO, 2005. ECHO Manual - Project Cycle Management. Available at: ec.europa.eu/echo/files/about/actors/.../project_cycle_mngmt_en.pdf [Accessed November 1, 2011].

Erdmann, M.K., 2003. *Supply Chain Performance Measurement*, Lohmar: Eul Verlag.

Eriksson, J. et al., 1996. *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience*, London: Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda.

Fitzgerald, L. et al., 1991. *Performance Measurement in Service Businesses*, London: Chartered Institute of Management Accountants.

De Geuser, F./Mooraj, S./Oyon, D., 2009. Does the Balanced Scorecard Add Value? Empirical Evidence on its Effect on Performance. *European Accounting Review*, 18(1), pp. 93–122.

GHD Initiative, 2003. Principles and Good Practice of Humanitarian Donorship. Available at: http://www.goodhumanitariandonorship.org [Accessed November 1, 2011].

Gladen, W., 2005. *Performance Measurement. Controlling mit Kennzahlen,* 3rd ed., Wiesbaden: Gabler.

Gnaedinger, A., 2007. Humanitarian principles - the importance of their preservation during humanitarian crises. Available at: http://www.icrc.org/eng/resources/documents/statement/humanitarian-principles-statement-121007.htm [Accessed November 9, 2011].

Greaney, P./Pfiffner, S./Wilson, D. eds., 2011. *The Sphere Project - Humanitarian Charter and Minimum Standards in Humanitarian Response,* 3rd ed., London: Practical Action Publishing.

Griekspoor, A./Sondorp, E., 2001. Enhancing the quality of humanitarian assistance: Taking stock and future initiatives. *Prehospital and Disaster Medicine*, 16(4), 209–215.

Hallam, A., 1998. Evaluating Humanitarian Assistance Programmes in Complex Emergencies. *Good Practive Review* no. 7.

Halligan, J., 2006. The Reassertion of the Centre in a First Generation NPM System. In T. Christensen/P. Laegreid, eds. *Autonomy and Regulation: Coping with Agencies in the Modern State*. Cheltenham: Edward Elgar Publishing, pp. 162–180.

HAP, 2011. Humanitarian Accountability Partnership International. *HAP International*. Available at: http://www.hapinternational.org [Accessed November 4, 2011].

HAP, 2010. HAP Standard in Accountability and Quality Management. Available at: http://www.hapinternational.org/pool/files/2010-hap-standard-in-accountability.pdf [Accessed November 4, 2011].

Harmer, A./Cotterell, L./Stoddard, A., 2004. From Stockholm to Ottawa: A progress review of the Good Humanitarian Donorship initiative. *HPG Policy Brief* no. 18.

Harrell-Bond, B., 1986. *Imposing Aid: Emergency Assistance to Refugees*, , Oxford, Oxford Medical Publications.

Hilhorst, D., 2002. Being Good at Doing Good? Quality and Accountability of Humanitarian NGOs. *Disasters*, 26(3), pp. 193–212.

HNTS, 2011. Health and Nutrition Tracking Service. *The HNTS*. Available at: http://www.thehnts.org/ [Accessed November 6, 2011].

Hoffmann, O., 1999. *Performance Management: Systeme und Implementierungsansätze,* 3rd ed., Bern: Haupt Verlag.

Hofmann, C.A. et al., 2004. Measuring the impact of humanitarian aid: A review of current practice. *HPG Report* no. 17.

Horváth, P./Kaufmann, L., 1998. Balanced Scorecard - Ein Werkzeug zur Umsetzung von Strategien. *Harvard Business Manager*, 20(5), pp. 39–50.

IFRC, 2011. Analysis of the Fundamental Principle of Humanity. *IFRC*. Available at: http://www.ifrc.org/en/who-we-are/vision-and-mission/the-seven-fundamental-principles/humanity/ [Accessed November 9, 2011].

IFRC/ICRC, 2010. The Code of Conduct for the International Red Cross and Red Crescent Movement and Non-Governmental Organisations in Disaster Relief. Available at: www.ifrc.org/Docs/idrl/I259EN.pdf [Accessed November 1, 2011].

Johnsen, Å., 2001. Balanced scorecard: theoretical perspectives and public management implications. *Managerial Auditing Journal*, 16(6), pp. 319–330.

Kaplan, R.S., 2001. Strategic Performance Measurement and Management in Nonprofit Organizations. *Nonprofit management and Leadership*, 11(3), pp. 353–370.

Kaplan, R.S./Norton, D.P., 1997. *Balanced Scorecard: Strategien erfolgreich umsetzen*, Stuttgart: Schäffer-Poeschel.

Kaplan, R.S./Norton, D.P., 1996. *The Balanced Scorecard: Translating Strategy into Action*, Boston: Harvard Business School Press.

Kaplan, R.S./Norton, D.P., 1992. The Balanced Scorecard - Measures That Drive Performance. *Harvard Business Review*, 70(1), pp. 71–79.

Keegan, D.P./Eiler, R.G./Jones, C.R., 1989. Are your performance measures obsolete? *Management Accounting*, 70(12), pp. 45–50.

Kennerley, M./Neely, A., 2002. A framework of the factors affecting the evolution of performance measurement systems. *International Journal of Operations & Production Management*, 22(11), pp. 1222–1245.

Kennerley, M./Neely, A./Adams, C., 2007. Performance measurement frameworks: A review. In A. Neely, ed. *Business Performance Measurement: Unifying Theory and Integrating Practice*. Cambridge: Cambridge University Press, pp. 143–162.

Klingebiel, N., 2000. *Integriertes Performance Measurement*, St. Gallen: Deutscher Universitätsverlag.

Klingebiel, N., 1998. Performance Management - Performance Measurement. *Zeitschrift für Planung*, 9(1), pp. 1–15.

Kloot, L./Martin, J., 2000. Strategic performance management: A balanced approach to performance management issues in local government. *Management Accounting Research*, 11(2), pp. 231–251.

Lange, C./Schaefer, S./Daldrup, H., 2001. Integriertes Controlling in Strategischen Unternehmensnetzwerken. *Controlling*, 2(2001), pp. 75–83.

Lawrie, G./Kalff, D./Andersen, H., 2006. Balanced Scorecard and Results-Based Management: Convergent Performance Management Systems. In *3rd Annual Conference on Performance Measurement and Management Control at the European Institute for Advanced Studies in Management (EIASM)*. Nice, France September 2005. Maidenhead: 2GC Limited.

Lee, H.L./Padmanabhan, V./Whang, S., 1997. Information Distortion in a Supply Chain: The Bullwhip Effect. *Management science*, 43(4), pp. 546–558.

Locke, E.A., 1982. The Ideas of Frederick W. Taylor: An Evaluation. *The Academy of Management Review*, 7(1), pp. 14–24.

Lynch, R.L./Cross, K.F., 1992. *Measure up! The essential guide to measuring business performance*, London: Mandarin.

Macrae, J. et al., 2002. Uncertain power: The Changing Role of Official Donors in Humanitarian Action. *HPG Report* no. 12, pp. 1–8.

Mitchell, J., 2008. Collective action and performance - a personal view. In M. Herson/J. Mitchell/B. Ramalingam, eds. *ALNAP 7th Review of Humanitarian Action*, pp. 11–19.

NAO, 2007. World Food Programme - Managing for results: A second review of progress in implementing results-based management. Available at: http://www.mande.co.uk [Accessed November 16, 2011].

NCVO, 2011. Key Learning: QSTG 1997-2004. *NCVO*. Available at: http://www.ncvo-vol.org.uk/qstg [Accessed November 5, 2011].

Neely, A., 1999. The performance measurement revolution: why now and what next? *International Journal of Operations & Production Management*, 19(2), pp. 205–228.

Niven, P.R., 2008. *Balanced Scorecard: Step-by-Step for Government and Nonprofit Agencies*, Hoboken: John Wiley and Sons.

Norrie, J./Walker, D.H.T., 2004. A balanced scorecard approach to project management leadership. *Project Management Journal*, 35(4), pp. 47–56.

OECD DAC, 1999. Guidance for Evaluating Humanitarian Assistance in Complex Emergencies. Available at: http://www.oecd.org/dataoecd/9/50/2667294.pdf [Accessed November 1, 2011].

Omolo, A., 2010. Project Tugure T. Mung'ou, ed. *Reach Out 41 - Quarterly Publication of Kenya Red Cross Society* no. 41, p. 5.

Oxford Dictionaries, 2011. Oxford Dictionaries Online - English Dictionary and Language Reference. *Oxford Dictionaries*. Available at: http://oxforddictionaries.com/ [Accessed October 14, 2011].

Proudlock, K./Ramalingam, B./Sandison, P., 2009. Improving humanitarian impact assessment: bridging theory and practice. *ALNAP 8th Review of Humanitarian Action: Performance, Impact and Innovation*.

Radin, B., 2000. *Beyond Machiavelli: Policy Analysis Comes of Age*, Washington D.C.: Georgetown University Press.

Ramalingam, B. et al., 2009. Counting what counts: performance and effectiveness in the humanitarian sector. *ALNAP 8th Review of Humanitarian Action: Performance, Impact and Innovation*.

ReliefWeb, 2008. ReliefWeb Glossary of Humanitarian Terms. *ReliefWeb*. Available at: http://reliefweb.int/node/23696 [Accessed November 1, 2011].

Richert, J., 2006. *Performance Measurement in Supply Chains: Balanced Scorecard in Wertschöpfungsnetzwerken*, Wiesbaden: Springer.

Rigby, D./Bilodeau, B., 2007. Bain's global 2007 management tools and trends survey. *Strategy & Leadership*, 35(5), pp. 9–16.

Rigby, D./Bilodeau, B., 2005. The Bain 2005 management tool survey. *Strategy & Leadership*, 33(4), pp. 4–12.

Sheldon, O., 1924. Taylor, the creative leader. In D. Del Mar/Taylor Society, eds. *Classics in Scientific Management: A Book of Readings*. Tuscaloosa: University of Alabama Press, pp. 35–51.

Sibbet, D., 1997. 75 Years of Management Ideas & Practice 1922-1997. *Harvard Business Review*, 75(5), pp. 2–12.

Slim, H., 1997. Relief agencies and moral standing in war: Principles of humanity, neutrality, impartiality and solidarity. *Development in Practice*, 7(4), pp. 342–352.

SMART, 2011. Standardized Monitoring and Assessment of Relief and Transitions. *SMART Indicators*. Available at: http://www.smartindicators.org/ [Accessed November 6, 2011].

Stockton, N., 2000. Speaking Notes: *Keeping the International Community Involves in Zones of Conflict No Longer in Fashion*. Taplow Court, 04 March 2000.

Stölzle, W./Heusler, K.F./Karrer, M., 2001. Die Integration der Balanced Scorecard in das Supply Chain Management-Konzept. *Logistik Management*, 3(2/3), pp. 73–85.

Street, A./Parihar, G., 2007. The UN Cluster Approach in the Pakistan earthquake response: an NGO perspective. *Humanitarian Exchange* no. 37, pp. 32–34.

Talbot, C., 1999. Public Performance - towards a new model? *Public Policy and Administration*, 14(3), pp. 15–34

Taylor, F.W., 1994. Scientific management. In F. Fischer/C. Sirianni, eds. *Critical Studies in Organization & Bureaucracy*. Philadelphia: Temple University Press, pp. 44–54.

Taylor, F.W., 1970. What is scientific management? In H. G. Merrill, ed. *Classics in management*. New York: American Management Association, pp. 67–71.

Van Dooren, W., 2009. A Politico-administrative Agenda for Progress in Social Measurement: Reforming the Calculation of Government's Contribution to GDP. *Journal of Comparative Policy Analysis*, 11(3), pp. 309–326.

Van Dooren, W./Bouckaert, G./Halligan, J., 2010. *Performance Management in the public sector*, New York: Routledge.

Weber, J./Schäffer, U., 1998. *Balanced Scorecard*, Vallendar: Wissenschaftliche Hochschule für Unternehmensführung.

Weber, J./Bacher, A./Groll, M., 2002. Konzeption einer Balanced Scorecard für das Controlling von unternehmensübergreifenden Supply Chains. *Kostenrechnungspraxis* no. 46, pp. 133–141.

Werner, H., 2000a. Die Balanced Scorecard im Supply Chain Management (Part 1). *Distribution*, 31(4), pp. 8–11.

Werner, H., 2000b. Die Balanced Scorecard im Supply Chain Management (Part 2). *Distribution*, 31(5), pp. 6–20.

WFP, 2006. Overview of the SENAC project. Available at: http://reliefweb.int/sites/reliefweb.int/files/resources/BD66D7361D99EAEA4925722900028192-wfp-senac-overview.pdf [Accessed November 1, 2011].

Zimmermann, K., 2003. *Supply Chain Balanced Scorecard: Unternehmensübergreifendes Management von Wertschöpfungsketten*, Wiesbaden: Deutscher Universitätsverlag.

IFHV Working Paper Series

URL: http://www.rub.de/ifhv/6-publications/6-workingpapers.html

ISSN: 2199-1367

List of published IFHV Working Papers

Vol. 1, No. 1
03/2011
Huseyn Aliyev
Aid Efficiency in an Armed Conflict
The Role of Civil Society in the Escalation of Violence in the North Caucasus
http://www.rub.de/ifhv/documents/workingpapers/wp1_1.pdf

Vol. 1, No. 2
08/2011
Matteo Garavoglia
Germany's Humanitarian Aid and Media Reporting on Natural Catastrophes
An Investigation of the Relationship between Issue Salience and the Provision of Humanitarian Aid at the Beginning of the 21st Century
http://www.rub.de/ifhv/documents/workingpapers/wp1_2.pdf

Vol. 2, No. 1
05/2012
Jan Wulf
A Balanced Scorecard for the Humanitarian Sector?
Adaptability of the Balanced Scorecard Model to Sector-Wide Performance Management in Humanitarian Aid – Feasibility and Implications
http://www.rub.de/ifhv/documents/workingpapers/wp2_1.pdf

Manuscripts can be submitted to the editors:
Prof. Dr. Pierre Thielbörger, ifhv@rub.de
Dr. Kerstin Rosenow-Williams, ifhv@rub.de

IFHV Working Paper Series

URL: http://www.rub.de/ifhv/6-publications/6-workingpapers.html

ISSN: 2199-1367

Institute for International Law of Peace and Armed Conflict (IFHV)

The Institute for International Law of Peace and Armed Conflict (Institut für Friedenssicherungsrecht und Humanitäres Völkerrecht, IFHV) was established in 1988 by decision of the University Senate as a central research unit ('Zentrale Wissenschaftliche Einrichtung') of the Ruhr University Bochum. The IFHV is responsible directly to the Rector and the Senate of the Ruhr University Bochum, but works in close cooperation with the different faculties, in particular the faculties of law, social science, geosciences and medicine.

The IFHV carries out research and teaching on the problems of peace and armed conflict from an inter-disciplinary perspective. Based on its strong international humanitarian law tradition, the IFHV is the only institute in Germany, and one of very few in Europe and the world, which is dedicated to the discipline of humanitarian studies. The IFHV combines its strong emphasis on international humanitarian law, the law of peace and human rights law with sociological and political perspectives on humanitarian crises, actors and activities.

IFHV Working Paper Series

In 2011, the IFHV and the Ruhr University Bochum decided to set up an IFHV Working Paper Series in humanitarian studies. In line with the IFHV's multidisciplinary profile, we intend to publish a broad range of papers in the field of humanitarian studies. Our Working Paper Series publishes 'work in progress'. The Working Paper Series intends to stimulate the humanitarian discourse, contribute to the advancement of the knowledge and understanding of the practices, policies and norms of humanitarian action, and last but not least seeks to attract comments, which improve the content of the working paper for further publications.

The Working Paper Series allows IFHV staff and students, and like-minded researchers in the field of humanitarian studies to bring their work and ideas to the attention of a wider audience. In order to publish high level working papers, the papers offered for publication will be technically screened by the editors of the working paper series and subjected to an internal blind peer review process.

Contact:

Institute for International Law of Peace and Armed Conflict (IFHV)
NA 02/33
Ruhr University Bochum
Universitätsstraße 150
44780 Bochum
Germany

Telephone: +49(0)234 32 27366
Fax: +49(0)234 32 14208
Email: ifhv@rub.de
Web: www.ifhv.de
Facebook: www.facebook.com/rub.ifhv