Key Performance Indicators in Humanitarian Logistics

by

Anne Leslie Davidson

B.S. Accounting and Information Systems, B.A. French
Virginia Polytechnic Institute and State University, 2002

Submitted to the Engineering Systems Division in Partial Fulfillment of the Requirements for the Degree of

MASTER OF ENGINEERING IN LOGISTICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2006

Signature of Author .....
Engineering Systems Division
May 10, 2006

Certified by .....
Edgar Blanco
Research Associate, MIT Center for Transportation and Logistics
Thesis Supervisor

Accepted by .....
Yosef Sheffi
Professor, Engineering Systems Division
Professor, Civil and Environmental Engineering Department
Director, MIT Center for Transportation and Logistics

Key Performance Indicators in Humanitarian Logistics

by

Anne Leslie Davidson

Submitted to the Engineering Systems Division
on 12 May 2006 in Partial Requirements for the Degree of Master of Engineering in Logistics

Abstract

Non-profit humanitarian relief organizations have typically been unable to measure the performance of their supply chains due to an inability to centrally capture data from operations. With the recent development and implementation of information technology systems that can support the logistics function of these organizations, the data is now available to measure performance, but what is still lacking is a central framework of metrics that measures performance according to the organization's strategic goals.

First, this thesis reviews the best practices noted in performance measurement systems of the logistics functions in military and commercial organizations and applies them to the humanitarian sector. Second, the thesis suggests a framework of key performance indicators to be implemented in an international non-profit humanitarian organization based on the unique strategic goals of the sector. The thesis then applies this proposed framework to two actual operations performed by this organization. The analysis performed herein proves that a measurement system would help strengthen the organization's ability to deliver goods to beneficiaries more efficiently and effectively. Finally, the thesis addresses feasibility issues of implementing a measurement system in the non-profit sector and also describes the next steps of opportunities related to measurement systems within humanitarian logistics.

Thesis Supervisor: Dr. Edgar Blanco
Title: Research Associate, MIT Center for Transportation and Logistics

Acknowledgements

The completion of this thesis would not have been possible without the support and valuable input from the following people: Dr. Edgar Blanco of MIT, Dr. Laura Kopczak of the Tuck School at Dartmouth and the Fritz Institute, Martin Bush of the International Federation of the Red Cross and Red Crescent Societies, and Mitsuko Mizushima of the Fritz Institute. I would also like to thank my fiancé, my family, and my fellow MLOG students for their continued support throughout this program.

Biographical Note

Anne Davidson completed a Bachelor of Science degree in Accounting and Information Systems, as well as a Bachelor of Arts degree in French from Virginia Tech in August of 2002. Following graduation, she became a Certified Public Accountant while working in Assurance & Advisory Services with Deloitte & Touche, LLP, located in Richmond, Virginia. She returned to Blacksburg, Virginia, in August of 2004, to serve as a campus ministry intern with the Baptist Collegiate Ministry at Virginia Tech. During this internship on a disaster relief construction trip, it was then that she discovered the challenging and fascinating field of humanitarian logistics.

Table of Contents

List of Tables

List of Figures

1 Introduction

When disasters strike, humanitarian organizations respond. When disasters strike and organizations don't respond fast enough in the eyes of the media and the local government, the blame that is placed on these organizations tarnishes their name and reputation, which affects their base of donors. Without the financial support of donors, a non-profit humanitarian organization's entire ability to continue operations is placed in danger.

Therefore, actual performance of operations after a disaster, as well as the *perception* of the organization's performance, are both clearly essential to "humanitarian logistics": the function of a non-profit organization that oversees "delivering the right supplies to the right people, at the right place, at the right time, and in the right quantities" (Russell, 2005). The question then becomes, "How well did the organization actually perform?"

Given the volatile and variable environment of disaster occurrence, answering this seemingly simple question becomes a very complex problem. The need for standard performance metrics in humanitarian logistics is a deeper issue than merely the media's perception of poor or slow performance after a disaster. It is a very real need that has been acknowledged in recent studies by the Fritz Institute after the 2004 tsunami in South Asia (Thomas & Ramalingam, 2005) and in the "*Humanitarian Response Review*" commissioned by the United Nations Office for the Coordination of Humanitarian Affairs (OCHA, 2005).

The question of how to measure the performance of an organization's response to a disaster has surfaced frequently in the past two years, mainly because of the following three

high-profile disasters: the 2004 tsunami in Asia, the 2005 earthquake in South Asia, and Hurricane Katrina in 2005. However, the lack of performance metrics has been a standing problem in the humanitarian sector long before these disasters occurred. Several factors cause the performance of humanitarian organizations to be both difficult to define and to measure:

- Lack of centrally-captured data from operations
- Limited information technology infrastructure
- Lack of funding for IT infrastructure
- Variability / chaotic environment after each disaster
- External factors (e.g. geography, state response, etc.)
- Lack of incentive for measurement in non-profit sector
- Potential negative media exposure
- Human resources issues
- Organizational culture
- Long-term vs. short-term goals of disaster response

Although the above factors are significant obstacles in overcoming the problem of performance measurement, the hypothesis of this research is that measurement standards *can* be defined and implemented to improve the performance of disaster relief operations. Significant research has been performed relating to performance measurement of the logistics function in both the private and military sectors, and many of these operations have been transformed through the implementation of performance measurement systems. It is the goal of this research to draw from that knowledge and apply it to the humanitarian sector.

The remainder of this thesis is structured into five chapters. Chapter 2 will explore the current context of the humanitarian sector with regard to implementing a system of performance

measurement and address the question, "How willing and able is the sector as a whole to implement such a system?" Chapter 3 documents the best practices of performance measurement systems that have been implemented in business and military logistics operations and describes how these principles can appropriately be applied to disaster relief operations. Chapter 4 describes the business processes of a disaster operation at the International Federation of the Red Cross and Red Crescent Societies (IFRC) and proposes a framework of key performance indicators to be used by the logistics department of this organization. To demonstrate how this framework would be used and interpreted by the IFRC, the scorecards are presented to recreate a walk-through of two different IFRC emergency operations. Chapter 5 tests the proposed IFRC framework by giving a more in-depth background and analysis of these two operations. The performance of the logistics function in these two operations is compared and contrasted in terms of the metrics used in the proposed framework. Chapter 6 discusses how the IFRC framework can be applied more generically to other non-profit humanitarian organizations. It also discusses feasibility issues of implementing a framework of key performance indicators. Chapter 7 concludes by describing opportunities for future research that will further promote the use of key performance indicators in the humanitarian sector.

2 Understanding the Context of the Humanitarian Sector

2.1 Industry Climate: Recognizing the Need for Metrics

There are several indications that the culture of the non-profit sector is on the brink of being able to embrace a quantitative system of performance measurement. In the past, goals for humanitarian operations have been set very qualitatively, which makes it difficult for these organizations to truly gauge their performance upon the completion of an operation. Understandably, due to the chaos in which most workers find themselves in the midst of a relief operation, there has been much more emphasis placed in non-profit organizations on the necessity to deliver goods and services as quickly as possible to beneficiaries rather than emphasizing the need to perform the paperwork and review operational performance (Taylor, 1997). However, when the record-keeping suffers, knowing with confidence what the current inventory levels are as well as knowing when future deliveries will arrive becomes much more difficult, which often leads to an environment of intense frustration for many relief workers on the ground.

Enabling better supply chain visibility would alleviate many of these frustrations. The only way to support such visibility involves two key steps: to establish systematic organizational processes **and** to define benchmarks of performance. These steps should standardize the flow of the relief chain as much as possible, although it is important to note that all variability can never

be eliminated due to the very nature of responding to disasters. It is also important to highlight that while the focus of this paper is to define the key performance indicators that enable an organization to determine what constitutes an effective response, it is absolutely imperative that standardized business processes are in place throughout these organizations as well. Oftentimes in the humanitarian sector, business processes are either not pre-defined for individuals to follow or they are not enforced, which leaves the action up to the individual's discretion. The importance of standard and clearly defined business processes and lines of authority in these organizations cannot be stressed enough as an essential part of gaining visibility and standardization in the relief supply chain.

The need for benchmarks and performance measurement is also augmented as the scope and size of an operation increases. Such was the case during and after the response to the December 2004 tsunami. As mentioned earlier, the need for establishing benchmarks of an effective operation was one of the top-line findings of a recent report from the Fritz Institute (Thomas & Ramalingam, 2005). In this operation it was found that while the beneficiaries themselves may serve as an important source of gauging operational performance, it is usually subjective and should be used with caution. Because the beneficiaries' perceptions may be affected by various factors such as cultural expectations, the unique circumstances of the disaster/event, and how the aid is actually administered, organizations should rely on a more objective and standard basis of performance measurement.

2.2 UN Efforts Towards Global Benchmarks

A similar recommendation was made in a 2005 report commissioned by the United Nations Office for the Coordination of Humanitarian Affairs. This review highlighted a number of long-standing gaps already known to be present in the humanitarian sector, particularly in the areas of preparedness, benchmarking, and coordination. Specifically related to benchmarking, while OCHA recognized the difficulty of gaining consensus around objectives from various international organizations, it nonetheless recommended that a global set of benchmarks and indicators be created and implemented to measure performance. It notes that "other key recommendations cover, first and foremost, the development and application of benchmarks and indicators to measure performance" (OCHA, 2005, p12). While this is clearly a powerful indication of a cultural shift towards performance measurement from an organization at a global level in the humanitarian sector, the complexity of following through with this recommendation to actually develop and implement any such system remains.

The OCHA report itself actually highlights the issue of vague quantitative benchmarks and performance goals that have typically characterized the humanitarian sector, most surely without intending to do so. It indicates as one of its main recommendations that benchmarks and indicators should address the following:

- access and coverage of population in need,
- identification of responsibilities in delivery of assistance and in coordination,
- resources mobilization (e.g. human, assets, financial),
- identification of relevant lifesaving activities, and
- protection aspects, where needed (UNOCHA, 2005, p16).

The report also indicates that the needed elements of benchmarks should relate to "speed,

quantity, relevance, appropriateness, and sustainability" (UNOCHA, 2005, p21). The gap in the analysis is that the possible indicators that are listed on pages 23-25 of the report do not cover all of the bulleted objectives listed above, nor do they relate to all of the needed elements of benchmarks that were given. For the possible indicators that have been listed in the report, there is no real structure around them so that organizations can build a framework of metrics that will enable performance measurement, even though the report acknowledges that such a framework is necessary.

The other questionable finding from the OCHA report is its repeated focus on donor-based benchmarks. The underlying theme related to benchmarking in the report is that donors should be stringently involved in both the development and the monitoring process of benchmarks used in humanitarian organizations. The research performed herein will take a different approach to this idea, proposing that performance indicators should first be developed internally by an organization and later introduced as external indicators for donors. The reason for this difference is because the research findings from the private sector related to performance measurement strongly indicate that *it is the organization itself which must define its operational goals based on its strategic objectives. Any metric framework introduced should first flow from the organization's goals and not from any goals imposed by an outside party.* This idea will be explained further in Chapter 3.

2.3 Development of Software for Humanitarian Organizations

Another shift that is gradually occurring in the humanitarian sector is the development and implementation of software packages that are specialized to humanitarian logistics. Such software packages can provide visibility in the relief supply chain since they can capture the data

of an operation in a centralized location. The lack of complete operational data has long been the single largest inhibitor of standardized performance measurement, so the development and implementation of logistics software is truly a giant leap forward for the humanitarian sector. In 2003, the logistics department of the IFRC implemented Humanitarian Logistics Software (HLS), which was developed jointly with the Fritz Institute. HLS captures data and tracks information related to the procurement and distribution of goods to field personnel. It has been in use for all major emergency IFRC operations since late 2003. This system was the data source for the research performed in this project and documented in Chapters 4 and 5.

Performance measurement for the entire humanitarian sector may still be a long way off, but the trends and motivation exist to begin investigating these possibilities today. The need for benchmarks has been acknowledged after the three well-publicized disaster response operations of the past two years. There is still a gap between motivation and execution, though, until more organizations begin to implement computerized logistics systems. How quickly other organizations implement such systems will most certainly affect when it will become standard for the sector as a whole to use performance metrics.

3 Best Practices of Performance Measurement Systems Used in Business & Military Logistics

3.1 Evaluating the Applicability of Metric Frameworks Used in Business and Military

3.1.1 Military Logistics and "Velocity Management"

In many ways military logistics are more similar to international humanitarian logistics than supply chains designed for the business sector. Military supply chains and humanitarian relief chains share the following:

- Both organizations must be flexible enough to create and establish a working supply chain in a given geographic region with very short notice and/or time for operation-specific preparation.
- Both kinds of operations are working in potentially chaotic and hazardous environments.
- Neither type of organization are working to make a profit from its operations, so they "measure success" differently than a business. For the military, a successful mission would be

one that accomplished its goal, rather than to show an increase in Net Income or a decrease in Total Costs.

A key characteristic of military supply chains is that they must be both agile and flexible (Edmonson, 2005). Edmonson acknowledges that a military supply chain will never be able to achieve the kind of "lean" efficiency found in commercial supply chains because of the higher degree of flexibility a military operation requires. This is the trade-off found due to operating in times of emergency, which both military and humanitarian organizations are forced to do on a regular basis. Before the commercial focus on supply chain management during the past 10-15 years, the military used to surpass the commercial sector in terms of supply chain technology, but now it actually lags behind in cutting-edge initiatives such as the use of RFID (Radio Frequency Identification) to track the flow of goods in supply chains. Edmonson's assessment of the military's current progress in supply chain management also validates the importance of researching commercial supply chain best practices as part of the search for key performance indicators, even though military supply chains do share more in common with humanitarian supply chains than those found in the commercial sector.

One best practice applicable to humanitarian organizations that Edmonson notes is the military's focus on packaging items for the end user as early as possible in the supply chain. Although sometimes humanitarian organizations are able to implement this principle, the receipt of unsolicited goods often presents a problem that is not faced in the military sector. Humanitarian organizations typically receive a variety of unsolicited goods after a disaster. Edmonson cites Dalton Cunningham, a major in the Salvation Army: "When donations come in unsolicited, it's generally a mix-match of multiple items that people believe are needed. All of it has to be sorted, separated, and then it can be palletized, shrink-wrapped and distributed to the

areas where the need is" (Edmonson, 2005). Despite the issue of unsolicited goods, many organizations do pre-package "kits" of items upstream in the supply chain to create hygiene kits or kitchen sets, which does save time during the distribution phase.

The difficulty for a humanitarian organization then becomes having all the necessary items on hand for a one-time distribution to a family, for example, that is in need of multiple *types* of items (e.g. a kitchen set and a hygiene kit). Obtaining such efficiencies in the distribution process requires a high level of coordination between the various departments of an organization, particularly between the logistics department and the field personnel. The field personnel must communicate what various types of goods need to be distributed to beneficiaries at a specific time, and the logistics department must be able to have them all delivered before the actual distribution, to prevent the redundancy of having multiple distributions to the same beneficiaries.

Another best practice from military logistics is the customer focus of "Velocity Management," which is a performance measurement initiative that was implemented in U.S. Army supply chains in the late 1990's. As noted in study conducted by the RAND Arroyo Center for Research, Velocity Management is a streamlined version of the Six Sigma approach that uses structured methodologies to measure and improve performance (Dumond, 2000). This approach was adopted to improve customer response times in the supply chains, so that goods can flow to the end user as quickly as possible.

The Velocity Management initiative centers around the processes of "Define, Measure, and Improve." The various processes involved in Army logistics are to be explicitly defined and scoped, then measured using appropriate metrics, then analyzed to be improved upon. The initiative is structured to be an iterative process so that improvement continues over time. RAND

noted in its research the value gained simply from performing group walkthroughs of the various processes, because it forced communication from different divisions involved in a given process and facilitated collective insights about each process. RAND also noted that the "most critical aspect of measurement is the development and implementation of appropriate metrics that span the full process and reflect key customer values" (Dumond, 2000).

Another key insight from Velocity Management is that it includes the use of multiple metrics that cover the various dimensions of time, quality, and cost. These metrics look at the median values and not just the mean, because a primary goal of the initiative is to reduce variability in the supply chain as well as find the trade-off between time, quality, and cost. The idea of analyzing the median values can and should be directly applied to humanitarian logistics measures. With the inherent variability that disaster situations naturally include, humanitarian organizations should seek to reduce the variability of their own performance as much as possible in order to streamline their supply chains.

Velocity Management also establishes a "baseline performance level" as part of its measurement process. It is unclear if this should be applied in all areas of humanitarian logistics because baseline performance may vary significantly based on the type of disaster, geographic region, and scope of the disaster. Determining reasonable baseline levels of performance may be the next phase of implementing key performance indicators in humanitarian logistics. First, over-arching performance goals for all disaster operations should be determined, and then baseline measures could be determined at a more detailed level for various disaster types in various regions. This could potentially be a deep area for future research.

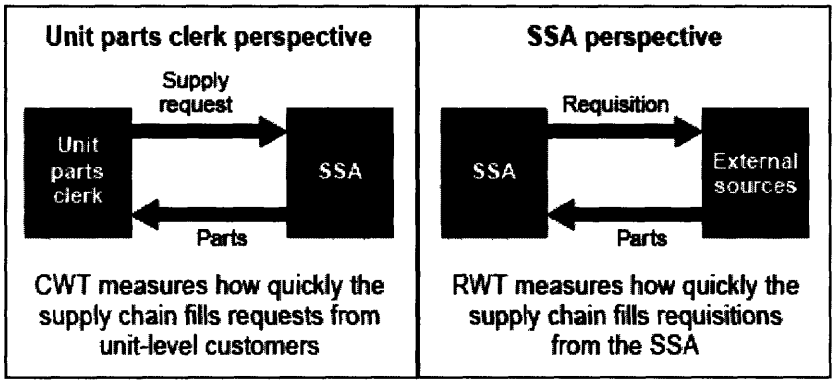
RAND also notes that the data should be analyzed in both short and long-term feedback cycles in order to gain a deeper understanding of the actual trends of performance. This principle

is also applicable to humanitarian logistics. Organizations should define what specific time periods are implied by "short-term" and "long-term" after a disaster, and performance should be measured and analyzed at both levels in order to find the true areas of improvement. This is important because examining operations over time may reveal counter-intuitive insights. For example, procurement times may actually increase over time if the disaster was of a magnitude to cause a regional or even international shortage of a given product. The 2005 Pakistan earthquake caused this to happen, because its magnitude triggered a worldwide shortage of tents and temporary shelters. With that knowledge in mind, one would expect procurement times to have increased over time after the earthquake, instead of decrease, which would be more typical of the time to procure a given item. In short, to truly understand what is happening in the supply chain, an organization must measure itself across both time-frames of the short and long-term, however those are defined.

Two specific metrics that have been key to improving supply chain performance as part of the Velocity Management initiative are Customer Wait Time (CWT) and Requisition Wait Time (RWT), (Brauner & Lackey, 2003). The RAND Arroyo Center for Research conducted a study specifically on these metrics and how they improved the U.S. Army's supply chain for spare parts. These metrics both measure order fulfillment time, but they measure it from different levels of perspective. CWT measures the total time it takes to satisfy a request for a part needed to make a repair, from the initial customer request to the time the request has been received by the customer. RWT measures the total time it takes for a request to be fulfilled by the Supply Support Activity (SSA). The SSA is the "military analog of a retail parts store," so the purpose of this metric is to determine how much time is required for an individual SSA to satisfy a requisition.

Figure 1 from the article represents the different purposes of these metrics:

Figure 1 – CWT vs. RWT Metrics



CWT focuses on supply requests. Its variability depends on the amount of stock supplied on-hand by the SSA, as well as the time it takes to order and receive goods not held on hand by the SSA. The Army examines the variability of CWT by looking at the mean, the median, and they also group the data in various percentiles in order to determine what parts are caused to be delayed in the supply chain.

RWT focuses on how well a supply unit in a specific location can meet its demands. It evaluates how well the logistics chain for spare parts serves the SSA: how much time is required to satisfy an SSA requisition for a part. Like CWT, RWT measures percentiles to not only examine the speed of the supply chain but also its reliability.

After implementing these metrics, the Army then set specific improvement goals based on the performance revealed by these metrics, both for averages and for percentiles. The RAND research noted that these metrics have enabled the Army to dramatically reduce wait times since 1995, when these systems were first implemented. For example, average RWT was 22 days in 1995. RWT had dropped to 10 days in 2002, with a goal to improve to 6 days after 2002.

Both of these metrics appear to be very applicable to humanitarian logistics, given that information systems are put in place which can support capturing this data. The benefit of using CWT and RWT is that it measures at not only the customer level, but also the service unit level. For an international non-profit organization, the analog of a customer would be the beneficiary and the analog of a service unit could be a storage warehouse on-site. CWT could help the organization measure the entire length of time it takes for goods to be delivered to the beneficiary, and RWT could help an organization isolate which geographically-specific warehouses have the most delays in the supply chain. Applying these metrics directly to humanitarian logistics will be discussed further in Chapter 4.

3.1.2 The Balanced Scorecard Approach

The Balanced Scorecard Approach originally emerged in 1992, as a framework for performance measurement in the *Harvard Business Review* (Kaplan & Norton, 1992). The Balanced Scorecard Approach forces managers to select only a small number of critical measures by which they can gauge performance. The Balanced Scorecard also excels in forcing organizations to focus on two key issues: first they must examine what their core competencies are, and second, they must look beyond purely financial metrics.

The Balanced Scorecard organizes the metrics into four perspectives that are addressed by asking the following questions, which should provide managers with a more "balanced" perspective of performance than just the proverbial "bottom-line results."

1. **Customer perspective**: How do customers see us?
2. **Internal Business perspective**: What must we excel at?
3. **Innovation & Learning perspective**: Can we continue to improve and create value?
4. **Financial perspective**: How do we look to shareholders?

The next question to therefore ask is, "Can the Balanced Scorecard be successfully applied to humanitarian logistics?" While the Balanced Scorecard was originally designed to be implemented in for-profit businesses, it has been successfully implemented in other non-profit organizations, such as local governments (Kreklow & Ruggini, 2005). Kreklow & Ruggini provide insight into the history of performance measurement in governmental systems and indicates that performance measurement in earnest began when local governments started to understand the importance of focusing their attention on customer service and product quality. To meet these goals, some city and local governments have used the questions above to implement their own version of the Balanced Scorecard approach. Hence, the transition from for-profit to non-profit industries using the Balanced Scorecard approach *is* possible.

The clearest benefit of using the Balanced Scorecard approach in humanitarian logistics is the importance it places on customer service. In the humanitarian sector, the customer's perception of service and the media's portrayal of this service greatly affect an organization's reputation and potential donations, so the customer perspective is critical in this sector. Because the Balanced Scorecard incorporates the customer perspective as a significant part of its framework, it follows that the Balanced Scorecard should, in theory, provide a framework of metrics that is consistent with a humanitarian organization's goals.

One key feature to remember about the Balanced Scorecard that was noted in the original Kaplan & Norton article is the importance of information systems (Kaplan & Norton, 1992). They noted that information technology systems play a critical role in helping managers implement a measurement system. This is a significant area of weakness in the humanitarian sector, because as described in Chapter 2, IT systems are often lacking the relief supply chain of various organizations. In other words, a performance measurement system is only as valid as the

data that it captures, and so if there is limited data being captured by the IT system or there are no IT systems in place, a performance measurement system like the Balanced Scorecard approach is most likely not feasible.

In 2002, an INSEAD study was performed to evaluate both the feasibility and the value of implementing a balanced scorecard at a selected division of the IFRC (Guyoton & Muirhead, 2002). Guyoton & Muirhead elaborate on many of the current systems of performance measurement, which are very informal and often anecdotal. They conclude that implementing a form of the balanced scorecard would very valuable at the division examined (the Disaster Management and Coordination Division), but there are significant challenges such as executive support, IT infrastructure, and personnel retention issues, which must be overcome to achieve a successful implementation. They conclude with a suggested framework of metrics along the four Balanced Scorecard perspectives, but none of these metrics are really aimed at measuring supply chain performance or operational performance immediately after a disaster. They are more aimed towards the strategies of preparation, personnel retention, and knowledge-sharing. They do support the conclusion, though, that given a humanitarian organization with proper IT infrastructure and management support, the Balanced Scorecard could be implemented, provided that the metrics selected are quantitative enough to reveal where the true opportunities for improvement can be found.

3.1.3 Supply Chain Operations Reference Model (SCOR Model)

The SCOR model is a framework that groups the phases of supply chain management into 5 management processes: plan, source, make, deliver, return. The SCOR model breaks down each management process at various organizational levels and establishes metrics at each of these levels (SCC, 2004). While this model provides enough detail and choices of metrics that it

may be adaptable to the humanitarian supply chain, it is not the best choice of metric systems mainly because its complexity could hinder its ability to be implemented in a humanitarian organization.

The SCOR model would also have to be analyzed on a "per disaster" basis in order to be implemented. According to Beamon, businesses operate in a static world where supply chains and distribution networks are pre-defined and planned before demand from the consumer occurs. Humanitarian organizations may have some pre-defined distribution networks and agreements with vendors in place before disaster strikes, but nearly all operations contain some kind of "ad-hoc" elements to the supply chain, given the unpredictability of operation timing, location, needs, etc. (Beamon, 2004). Therefore, the management processes of the SCOR model may vary per item within an organization.

For a simple example, a flood occurs in South Asia and a humanitarian organization learns that the two items most needed from them are water purification kits and sandals. This organization has a pre-existing relationship with a regional vendor to provide the purification kits in South Asia in case disaster strikes there. The same organization has no pre-existing agreement with any regional vendors for sandals. Using the SCOR model, which emphasizes the "Plan – Source – Make – Deliver – Return" timeline, the point of customer demand for these two items occurs during different phases of the creation of the supply chain. In the case of the purification kits, the point of demand occurs after the sourcing has already been determined. What remains for the organization is to activate this supply chain and deliver the goods to the customers. In the case of sandals, since no pre-existing relationship exists, the supply chain has not fully been established yet. Hence, the point of demand occurs during the "Plan" phase of the supply chain and the items must first be sourced before they can be delivered to customers.

Based on the above example, the SCOR model itself provides a framework that is perhaps too rigid for use in humanitarian agencies, given the inherent variability of their operations and supply chains. The SCOR model does, however, contain numerous metrics that may be of use for the humanitarian sector. Such metrics include the following, which will be re-visited in Section 4.1.2.

- Receiving Cycle Time
- Verification Cycle Time
- Transfer Cycle Time
- Transfer / Product Storage Costs as a Percentage of Product Acquisition Costs
- Delivery Flexibility

3.2 Principles of Performance Measurement

3.2.1 Best Practices of Performance Measurement Systems

Because the commercial sector has been using performance measures as part of their management accounting systems for the past 75 years (Johnson & Kaplan, 1987), there is a wealth of literature written on this topic, specifically about the best practices that have been found in companies that tend to perform well.

The most obvious and most important best practice related to performance measurement is that the metrics selected must be aligned with the organization's over-arching goals (Lambert, 2001). Because this point has been noted again and again in both literature and in actual presentations related to supply chain management given by Fortune 500 companies at MIT, this point cannot be stressed enough. If a metric is not critical to fulfilling your specific organization's goals, then it should not be included in your system of performance measurement.

Organizations should therefore begin by analyzing what their true goals are as an organization and determining what attributes of a supply chain are most valuable in helping them achieve those goals. For example, a humanitarian organization should determine which is more important to fulfilling its goals: Is it more important to get goods to beneficiaries as fast as possible or to get them the right quantity of the most-needed goods? While the initial answer by an organization to this question may be "both," there is an inherent trade-off between one and the other. Speed of delivery to beneficiaries requires a faster initial assessment and fast communication upstream in the supply chain to request goods. Conversely, focusing on accurate quantities of items requires more detailed assessments on the ground, which inherently requires more time. Hence, fulfilling the second goal will decrease the chances of meeting the first one.

The best way to solve the above dilemma is to determine what the organization's true goals and core competencies are. One principle of the Balanced Scorecard Approach is to select only those key metrics that help you achieve your core competencies. If too many metrics are selected, the scorecard can become too cluttered and too cumbersome to truly gauge performance. There is a tendency in designing performance measurement systems of "more is better," but this is simply not true. As explained below, measurement systems require trade-offs of all kinds, so with each metric that you add to the system, there must be some kind of responsive action somewhere else.

The second best practice noted from industry is that the trade-offs between various metrics selected in a measurement system should be taken into account when designing the system, and these trade-offs can occur across multiple dimensions (Caplice & Sheffi, 1994). The authors in their study noted the top eight criteria that had been recorded from other previous studies of logistics metrics. Therefore, the following list captures many ideas from prior work

related to metrics and what to consider before including a specific metric in a measurement system:

- **Validity**: The metric accurately captures the events and activities being measured and controls for any exogenous factors.
- **Robustness**: The metric is interpreted similarly by the users, is comparable across time, location, and organizations, and is repeatable.
- **Usefulness**: The metric is readily understandable by the decision maker and provides a guide for action to be taken.
- **Integration**: The metric includes all relevant aspects of the process and promotes coordination across functions and divisions.
- **Economy**: The benefits of using the metric outweigh the costs of data collection, analysis, and reporting.
- **Compatibility**: The metric is compatible with the existing information, material, and cash flows and systems in the organization.
- **Level of detail**: The metric provides a sufficient degree of granularity or aggregation for the user.
- **Behavioral soundness:** The metric minimizes incentives for counter-productive acts or game-playing and is presented in useful form.

Some of the key trade-offs to consider from the above criteria are among the first four listed. The more valid and specific a metric is to a particular issue, the less robust it will be, because it will be less comparable to other issues. The more integrative a metric is across various functions or processes, the less useful it will be, because it will not point to one specific corrective action to be taken. Both of these trade-offs are applicable to humanitarian logistics

metrics. For example, metrics that incorporate specific actions related to one particular type of disaster may want to be avoided because although they would be more *valid* to this type of disaster, they would be less comparable and therefore less *robust* in measuring the organization's overall performance. There is also a secondary trade-off to note between usefulness and validity: the more valid a metric is, in that it accounts for all activities and exogenous factors, the less likely it will be understandable for decision-making, or the less useful it will be.

The third best practice noted from industry is the idea that measurement systems should be evaluated at the system-level, in addition to the metric-level (Caplice & Sheffi, 1995). While the discussion above centered around how to select specific metrics for use in a measurement system, once the metrics have been determined the entire system should be examined to ensure that the organization is indeed measuring what they were seeking to measure, and the system will drive behavior in the direction of their strategy. The authors note that a performance measurement system must be "cohesive, comprehensive, and complementary." The term "cohesive" indicates that the measurement system should be both vertically integrated to the overall firm strategy and horizontally integrated to measure *processes* that include various activities, functions, and departments, rather than measuring by function alone. The term "comprehensive" indicates that a measurement system should capture performance from more than one perspective, or from all relevant stakeholders. To apply this principle to a humanitarian organization, this means that stakeholders such as beneficiaries and donors should be considered when designing a broad set of performance measures. Finally, the term "complementary" refers to the trade-offs of the individual metrics noted in the previous paragraph.

In short, between the phase of design and implementation phases of a performance measurement system, an organization should first go back and re-evaluate the proposed system

as a whole. Only from this perspective can the organization see if they are capturing what they need to capture and if the set of metrics align with the organization's overall goals. The organization should also go back and consider what their measurement system is *not* capturing as well. Through this type of review process, an organization may discover an issue of importance to stakeholders that had not been previously included.

Another best practice noted from industry is there exists a need to change the targets of the metrics periodically in order to continue the improvement process as well as to prevent people in organizations from simply "gaming the metrics" (Meyer 2005). Meyer takes a more pessimistic viewpoint about measurement systems in that he believes that "any definition of performance that makes it easy to measure would lock firms into rigid and maladaptive patterns of behavior." For the purposes of humanitarian logistics, however, so little has been actually implemented regarding quantitative performance metrics that this caution is not immediately applicable, but his point is well-taken. Because it is people that comprise any organization and people tend to adapt their behavior to deliver "what is measured rather than the performance that is sought but measured imperfectly," it is advisable to humanitarian organizations that they revise their targets regularly in order to prevent the metrics from becoming obsolete or irrelevant.

Meyer also brings to light the interesting issue of the effect of time on performance metrics. He writes that, "most performance studies do look forward, but to do so, they shift the time frame backward and look forward to the present from the vantage of the past." This insight is extremely applicable to humanitarian organizations, because if only project "post-mortems" are performed after an operation has been completed, the organization can only apply knowledge to future operations based on past operations. It would therefore be advisable to design a system that will measure the performance of the supply chain *during* operations while goods are still

being delivered to beneficiaries. Implementing a "dashboard" set of metrics to monitor current performance during an operation would alleviate some of the problems created by only looking at performance from the vantage of the past.

3.2.2 Current Trends in Supply Chain Management Relevant to Performance Measurement

There are two current trends in supply chain management today that stand out as the most applicable to humanitarian supply chains. The first is the important trend of building flexibility into commercial supply chains, and the second a broad trend of supply chains being used as a competitive advantage to aid companies in the pursuit of their strategic goals.

Clearly, the elements of flexibility and resiliency in supply chains have received ample attention in supply chain management within the past few years due to threats of terrorism and natural disaster. Sheffi proposes several strategies for making supply chains more resilient after documenting numerous examples of companies whose operations were dramatically interrupted due to such unexpected events (Sheffi, 2005). While the advantages of having a flexible supply chain that can respond to a variety of situations are obvious under these conditions, incorporating flexibility which can be measured into a supply chain poses both an important yet complex challenge (Beamon, 1999). Beamon indicates that supply chain performance metrics should center around measuring **resources**—metrics that aim for efficiency, measuring **output**—metrics that aim for high customer service, and measuring **flexibility**—metrics that target the ability to respond to a changing environment. Beamon further breaks down flexibility metrics into two categories of "range flexibility" and "response flexibility". The former measures to what extent the operation can be changed (for humanitarian operations this could mean by disaster type, geographic region, etc.). The latter measures how easily the operation can be changed,

presumably measured by the additional amount of time and money it takes to respond to a changing environment.

The second key trend that appears across in the commercial sector in supply chain management is the idea that a streamlined and well-coordinated supply chain remains for many businesses to be an untapped wealth of competitive advantage. Lambert (2001) writes that what most companies refer to as "supply chain metrics" are in actuality mostly internally-focused logistics measures such as lead time, fill rate, or on-time performance. Lambert proposes more holistic measures that require not only coordination between parties, but a true integration between the parts of the supply chain in order to measure its performance functioning as a whole. He indicates that in order for this kind of "global optimization" to occur, it requires parties to align both their activities and their strategies in order to create the ideal supply chain, given a particular industry and market position. While Lambert's ideas are pushing the boundaries in the commercial sector of supply chain management, they may seem far too advanced for the humanitarian sector, but I believe there are indeed applications to the topic at hand.

The part of Lambert's idea that is applicable is the underlying theme that the proper use of supply chain information can create competitive advantage. In the humanitarian sector, due to the donors demanding more and more proof of operational effectiveness and due to the media which stands by attentively watching, there will be smaller allowances for inefficiencies in the future. There will, however, be greater chances for organizations to differentiate themselves through the use of streamlined supply chains and well-measured operations.

4 A Performance Indicator Framework for the Humanitarian Supply Chain

4.1 Explanation and Description of Proposed Framework

4.1.1 Understanding the Business Processes at the IFRC

The framework and the process of how it was derived are described in this chapter in terms specific to the IFRC. However, any organization implementing a measurement system should follow first determine what elements of performance are most important them and then define metrics that support these priorities. This should provide a basic template that is applicable to other humanitarian organizations, given the similar over-arching goals of delivering goods in the disaster relief environment.

One theme of the research thus far has been that organizational goals should determine what metrics are put into place. In order to apply this methodology to the IFRC, it was first necessary to determine from a broad level what is important to the organization and how they define a successful operation. To determine their goals, I interviewed individuals from several departments at the IFRC headquarters in Geneva, Switzerland, during the week of January 24-27, 2006. The purpose of these interviews was to obtain different perspectives of how the IFRC

defines a successful emergency operation and how best to measure the performance of their logistics function given these goals. Individuals were asked questions such as the following:

- From your perspective, what is the most important aspect that defines a successful emergency operation?
- What should be done to streamline operations and the current process of obtaining items for distribution to beneficiaries?
- Where do you see the "hold-ups" in the system? Where does the flow of goods from procurement to distribution tend to slow down or break down?

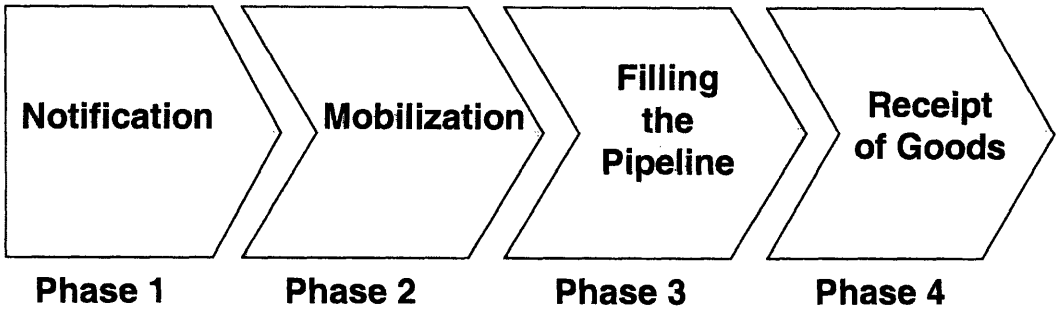
The positions and departments represented by the individuals interviewed at IFRC headquarters in Geneva are presented in Table 1 below.

Table 1 – Individuals Interviewed at IFRC Headquarters

IFRC Department	Position / Title
Disaster Preparedness & Response	Senior Officer, Disaster Policy
Finance	Department Head
Information Systems	Department Head
Logistics & Resource Mobilization	Department Head
Logistics & Resource Mobilization	Senior Officer, Procurement
Logistics & Resource Mobilization	Senior Officer, Mobilization
Logistics & Resource Mobilization	Senior Officer, In-Kind Donations
Logistics & Resource Mobilization	Senior Officer, Systems & Processes
Monitoring & Evaluation	Manager
Operations Support	Coordinator, Americas Region
Operations Support	Senior Officer, Reports & Appeals
Support Services	Director
Water & Sanitation	Manager

Based on the interviews conducted with these individuals an emergency operation at the IFRC can be grouped into four phases, which are shown in Figure 2.

Figure 2 – Four Phases of an IFRC Emergency Operation



Phase 1: Notification

The IFRC is first notified of an event by the national society of the country or countries in which the disaster took place. If the items required for aid are greater than the capacity of what the national society can provide and they request aid, then the IFRC establishes involvement in mobilizing goods. Otherwise, the operation will be managed at the national level without any visibility to the IFRC. In certain large-scale disasters in which it is clear that the requirements will exceed local capacity, the IFRC may skip the "notification phase" and directly enter the "mobilization phase." In most cases, however, the mobilization phase does not begin until the IFRC logistics department receives the initial list of requested items and item quantities from Operations Support, the department that supervises assessment and distribution activities in the field.

Phase 2: Mobilization

The "Appeal Creation Date" is the date on which the IFRC commits to responding to an emergency. On this date, the logistics department creates a new project in the HLS system. Items requested from the field are entered into a "Mobilization Table," which is the tool in HLS used to track the requested items, item quantities, and expected unit costs. The Mobilization Table can only be updated by the IFRC headquarters. After it is initially populated at IFRC headquarters, it is published online to the national societies of the Red Cross and Red Crescent. These national

societies serve as the direct donors to the IFRC for "in-kind donations" (donations of items instead of cash), which are requested on the Mobilization Table. If a national society wishes to donate items listed on the table, they contact the IFRC via telephone or e-mail to record the donation. The IFRC logistics department updates the Mobilization Table as donors make pledges for items, as well as when new information arrives from the field about the items and quantities needed on-site.

The IFRC receives in-kind donations as well as cash donations. When cash donations are given toward an operation, the IFRC has a formalized process of procuring these items from suppliers and shipping them directly to the destination country. The framework proposed in this chapter, however, is based more on in-kind donations. In the two IFRC operations examined, 100% of the Sahel Food Security Crisis shipments and 80% of the South Asia Earthquake Operation shipments were in-kind donations. Because of the relatively small proportion of procured items, no additional metrics were incorporated into the IFRC framework to account for the time it takes to raise money before procurement can occur. Unlike the IFRC, however, some non-profit humanitarian organizations only receive cash donations instead of in-kind donations. Section 6.1 describes how the proposed framework can be modified to be more applicable to these other organizations.

Phase 3: Filling the Pipeline

After donors register their donation with the IFRC, they notify the IFRC again when the shipping information is available. A shipment of goods is referred to as a "consignment" in HLS. The donor provides information to the IFRC for each consignment, such as item codes, item quantities, the expected shipment date, the expected arrival date, and the mode of transportation. The above information is then entered into HLS by the IFRC logistics department.

As shipments are recorded, the logistics department creates the "Pipeline Report" using HLS and distributes it periodically to logistics personnel as well as the Operations Support personnel working in the field. The purpose of this report is to provide visibility in the supply chain so that the field personnel can see when they can expect to receive future shipments. How often the Pipeline Report is distributed depends on the scope of the operation. In large-scale disasters where the pipeline is changing rapidly due to several shipments arriving via airfreight during the first weeks after a disaster, this report may be distributed as often as 2-3 times per day.

Phase 4: Receipt of Goods

When a consignment is received in the destination country, the logistics personnel on-site create a "Goods Received Note" or "GRN" for it and then notify the IFRC logistics department of its arrival. The GRN contains the date the goods arrived in-country, a list of the item codes in the consignment, and item quantities received. The logistics department then records the GRN data in HLS which modifies the consignment to appear as "closed." A copy of the GRN is also sent back to the donor to show them when their goods arrived on-site.

The arrival of the goods in the destination country and the creation of the GRN is as far downstream in the relief supply chain as HLS currently captures. HLS does not currently cover the "last mile of delivery," which is from the arrival of the goods in-country to the actual distribution to beneficiaries within that country. Hence, one limitation of the proposed framework of metrics is that there may be delays in this last mile that it will not observe, but the delays will occur nonetheless in preventing items from reaching their final destination.

It is also worthy to note that there may be a delay from the date a consignment arrives in-country and the date a Goods Received Note is recorded. This may be due to delays in customs or there may be too many arrivals at a given time for the field personnel to process all deliveries

as they are arriving. In either of these possibilities, while accuracy of the true arrival date is important, the GRN date will still be useful to the metrics regardless, because it notifies the IFRC that the goods shipped have, in fact, been delivered.

4.1.2 Proposed Framework and Definition of Key Indicators

Some articles related to humanitarian logistics suggested that financial efficiency and responding to the requests of donors were equal in importance as the effectiveness of the organization's actual response to beneficiaries. Financial efficiency is clearly an important goal at the IFRC, but the over-arching finding based on my interviews is that a timely response to the needs expressed in the field is much more important to the organization than financial efficiency alone. Hence, this must be kept in mind when designing and implementing a system of performance measurement.

Although a Balanced Scorecard approach was a candidate upon which to base a framework of metrics, it was not the final solution chosen. This is because what was determined to be important to the IFRC was not in complete alignment with the four quadrants of the Balanced Scorecard. Also, in an interview with Ray Archer, Vice President of DAO Operations at Dell, Inc., he noted that the Balanced Scorecard is often best used to track how well an organization is meeting its goals for *change management*, but it is not necessarily the best method to simply measure performance. Mr. Archer indicated that for an organization that is not familiar with performance measurement, it is best to start by identifying what is important to that specific organization and designing metrics that address these goals and leverage the organization's core competencies. Because the goal of this research is more closely related to performance measurement of a department rather than change management at an organizational level, the idea of forcing a framework for humanitarian logistics strictly into a Balanced

Scorecard was therefore abandoned. It is certainly worthy to note, however, that it may be modified for application in humanitarian organizations, depending on the business processes the organization is seeking to change.

A template of the proposed framework to be implemented at the IFRC is shown below in Figure 3, which is followed by a detailed explanation of the four performance indicators and how they would be used to measure performance throughout an operation.

Figure 3 – Proposed Framework for IFRC

Operation Name Appeal Date				
Status Update: Final ***Current Date:***	**Operation Total (Weighted)**	**Priority 1 Item Group**	**Priority 2 Item Group**	**Total Op Target**
Percent of Appeal Coverage (in quantity of items)				
After 1 week	%	%	%	%
After 2 weeks	%	%	%	%
After 1 month	%	%	%	%
After 2 months	%	%	%	%
After 3 months	%	%	%	%
Percent of Items Delivered (in quantity of items)				
After 1 week	%	%	%	%
After 2 weeks	%	%	%	%
After 1 month	%	%	%	%
After 2 months	%	%	%	%
After 3 months	%	%	%	%
Donation-to-Delivery Time				
Mean (# days)	# days	# days	# days	# days
Median (# days)	# days	# days	# days	# days
Financial Efficiency				
(Donor Cost - Budget Cost) / Budget Cost	%	%	%	%
Actual $ Spent - Budget $	$	$	$	$
Transportation Cost / Total Product Cost	%	%	%	%
Assessment Accuracy: Revised Budget / Original Budget				
After 2 weeks	%	%	%	%
After 1 month	%	%	%	%
After 2 months	%	%	%	%
After 3 months	%	%	%	%

The proposed framework relies upon the following four performance indicators, which are each described in more detail below:

1. Appeal Coverage
2. Donation-to-Delivery Time
3. Financial Efficiency
4. Assessment Accuracy

Appeal Coverage: This indicator uses two specific metrics: 1) *percent of appeal coverage* and 2) *percent of items delivered.*

The first metric—*percent of appeal coverage*—is defined as the percentage of the quantity of items that donors have pledged out of the total number of items requested for the operation at a given point in time. It is expressed in fractional form as follows:

$$\frac{\textit{Quantity of Items Pledged}}{\textit{Total Items Requested}}$$

The purpose of this metric is to indicate to the organization how well and how quickly they are finding pledges for the requested items.

The second metric—*percent of items delivered*—is defined as the percentage of items that have actually been delivered on-site out of the total number of items requested for the operation at a given point in time. This metric is expressed in fractional form as follows:

$$\frac{\textit{Quantity of Items Delivered}}{\textit{Total Items Requested}}$$

These two metrics together indicate how well an organization is both finding donors and delivering goods to the destination country at a specific point in time.

Donation-to-Delivery Time: The second performance indicator is straightforward, measuring how long it takes for an item to be delivered after a donor has pledged to donate it. Based on the findings from the military logistics research, it is important to measure not only the mean delivery time but also the median delivery time, in an attempt to encourage consistency of delivery times as well as to gauge how long it takes on average for items to be delivered. As described further in Section 7.1, this metric could later be modified to become a more "relative" metric, such as "On-Time Delivery," which would allow for more comparability across operations. During this initial implementation, however, it is best to begin with an absolute metric so that the actual lead time in the number of days is visible to the logistics department.

Financial Efficiency: Three metrics are included as part of the indicator of financial efficiency. As mentioned earlier, how important this indicator is depends on the organization, but since all non-profit humanitarian organizations are held accountable to spending their donors' money wisely, this is a universal indicator which should be included in the framework. The first financial efficiency metric is defined as:

$$\frac{(Donor\ Cost - Budget\ Cost)}{Budget\ Cost}$$

This metric expresses the amount under or over budget as a percentage of budget cost. We suggest that at the beginning of an operation, an organization should establish a standard cost of what they expect to pay for a given item, based on current market values or based on the prices they have previously negotiated with suppliers. Although some organizations have created standard item catalogs that contain prices, these catalogs are often out of date and often do not reflect current market values.

The second metric of financial efficiency expresses differences to budget in more absolute terms. The second financial efficiency metric is defined as:

Actual Dollars Spent – Budgeted Dollars[1]

This metric should be used together with the first financial efficiency metric to present a fuller picture of the operation's financial state. For example, if the value of the relative metric at a point in time is 300%, this would indicate on its own that the operation is far exceeding its budget. If, however, the value of the absolute metric is $200,000, and one knows that the entire budget for the operation is $5 million, then one can conclude that despite the "relative" measure of overage, but the dollar-value of overage is not significant. Hence, these metrics used together help to show how financially efficient the operation is at a given point in time.

The third metric of financial efficiency incorporates the transportation cost of delivering the goods to the beneficiaries. Because certain items are considered to be critical needs immediately after a disaster strikes, it is expected that transportation costs will typically be very high in the first days of an operation. Goods are often shipped via airfreight immediately a disaster to get them to beneficiaries as quickly as possible. One can expect, however, that transportation costs should decline over the life cycle of the operation (e.g. one would not expect to see items shipped via airfreight late in the operation, because they could be shipped using a less expensive method after the initial phase of the relief effort). One would also expect the ratio of total transportation costs to total product costs to decrease over the course of the operation as well, because transportation costs should be decreasing as discussed above, and total product

[1] Note that the second financial efficiency metric is expressed in the generic term of "dollars", which has been replaced in the IFRC scorecards shown as Swiss Francs (CHF) because that is the currency used by the IFRC. Note that for other organizations using other currencies, any currency may be used in the proposed framework of metrics.

costs should be increasing as more and more items are delivered on-site. Hence, as a way to gauge financial efficiency, the third metric is expressed at a given point in time as:

$$\frac{\textit{Total Transportation Costs}}{\textit{Total Product Costs}}$$

Assessment Accuracy: Since the logistics function of an organization mobilizes its goods and donations based on the assessment of needs from personnel in the field, the amount of goods the logistics department can get into the pipeline relies on this information being accurate. For example, in the 2005 South Asia earthquake, the initial estimates of the death toll were far lower than the final estimation even up to several weeks after the earthquake occurred. In this case, the logistics department of any humanitarian organization would have been unable to respond with the appropriate quantity of goods with the first wave of items they delivered because the assessment of the earthquake's damage was incorrect.

The purpose of a metric related to assessment accuracy is to therefore measure how much the operation's final budget changed over time from the operation's original budget. This metric is expressed in the form of a percentage of:

$$\frac{\textit{Revised Operation Budget}}{\textit{Original Operation Budget}}$$

The original operation budget is defined as the budget one week after the first day of disaster operation (Original Appeal Date). The operation budget equals the budgeted cost of each item multiplied by the total number of items requested. In the breakpoints of time when the operation is subsequently evaluated (i.e. two weeks, one month, two months, etc.), the revised operation budget captures the items that were added to the list after the first week plus the original items requested. The percentage over or under 100% indicates how accurate the original assessment of needs was one week after a disaster occurred. This metric will help put the values

of the other metrics in the framework into context. For example, if it appears on the scorecard that the delivery lead time of a specific type of item was longer than the operation's average delivery lead time, the assessment accuracy metric will indicate if that late delivery was caused by a low initial estimation of the quantity needed of that type of item.

To understand how the data captured in HLS affects the values of the four indicators, Table 2 summarizes what elements of information are used to calculate each of the metrics.

Table 2 – Data Elements Used to Calculate the Four Performance Indicators

	Data Elements Used to Calculate the Four Performance Indicators								
Key Indicator	*Original Appeal Quantities*	*Revised Appeal Quantities*	*Pledged Quantities*	*Quantities of Items Received*	*Budgeted Item Costs*	*Actual Donor Item Costs*	*Transportation Costs*	*Pledge Date*	*Goods Received Note Date*
Percent of Appeal Coverage	X	X	X						
Percent of Items Delivered	X	X		X					
Speed of Delivery								X	X
Financial Efficiency - Budget vs. Actual	X	X	X		X	X			
Financial Efficiency - Transportation Costs				X		X	X		
Assessment Accuracy	X	X			X				

4.1.3 Features of the Framework

There are three additional features of this framework: a) the designation of high-priority items to be delivered during an operation, b) the use of specific breakpoints in time that are recommended to implement this framework, and c) the importance of how to weigh various types of items in calculating the operation's overall performance.

Item Group Priorities

At the outset of every operation, organizations should begin by designating what specific types of items will be their "priority items" to deliver to beneficiaries. The designation of certain

types of items as key priorities will help an organization define and clarify what the successful outcome of their operation will look like. Intuitively, without explicitly-stated priorities it is difficult for an organization to review their performance at the end of an operation and truly say that they were able to achieve their goals. The setting of priority-items helps them to determine which items will have the most critical impact on the beneficiaries, and it will help the logistics department internally to focus on procuring and delivering these items as rapidly as possible. In short, priorities will help an organization maintain strategic focus throughout an operation. For example, if the IFRC decides to focus their efforts on mobilizing water and sanitation items for a particular operation, not only will this save time from the donors asking what items are of the highest priority, but they will be able to continually focus their efforts in mobilizing and delivering these goods. They will also be better able to track and report their performance when the operation has been completed. This re-enforces one of the themes that has emerged throughout the course of this research, which is that spending more time in the planning phase of the supply chain causes a more consistent and predictable flow of goods throughout the supply chain, all the way to the distribution of the items to beneficiaries.

From interviews conducted with individuals in the logistics department at the IFRC, HLS does not currently have the capability to prioritize items. When a disaster strikes, individuals from various national societies contact the IFRC logistics headquarters via telephone or e-mail to ask the question, "Out of the list of items currently on the appeals list, which is the most important?" Oftentimes, a national society may have a donor step forward with a monetary donation with the intention to satisfy the most pressing needs first. The problem is that in the "Mobilization Table," which lists the current appeals for an operation, the items needed are not listed with any kind of priority or ranking system. Hence, such priorities are determined

manually through communication between Operations Support in the field, the logistics department at the IFRC headquarters, and then to the national society donors. This current process of communication is often very time-consuming for the individuals in the logistics department during the critical hours and days after a disaster. Thus, altering HLS to capture item-type priorities for each operation would not only help the organization measure their performance, it would also save time and effort of the logistics personnel.

Ideally in the future, the process of designating priority items could be further streamlined if several different non-profit organizations were to coordinate at the outset of a disaster and decide who would specialize in the different type of items needed (e.g. housing and shelter, water and sanitation, kitchen sets, etc.). This would help eliminate some of the redundancies that commonly occur in some operations, such as when multiple organizations deliver too many of one type of item and too few of another. Still, the most pressing need is clearly for priorities to be determined at the organizational level.

Breakpoints in Time

The second feature of the framework is how to use these scorecards at specific breakpoints in time. The purpose of setting breakpoints in time after the original appeal date is to understand how the operation's performance changes over time. This will enable an organization to evaluate performance *during* an operation, so that corrective action can be taken as necessary, depending on how well the targets on the scorecard are being met.

The necessity of using breakpoints in time after the original appeal date is driven by two of the key differences which exist between humanitarian operations and commercial operations. One key difference is that the assessment of the needs of the affected population may fluctuate significantly in the days and weeks following a disaster. This fluctuation affects the

organization's ability to respond accurately and ensure that the appropriate items and quantities are placed into the delivery pipeline. The second key difference is that the organization's targets of how many people they will assist in an operation may also change over time. For example, after a disaster occurs, an organization may initially decide to set a target goal of providing aid to 50,000 people. If the disaster's damage is later determined to be significantly worse than originally estimated, then the organization may change its target population to 80,000 people. One can see that either of these occurrences—the assessment of the damage changing or the target population changing—will affect how quickly the appropriate items and quantities arrive at their intended destination of the beneficiaries.

Given that the above situations are inherent in working in the sector of humanitarian aid and disaster relief, an organization should designate specific breakpoints in time at which they examine both the state of the operation and the performance of the logistics function. In the IFRC analysis, the first scorecards were calculated one week after the original appeal date. One week is a sufficient amount of time for field officers to obtain information from the local governments related to the extent of the disaster's damage, revise as new information is obtained, and communicate the most recent information to the logistics department in order to mobilize the goods that are requested.

Furthermore, I would caution placing a breakpoint in time of less than one week, because during this week the logisticians are concentrating their time and efforts to mobilize goods into the pipeline, and the information surrounding the event may still be unclear. Although the scorecards would be automatically generated by the IT system, the time spent by the logisticians evaluating the scorecards within the first week may actually detract from the performance itself, given the limited resources many organizations already face. This supports the idea of

"economy" to be used as a criterion when determining how to structure a framework of metrics, as described earlier in the 1994 Caplice & Sheffi article. This criterion states that "the benefits of using the metric outweigh the costs of data collection, analysis, and reporting." The benefits of creating and using a scorecard during the first week of an operation would not outweigh the costs of taking the time of the logisticians during this critical phase of an operation.

It is up to the individual organization how often after the initial week they should place their breakpoints in time to analyze their performance. For the IFRC, breakpoints were set arbitrarily at two weeks, one month, two months, and three months after the initial appeal date.

How Items are Weighted Affects Performance Metrics for Overall Operation

The purpose of the "Operation Total (Weighted)" column in the proposed framework is to help the user interpret how quickly the pipeline is being filled and items are being delivered for the entire operation. The other two columns focus specifically on the types of items that have been designated as high priorities. An important feature to recognize about the framework is that how *each* type of item has been weighted in the operation in total will therefore affect the value of the metrics in this column.

Ideally, at the beginning of an operation when an organization is defining their goals and targets, they would explicitly state how much of the total budget should be allotted for each type of item. For example, as shown below in Table 3 an organization could define their goals for Operation X as the following:

Table 3 – Example of Operational Budget Goals

Operation X Budget Goals	
Item Type	**Budget Goal**
Food	10%
Housing	25%
Transportation	15%
Water / Sanitation	50%
Operation Total	**100%**

If the goal is that half of the budget should go towards water and sanitation items, then in calculating the values for the metrics in the "Operation Total (Weighted)" column, the water and sanitation items would be weighed as worth 50% of the total.

To calculate the values shown on the scorecards in Section 4.2, the weights of each type of item were determined as the percentage of each item-type out of the total budget at each breakpoint in time. Since the budgets for the two operations were not initially set as "fixed" but rather they changed over time, for these calculations the weights had to change over time as well. Ideally, the weights should not change over time, but should be determined at the beginning of an operation. This will ensure consistency of the values presented on the various scorecards.

When determining how items should be weighed, an organization may want to adjust the weights to emphasize early delivery of certain lower-valued items. An example of this would be if an organization placed a high priority on early delivery of water / sanitation items to prevent the spread of disease after a disaster. Water / sanitation items may have a lower dollar value than shelter items, for example, but they could be weighted higher in the table of the operation's budget goals so they would arrive as early as possible on-site.

Finally, weights are important because they affect the organization's ability to compare results across operations. If item-types are designated to have different weights in various operations, then the results would not be truly comparable if one was to examine the Operation Total columns of two scorecards. This is an example of the trade-off between validity and robustness which was described in the 1994 Caplice & Sheffi article in Section 3.2.1. Validity is how specific a metric is to a particular issue. Robustness is how comparable the metric is across various issues (Caplice & Sheffi, 1994). In this case, as weights are more specific to an organization's goals for a particular operation, the less comparable the metrics will be across

multiple operations. To solve this dilemma, ideally an organization would determine a standard set of weights that represent their goals for item-types in *all* operations. The "Operation Total (Weighted)" column could then be accurately compared across multiple operations. To gain details about the performance of one operation, the organization could rely on the remaining two columns of the scorecard, which contain the item priorities which have been designated for that particular operation.

4.2 Walk-through of Two IFRC Operations Using Proposed Scorecards

4.2.1 Methodology

Two emergency operations were selected for the analysis of the HLS data. The first operation was the Sahel Food Security Crisis (referred to herein as "Sahel"), which had an appeal date of July 22, 2005. The second operation selected was the South Asia Earthquake operation (referred to as "South Asia"), which had an appeal date of October 9, 2005. Both operations occurred over 18 months after HLS was fully implemented, which indicates that the logistics department users were comfortable by then with both the system itself and with the business processes which the system helped to automate. The operations selected were significantly different in both the scope and disaster type, which provides a representative sample of the varying types of operations and disasters that the IFRC faces on a daily basis. Data was obtained from queries of the HLS system which were run by Martin Bush, Senior Officer of Systems and Processes of the Logistics and Resource Mobilization Department. The queries were imported directly into Microsoft Excel for analysis. In both operations, no priorities of item-types had officially been designated in the HLS system. "Priority 1" items were determined for this analysis based on interviews of individuals at the IFRC where they indicated what the most

important type of item was that the IFRC needed to deliver in this operation. "Priority 2" items were determined by analyzing the appeals list and selecting the type of items that was the second largest percentage of the budget, since the Priority 1 items took up the largest amount of the budget.

Because the two operations occurred in different geographic regions of the world, the level of difficulty for goods to be delivered to each country is very different. Sahel occurred mostly in a land-locked region of Africa. Of the four affected countries—Niger, Mauritania, Burkina Faso, and Mali—only Mauritania had direct seaport access. South Asia did have seaport access, but because of the mountain landscape, the remote villages, and the unexpectedness of the disaster, airfreight proved to be the most-used method of transport in this operation. Besides the location and disaster-type differences of Sahel and South Asia, one must also note that the other major difference between these operations is that South Asia attracted much more media attention on a global scale than did Sahel, which clearly affected the speed of donor response. Key factors relating to each case are summarized in Table 4.

Table 4 – Comparison of IFRC Operations Used as Case Studies

	Sahel	South Asia
Site of affected country	Third world region	Third world region
Seaport access?	Mostly in land-locked region	Yes
Sudden disaster?	No, slow onset disaster	Yes
Media exposure	Limited	High, global media exposure
Clear priority of IFRC item groups?	Yes	Yes
Duration of operation	5 months	6 months

4.2.2 Walkthrough of Proposed Framework

The following section will re-create a walk-through of what the proposed scorecards would have looked like for the Sahel and South Asia operations had this system been in place at that time. For each breakpoint in time (e.g. one week after appeal date, two weeks after appeal date, etc.), the data that was available at that time in HLS has been used to create the following scorecards, so the information presented is what the organization would have been able to see if they had been using this system. After each set of scorecards is presented, their results will be compared and contrasted. Following the walk-through, Chapter 5 will present additional analysis performed on these two operations to understand why the operations unfolded as they did.

The "Total Op Target" column has been left blank in the following illustrations. A key feature of using this system is that the organization would be forced to set targets for each metric at the outset of an operation in order to gauge their performance throughout the operation. Because this system was not in use during these two operations and goals were not set by the IFRC in this format, the column has been left blank.

Figure 4 – Sahel & South Asia : Week 1 Scorecards

Sahel Food Security Crisis Appeal Date: July 22, 2005				
Status Update: Week 1 ***Date: July 29, 2005***	**Operation Total (Weighted)**	**Priority 1 Food**	**Priority 2 Transport**	***Total Op Target***
Percent of Appeal Coverage (in quantity of items)				
After 1 week	5%	1%	6%	
Percent of Items Delivered (in quantity of items)				
After 1 week	0%	0%	0%	
Donation-to-Delivery Time				
Mean (# days)	N/A	N/A	N/A	
Median (# days)	N/A	N/A	N/A	

South Asia Earthquake Appeal Date: October 9, 2005				
Status Update: Week 1 ***Date: October 16, 2005***	**Operation Total (Weighted)**	**Priority 1 Housing**	**Priority 2 Kits & Sets**	***Total Op Target***
Percent of Appeal Coverage (in quantity of items)				
After 1 week	63%	61%	77%	
Percent of Items Delivered (in quantity of items)				
After 1 week	6%	1%	4%	
Donation-to-Delivery Time				
Mean (# days)	3	3	3	
Median (# days)	2	3	2	

One week after the organization had committed itself to each operation, 1% of the requested food items in Sahel had found donors as compared to 61% of the requested housing items in South Asia. In Sahel, no deliveries had been made after one week. Six percent of the total requested items had been delivered in South Asia.

There are no "Financial Efficiency" metrics or any "Assessment Accuracy" metrics included in the Week 1 Scorecard. Based on interviews with the IFRC, the immediate priority of the logistics department after an appeal is launched is to cover the appeal by finding donors and to monitor how quickly the donated items are delivered. In other words, since financial efficiency is not a priority at this point in time, it has not been included on these scorecards. Assessment accuracy is not included in the Week 1 Scorecard for a different reason; this metric

uses the total value of items requested after one week as the baseline value for future scorecards. Hence, assessment accuracy is not applicable after one week.

Figure 5 – Sahel & South Asia: Week 2 Scorecards

Sahel Food Security Crisis Appeal Date: July 22, 2005				
Status Update: Week 2 ***Date: August 5, 2005***	**Operation Total (Weighted)**	**Priority 1 Food**	**Priority 2 Transport**	***Total Op Target***
Percent of Appeal Coverage (in quantity of items)				
After 1 week	5%	1%	6%	
After 2 weeks	47%	54%	8%	
Percent of Items Delivered (in quantity of items)				
After 1 week	0%	0%	0%	
After 2 weeks	3%	0.1%	0%	
Donation-to-Delivery Time				
Mean (# days)	8	8	N/A	
Median (# days)	8	8	N/A	
Assessment Accuracy: Revised Budget / Original Budget				
After 2 weeks	103%	100%	100%	

South Asia Earthquake Appeal Date: October 9, 2005				
Status Update: Week 2 ***Date: October 23, 2005***	**Operation Total (Weighted)**	**Priority 1 Housing**	**Priority 2 Kits & Sets**	***Total Op Target***
Percent of Appeal Coverage (in quantity of items)				
After 1 week	63%	61%	77%	
After 2 weeks	47%	45%	18%	
Percent of Items Delivered (in quantity of items)				
After 1 week	6%	1%	4%	
After 2 weeks	9%	5%	2%	
Donation-to-Delivery Time				
Mean (# days)	5	6	6	
Median (# days)	5	6	4	
Assessment Accuracy: Revised Budget / Original Budget				
After 2 weeks	130%	118%	365%	

The information given on the Week 2 Scorecards must be carefully interpreted in order to draw the appropriate conclusions about each operation. The Week 2 Scorecard uses cumulative information from the Week 1 Scorecard, to give the user a perspective of the operation over time. At first glance, it appears that in the Sahel operation, requested items found more donors during the second week, while the donations decreased in the South Asia operation during Week 2. The

former statement is true: the requested items *did* find more donors in Week 2 of Sahel. 47% of the total requested items had found donors after Week 2, which was an increase from 5% after Week 1. From the Assessment Accuracy metric, one can also determine that the list of items requested (or the "Appeals List") did not change very significantly; hence, the original assessment after 1 week was relatively accurate and increased by only 3% from Week 1 to Week 2. The latter statement above, that donations somehow decreased during Week 2 of the South Asia operation, is *not* true. After Week 2, 47% of the total goods requested had found donors, which was a decrease from 63% after Week 1. By noting the Assessment Accuracy metric for this operation, however, which is at 130%, one can see that the Appeals List increased 30% during that week, which would explain why it appears that donations decreased.

Regarding deliveries that were made, a small number of deliveries were made during the second week of the Sahel operation, or 3% of the total number of goods requested. The number of deliveries made in the South Asia operation appears to have increased only to 9%, but one can deduce a significant number of deliveries *were* made that week. Again, this is because the total number of items requested increased by 30% during Week 2, which would be the denominator of the calculation of percent of items delivered (number of items delivered / total number of items requested).

Like the Week 1 Scorecards, the Week 2 Scorecards do not contain financial efficiency metrics, since it is our recommendation based on the IFRC's organizational goals that the coverage and the length of delivery metrics be emphasized when analyzing performance more than the financial metrics at this point in time. How highly the organization places an emphasis on financial efficiency will determine when they want to begin incorporating these metrics into

their analysis. Based on the research performed at the IFRC, we would recommend adding the financial efficiency metrics beginning in the Month 1 Scorecard, shown in Figure 6.

Figure 6– Sahel & South Asia: Month 1 Scorecards

Sahel Food Security Crisis Appeal Date: July 22, 2005				
Status Update: Month 1 ***Date: August 21, 2005***	**Operation Total (Weighted)**	**Priority 1 Food**	**Priority 2 Transport**	***Total Op Target***
Percent of Appeal Coverage (in quantity of items)				
After 1 week	5%	1%	6%	
After 2 weeks	47%	54%	8%	
After 1 month	96%	100%	67%	
Percent of Items Delivered (in quantity of items)				
After 1 week	0%	0%	0%	
After 2 weeks	3%	0.1%	0%	
After 1 month	7%	3%	4%	
Donation-to-Delivery Time				
Mean (# days)	11	12	10	
Median (# days)	10	10	10	
Financial Efficiency				
(Donor Cost - Budget Cost) / Budget Cost	-100%	-5%	265%	
Actual CHF Spent - Budget CHF	6,187,250	3,084,683	3,166,949	
Assessment Accuracy: Revised Budget / Original Budget				
After 2 weeks	103%	100%	100%	
After 1 month	124%	122%	100%	

South Asia Earthquake Appeal Date: October 9, 2005				
Status Update: Month 1 ***Date: November 8, 2005***	**Operation Total (Weighted)**	**Priority 1 Housing**	**Priority 2 Kits & Sets**	***Total Op Target***
Percent of Appeal Coverage (in quantity of items)				
After 1 week	63%	61%	77%	
After 2 weeks	47%	45%	18%	
After 1 month	74%	73%	51%	
Percent of Items Delivered (in quantity of items)				
After 1 week	6%	1%	4%	
After 2 weeks	9%	5%	2%	
After 1 month	33%	27%	8%	
Donation-to-Delivery Time				
Mean (# days)	11	12	12	
Median (# days)	11	11	12	
Financial Efficiency				
(Donor Cost - Budget Cost) / Budget Cost	-7%	-12%	44%	
Actual CHF Spent - Budget CHF	(3,570,139)	(5,531,198)	1,992,575	
Assessment Accuracy: Revised Budget / Original Budget				
After 2 weeks	131%	118%	365%	
After 1 month	139%	123%	377%	

Appeal Coverage: Between the two operations for the highest priority item-types, after one month donors had pledged 100% of the food items in the Sahel operation. In contrast, 73% of the housing items had been donated in the South Asia operation, but more of these items were reaching the destination country after one month. Only 3% of the food items had been delivered in the Sahel operation after one month, as opposed to 27% of the housing items that were delivered to South Asia.

Donation-to-Delivery Time: Both operations show comparable delivery lead times averaging at 11 days during the first month of the operation. Note that the "Total Op Target" column will be particularly useful for this metric. The organization will in essence be able to determine what method of transportation should be used at a given phase of an operation by how they set this target. For example, if the target is set at a mean of 3 days, then the organization is forcing themselves to ship the goods via airfreight in order to meet that goal. If after one month, the target is set at 15 or 20 days, the organization would be relaxing this rule and encouraging shipments to be made via other methods that have longer lead times, such as ocean-freight or truck.

Financial Efficiency: As described earlier, the first two financial efficiency metrics need to be interpreted together in order to understand the true financial dynamics of the operation. First, for the *(Donor Cost – Budget Cost) / Budget Cost* metric, positive percentages indicate that the donor cost exceeded the budgeted cost. Negative percentages indicate that the budgeted cost exceeded the donor cost. To meet the budget exactly, this metric would have a value of 0%.

In the Sahel operation, these values are seen to fluctuate much greater around 0% than the values in the South Asia operation. In the Operation Total column, the -100% indicates that weighted for the operation in total, the budgeted cost far exceeded the donor's cost for most items. The 265% in the Transport column, however, indicates that the donor's cost for transport items far exceeded the budgeted cost.

The second financial efficiency metric measured in CHF helps to put the first one into context. In comparing the two operations, one can interpret these values to mean that for the Sahel operation in total actual spend exceeded budgeted spend by 6.1 million CHF. Conversely, for the South Asia operation in total budgeted spend exceeded actual spend by 3.5 million CHF. It is interesting to note, however, that in the South Asia operation in the Kits & Sets column, actual spend exceeded budgeted spend by 1.9 million CHF. However, when one examines the Assessment Accuracy metric for Kits & Sets, it becomes apparent that the original needs assessment for this type of item was far lower than what had been requested after one month, which is why the actual spend exceeded the budget.

The third financial efficiency metric, *Total Transportation Cost / Total Product Cost*, is not included in the Month 1 scorecards in Figure 6, because transportation costs are not currently captured by HLS *during* the operation. Transportation costs currently are requested from donors after an operation has been completed, and they are then compiled for the total operation at IFRC headquarters. Because one would want to see this ratio decrease between the Month 1 scorecard and the Final Scorecard for an operation, we do recommend that this data should be captured at the consignment level and included in the future on the Month 1 scorecard. This idea is discussed further in Chapter 7 as one of the next steps in moving forward with implementing this proposed framework.

Assessment Accuracy: After Month 1, both of the operations experienced significant increases from their original Week 1 assessment of needs. The type of item that was most underestimated among the two operations was the Kits & Sets in South Asia, as mentioned above, with a 377% increase after one month from the Week 1 assessment. As described earlier this metric helps to put the other values on the scorecard into context for the logistics department, so that they can see the delays in the pipeline, as well as gauge what information they were lacking at the outset of an operation.

Figure 7 shows the Final Scorecards for the two operations, which give a picture of the operation's performance three months from the original appeal date. Denoting these Month 3 scorecards as "Final" is more for the purposes of illustration and clarity in this discussion. In actuality, these operations were not fully completed after three months. Items arrived in Sahel up to five months after the original appeal date, and shipments were still arriving in South Asia until March, 2006, which was when this data analysis was performed.

Figure 7 - Sahel & South Asia: Final Scorecards

Sahel Food Security Crisis Appeal Date: July 22, 2005				
Status Update: Final ***Date: March 18, 2006***	**Operation Total (Weighted)**	**Priority 1 Food**	**Priority 2 Transport**	***Total Op Target***
Percent of Appeal Coverage (in quantity of items)				
After 1 week	**5%**	**1%**	**6%**	
After 2 weeks	**47%**	**54%**	**8%**	
After 1 month	**96%**	**100%**	**67%**	
After 2 months	**100%**	**100%**	**100%**	
After 3 months	**100%**	**100%**	**100%**	
Percent of Items Delivered (in quantity of items)				
After 1 week	**0%**	**0%**	**0%**	
After 2 weeks	**3%**	**0.1%**	**0%**	
After 1 month	**7%**	**3%**	**4%**	
After 2 months	**38%**	**3%**	**98%**	
After 3 months	**52%**	**25%**	**98%**	
Donation-to-Delivery Time				
Mean (# days)	**29**	**36**	**20**	
Median (# days)	**15**	**18**	**9**	
Financial Efficiency				
(Donor Cost - Budget Cost) / Budget Cost	**131%**	**81%**	**482%**	
Actual CHF Spent - Budget CHF	**6,204,512**	**3,084,683**	**3,119,011**	
Transportation Cost / Total Product Cost	**37%**	**N/A**	**N/A**	
Assessment Accuracy: Revised Budget / Original Budget				
After 2 weeks	**103%**	**100%**	**100%**	
After 1 month	**124%**	**122%**	**100%**	
After 2 months	**123%**	**122%**	**107%**	
After 3 months	**123%**	**122%**	**107%**	

South Asia Earthquake Appeal Date: October 9, 2005				
Status Update: Final ***Date: March 18, 2006***	**Operation Total (Weighted)**	**Priority 1 Housing**	**Priority 2 Kits & Sets**	***Total Op Target***
Percent of Appeal Coverage (in quantity of items)				
After 1 week	**63%**	**61%**	**77%**	
After 2 weeks	**47%**	**45%**	**18%**	
After 1 month	**74%**	**73%**	**51%**	
After 2 months	**91%**	**92%**	**71%**	
After 3 months	**93%**	**99%**	**100%**	
Percent of Items Delivered (in quantity of items)				
After 1 week	**6%**	**1%**	**4%**	
After 2 weeks	**9%**	**5%**	**2%**	
After 1 month	**33%**	**27%**	**8%**	
After 2 months	**48%**	**46%**	**19%**	
After 3 months	**67%**	**72%**	**47%**	
Donation-to-Delivery Time				
Mean (# days)	**33**	**35**	**29**	
Median (# days)	**28**	**31**	**24**	
Financial Efficiency				
(Donor Cost - Budget Cost) / Budget Cost	**-5%**	**-11%**	**30%**	
Actual CHF Spent - Budget CHF	**(3,510,849)**	**(5,209,538)**	**1,810,531**	
Transportation Cost / Total Product Cost	**10%**	**N/A**	**N/A**	
Assessment Accuracy: Revised Budget / Original Budget				
After 2 weeks	**131%**	**118%**	**365%**	
After 1 month	**139%**	**123%**	**377%**	
After 2 months	**148%**	**127%**	**493%**	
After 3 months	**158%**	**127%**	**493%**	

Appeal Coverage: In the Sahel operation, donors had completely satisfied the requests on the appeals list after two months. In the South Asia operation, donors had satisfied 92% of the total requests after three months.

For deliveries that were made to the destination country, however, the scorecards tell a different story. In the Sahel operation after three months, only 52% of the total items had been delivered which included only 25% of the Priority 1 items (food). In the South Asia operation after three months, 67% of the total items had been delivered which included 72% of the Priority 1 items (housing/shelter).

Donation-to-Delivery Time: It is interesting to note that the Sahel operation actually shows a lower mean donation-to-delivery time after three months, but based on the information above, one can determine that this mean will increase because a significant portion of the goods have not yet arrived at the three-month breakpoint.

In the South Asia operation, because the median value is lower than the mean value for the operation total, the housing items, and the kits & sets, this indicates that there are fewer outliers with a very long delivery time than if the median was higher than the mean. As a "rule of thumb," an organization would desire the median value to be close to the mean if not slightly lower, to ensure consistency in the delivery times of goods that are in the pipeline.

Financial Efficiency: In addition to the two financial efficiency metrics discussed in the Month 1 scorecards, in the final scorecard the transportation cost metric has been added. At present the current HLS system is limited to providing transportation cost information at the level of "one donor for one operation." Since the scorecards are not presenting information at the donor-level, this essentially has limited this analysis to only obtaining transportation costs for the total operation and not for the individual types of items (e.g. food). A future enhancement of this system would be to capture transportation cost data at the item-type level throughout the operation.

Given this limitation, it should also be noted that at the time of publication of this thesis, the IFRC was still in the process of receiving transportation cost data from donors of the South Asia operation, which explains the low value of 10%. 53% of the donors had not yet reported to the IFRC their transportation costs for this operation at the time this analysis was performed.

Assessment Accuracy: One can see that after two months of the Sahel operation, the list of items appealed for did not change any further. In the South Asia operation, however, the list of items appealed for continued to grow over time, even as late as three months from the original appeal date.

The next chapter will take a deeper look at analyzing the contexts of each operation and will discuss some of the factors that influenced operational performance. These graphs and findings may be replicated for any major emergency IFRC operation in the future, since the data used to construct them was extracted from HLS.

5 In-Depth Analysis of Two IFRC Operations

Chapter 4 presented the operational results of the two operations based on the metrics defined, but the causes and contexts of each disaster must be more deeply explored in order to further understand each scorecard and to draw conclusions about each operation's performance. First, the basic facts and background of each disaster are given in order to provide the context of the relief operation which followed. The HLS data is then used to gain further insights about these operations to determine the common themes discovered from the two and how the framework of metrics will enable some of the "lessons learned" to be applied to future operations.

5.1 Basic Facts & Background of Two Operations

5.1.1 Sahel Food Security Crisis

In the African region of Sahel, which encompasses Niger, Mali, Mauritania, and Burkina Faso, severe food shortages were reported in the summer of 2005, due to a combination of low rainfall and a locust invasion in 2004. The IFRC launched its appeal on July 22, 2005, but it had been monitoring the situation since August of 2004. An estimated number of 8 million people were reported to be at risk in this region due to food shortages, with the most threatening shortages occurring in Niger. The IFRC committed to assisting 222,000 people for six months. In this food security crisis, the IFRC was coordinating not only with fellow Red Cross and Red

Crescent national societies, but also with various agencies of the United Nations such as the World Food Programme, the World Health Organization, and UNICEF (IFRC, 2005).

5.1.2 South Asia Earthquake

On October 8, 2005, an earthquake with a magnitude of 7.6 on the Richter scale struck 95 kilometers northeast of Islamabad, Pakistan. Tremors were felt throughout South Asia from Kabul, Afghanistan, to Delhi, India. Aftershocks continued for the next several months, some measuring 6.0 on the Richter scale (IFRC, 2005). The most recent estimate of casualties indicate that over 73,000 people were killed by this earthquake, and another 3.5 million people were rendered homeless (Ozsoy and Nelson, 2006).

Due to the geographical difficulties of accessing villagers in the mountains of northern Pakistan, this operation has posed significant challenges to humanitarian organizations in terms of scope and delivery. Assessing the true extent of the earthquake's damage was a massive challenge in this disaster, because so many remote villages were affected which already had limited access to any outsiders, whether they were officials from the local governments or personnel from non-profit organizations.

Another key item to note about this operation is why so many people were rendered homeless and why housing / shelter items were the highest priority for the IFRC to deliver. A study performed by the Earthquake Engineering Center of the University of Engineering and Technology, Peshawar, indicated that approximately 60% of buildings in urban areas were constructed using unreinforced solid concrete. More than 60% of these buildings collapsed from the earthquake, which caused the majority of the casualties in this disaster (EERI, 2005). Hence, providing shelter for so many thousands of people rendered suddenly homeless before the onset of winter became the highest priority for the IFRC in this operation.

Because the extent of this earthquake's damage was so difficult to quantify, the organization changed their target population to serve as time passed after the initial earthquake. The preliminary appeal launched by the IFRC was on October 9, 2005, to aid 30,000 families for four months. As new information was received about the extent of the damage, this target increased to as much as 150,000 families in October of 2005, and the scope has since been narrowed to 81,000 families for six months (IFRC, 2006).

5.2 Comparison and Contrast of Logistics Performance in Sahel and South Asia

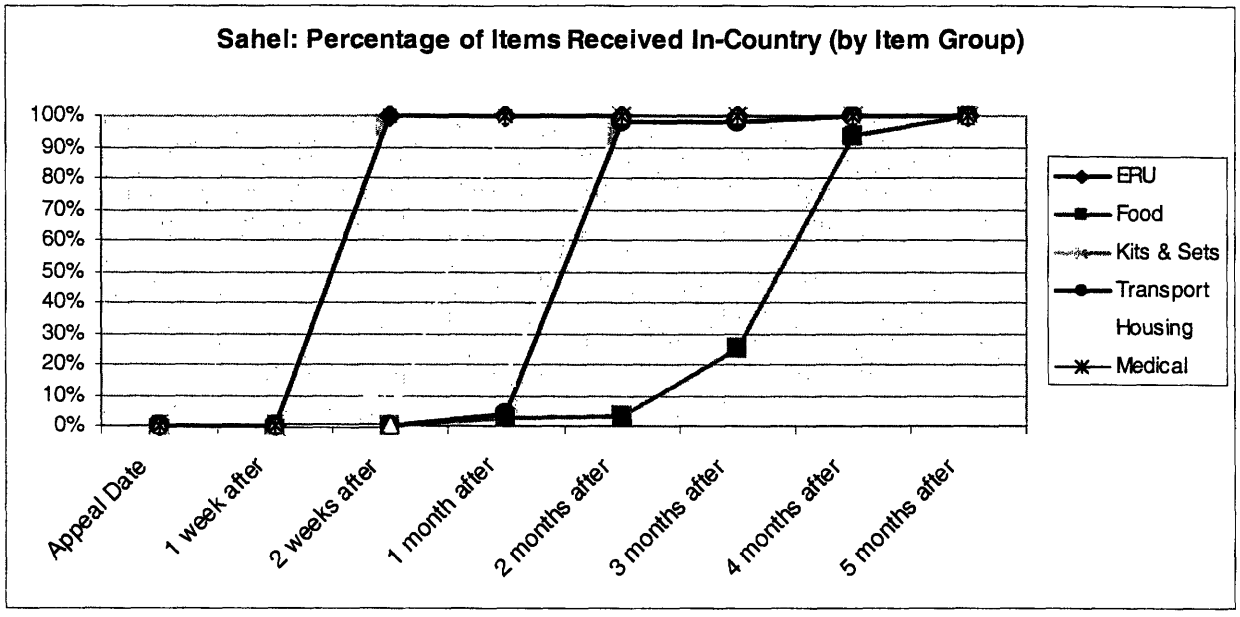
Based on the very nature of the Sahel operation being a food security crisis, food was clearly the type of item that was the top priority to deliver. However, the importance of explicitly identifying item-type priorities throughout the mobilization phase of an operation cannot be understated. In the Sahel operation, although food was the most important type of item to be delivered, it was the last type of item to arrive during the operation. Figure 8 below displays the length of time it took for 100% of all items of each type to be delivered after the original date the appeal was created. Based on discussions with individuals at the IFRC, transportation items were the second highest priority to deliver after food in this operation, because there was limited transportation in the area and the organization needed to be able to distribute goods to beneficiaries far inland into Africa.

Another key factor to consider with regard to the performance of the Sahel operation is that this was a slow-onset disaster. It appears that the IFRC decided to respond when it did due to the high malnutrition rates revealed from a nutritional survey that took place in the summer of 2005, in Niger. The coordinator of this operation also indicated in the IFRC press release that if the appeal was launched in July, the organization had a three-month "window of opportunity"

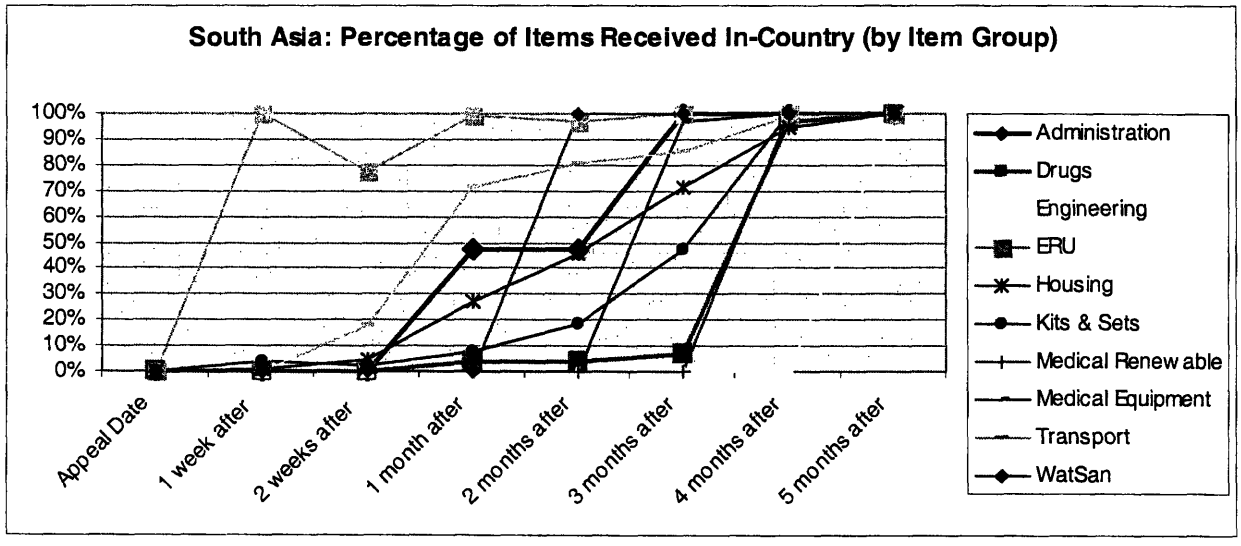
during which they could assist people in planting for the final harvest coming up in October (IFRC, 2005). Because media exposure gravitates toward sudden-onset disasters, a food shortage such as this one that has been seen coming for some time will typically elicit a slower response from donors as opposed to a sudden-onset disaster, such as an earthquake or a tsunami. For this reason, when an organization decides to commit to a particular operation, they must ensure that the targets they set are a reasonable match to the specific context of finding donors and delivering goods for that operation.

In the South Asia operation, delivering housing and shelter was clearly identified as the IFRC's top priority. Because of the scope of this disaster, however, which not only left thousands of people homeless, but also left them without any material possessions at all, kits and sets emerged as the second-highest priority for the IFRC. One of the largest challenges, however, was determining how many goods to get into the pipeline and how much to request from donors. Accurately assessing needs and identifying how many people the IFRC would commit to provide aid to was a significant challenge, because the magnitude of this earthquake was not truly understood by the humanitarian community until weeks and even months after it had occurred. However, the scope of this disaster prompted a strong worldwide response, which explains why the donations kept coming despite the increasing list of appeals, which grew rapidly throughout the first three months of the operation.

Figure 8 – Sahel & South Asia: Percentage of Items Received In-Country (by Item Group)



Note: in the above graph, there was only one shipment each for ERU items, housing items, and medical items, which explains why several line items in the graph go directly from 0% to 100%. In this operation, there were multiple shipments of food, kits & sets, and transport items.



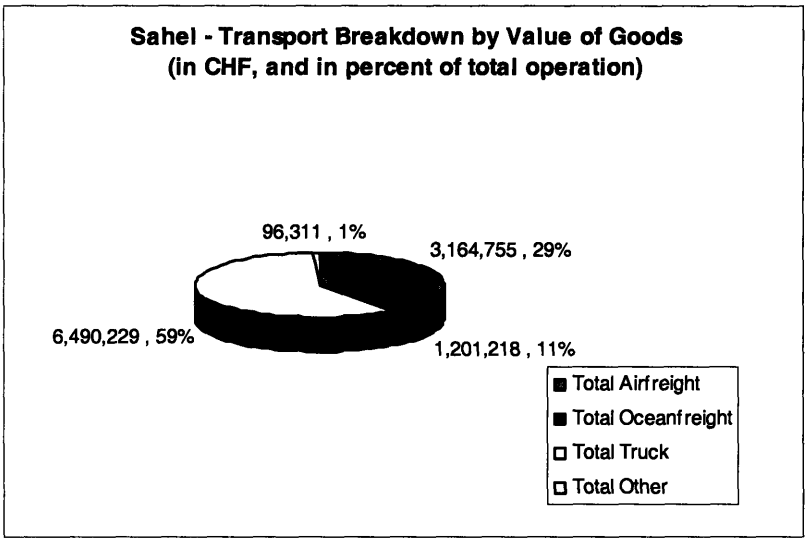
Note: In the above graph, certain types of items were added to the appeals list several months after the initial appeal date. For example, drugs, engineering items, and medical renewable items were added one month after the initial appeal date, which is why none of these items arrived until two months or more after the original appeal date.

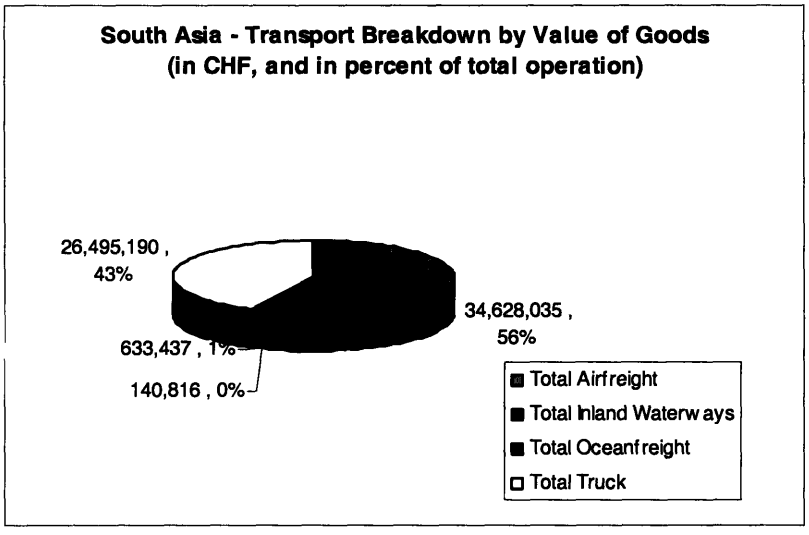
As shown in Figure 8, 50% of the Priority 1 items (food) in the Sahel operation were received 3.5 months after the original appeal date. In the South Asia operation, 50% of the Priority 1 items (housing) had been received 2 months after the original appeal date. For the Priority 2 items in Sahel (transport), 50% had been received 2 months after the original appeal date, while the Priority 2 items in South Asia (kits & sets), were received 3.5 months from the original appeal date.

To understand why the highest priority items were the last to be delivered in the case of the Sahel operation, and why the kits & sets took so long to arrive in the South Asia operation, the data can be broken down for further analysis. First, for Sahel, each consignment was analyzed by the method used for delivery: airfreight, ocean, truck, and other. As shown below in Figure 9, 59% of the total value of items shipped in this operation were delivered by truck, which accounts for a total of 6.4 million CHF of goods. 29% of the total value of goods were shipped via airfreight, which accounts for a total of 3.1 million CHF of goods.

In the South Asia operation, 56% of the total value of goods shipped were via airfreight, and 43% of the total value of goods were shipped via truck. These account for 34.6 million CHF and 26.4 million CHF of goods, respectively. The higher percentage of goods shipped via airfreight in the South Asia operation intuitively helps to explain why the high-priority items arrived at their destination faster.

Figure 9 – Sahel & South Asia: Transport Breakdown by Value of Goods





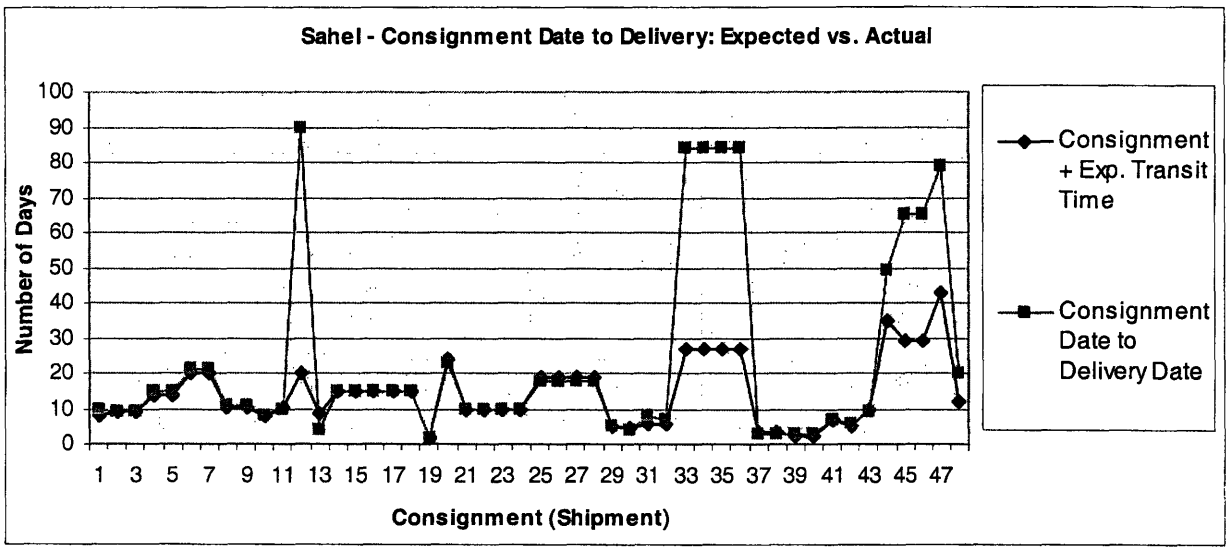
5.3 Lessons Learned from the Two Operations

5.3.1 Sahel Food Security Crisis

Further analysis of the Sahel data demonstrates not only that airfreight will move faster than trucks, which is obvious, but that some particular shipments of trucks arrived significantly later than anticipated. Using the data from HLS, one can compare the expected transit time of each shipment to the actual time in transit. Figure 10 below indicates that of the 48 shipments in

total for this operation, 39 shipments arrived on-time (or 81% of total shipments). 8 of the remaining 9 shipments missed the target delivery date by more than 30 days. The shipments shown below were sorted in order of the shipment date, so the graph presents a snapshot of the course of the operation from its beginning to end. In other words, the few shipments that arrived late arrived extremely late (> 30 days late), and these late shipments were the ones that contained the bulk of the food items.

Figure 10 – Sahel - Expected Delivery Time vs. Actual Delivery Time



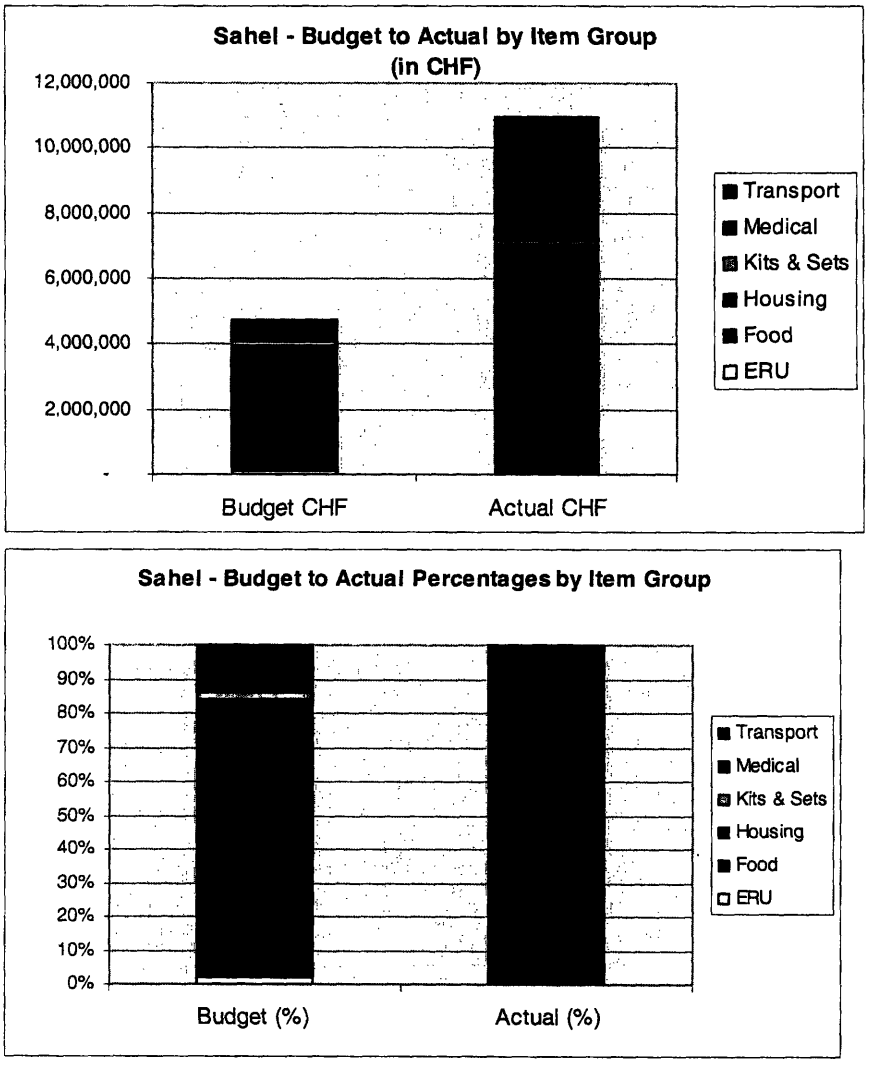
Note: Dark line represents expected length of shipment times, light line represents actual length of shipment times.

Figure 11 below provides a strong case for establishing official priorities of item families at the outset of an operation. The following two graphs indicate two basic facts relating to the outcome of the operation as a whole: 1) the value of the goods shipped during the operation exceeded the budgeted value and 2) that the money spent on each item group was not spent proportionally to the operation total as originally intended.

From the first graph in Figure 11, one can see that the original budget of the Sahel operation was 4.7 million CHF, while the actual amount of money that was spent during this operation was 10.9 million CHF. In the second graph shown below, one can see that the

percentage of the total monies spent in this operation towards food items was not as high as originally intended. Clear priorities should be established at the outset of an operation so that the logistics department and the donors will be better able to meet the original targets of how much the organization wants to spend in total on each type of tem.

Figure 11 – Sahel: Budget to Actual Values by Item Group in Dollar Form and in Percentages

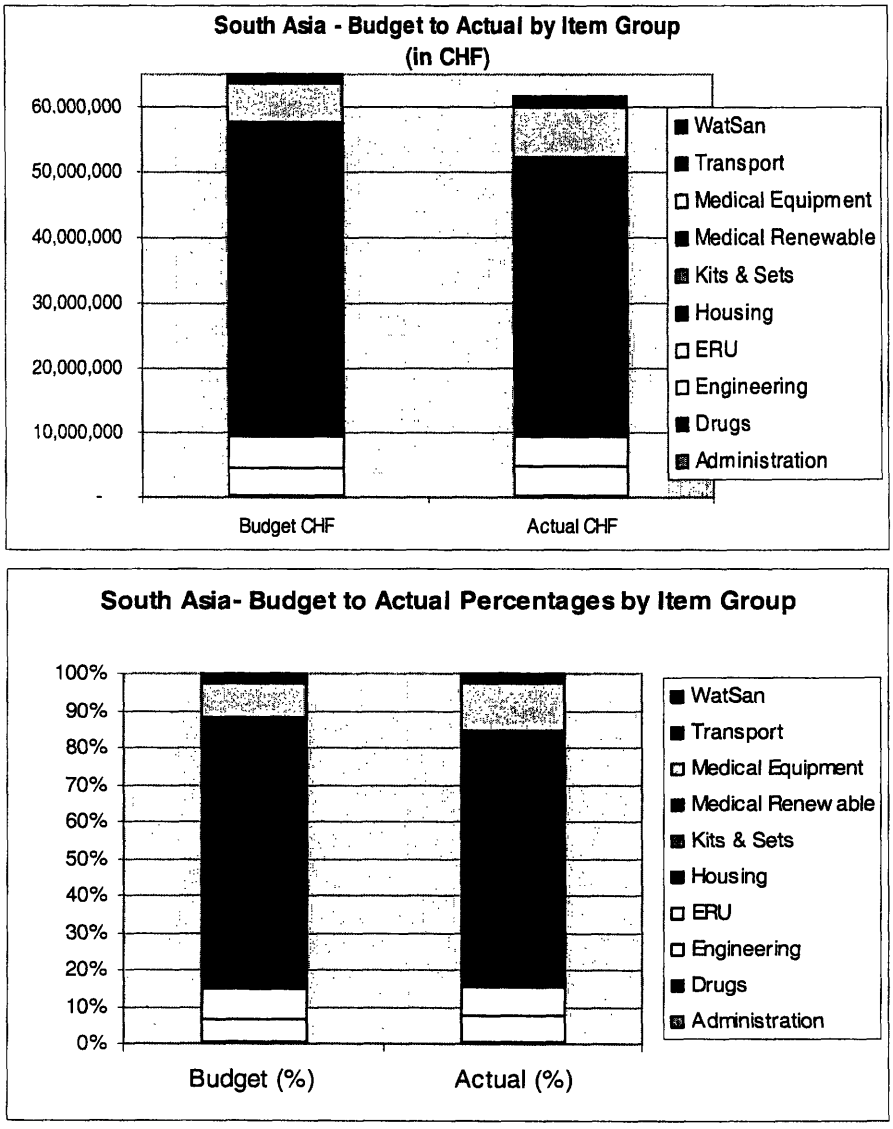


5.3.2 South Asia Earthquake

In the South Asia earthquake operation, a different story can be told relating to the adherence to the budget for the item-types that were the highest priority to delivery. As shown below in the first graph of Figure 12, the actual cost of items incurred was less than expected for

this operation (61.5 million CHF compared to 65 million CHF). However, despite the 3.5 million CHF difference, the second graph illustrates that the proportions of the various types of items were very close in comparing budget spend to actual spend. Housing items were budgeted to be 74% of total spend for this operation, and they actually accounted for 69% of the total spend. Based on these graphs, it can be inferred that the organization performed well in adhering to the original priorities of the items they wanted to deliver in this operation.

Figure 12 – South Asia: Budget to Actual Values by Item Group in Dollar Form and in Percentages



One of the benefits to the humanitarian sector of using a quantitative system of metrics is that issues noted anecdotally during an operation can now be examined using the data captured

to determine their true effects on the operation. In the past in many humanitarian organizations, observations that are made and reported by personnel in the field often make it into the final reports after an operation has been completed, but the extent to which issues have been noted are unclear because the data has not been quantified. Capturing operational data and being able to accurately report back to headquarters with it will ameliorate this problem. In this way, the true issues that affected the operation can be revealed and issues that were noted but revealed by the data to be small issues can be noted as such.

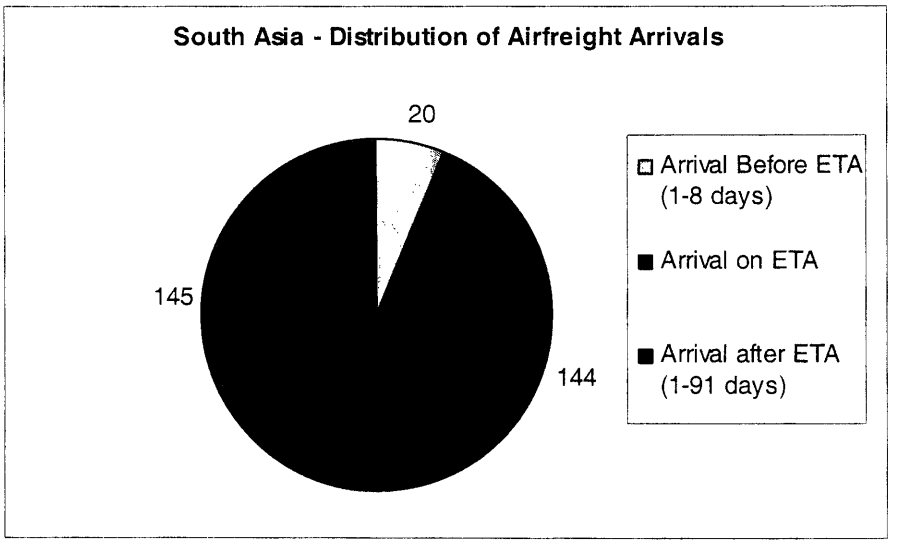
An example of how data can be used to determine the extent of an issue noted is the distribution of the arrival of airfreight shipments was noted in the final reports for the South Asia operation. This issue was noted both through interviews with individuals at the IFRC and in the "End of Mission Report – South Asia Earthquake Operation." The following is an excerpt from the "End of Mission Report" which explains this issue:

> "Despite the shipping instructions and the daily pipeline [report] sent on a daily basis to the field operations, a lot of planes arrived in Islamabad on the wrong date and at the wrong time which leads to a lot of frustration at the airport. In addition, some items came with poor or no documentation, quality of goods were not according to standards and some goods came unannounced or with wrong information. All this resulted in major difficulties to handle the goods, delays in customs clearance (as wrong or no documentation) and non-suitable items to deal with" (IFRC, 2006).

By sorting the HLS data into shipments by transportation mode, 309 shipments were recorded as airfreight recorded with a Goods Received Note. Of these shipments, 144 shipments arrived on the expected date of arrival. The remaining 165 shipments arrived either before the

scheduled arrival time or after it (see Figure 13). In this scenario where the airport is likely to be overcrowded by relief shipments, an arrival either before or after the expected date may cause complications . The shipments ranged from as much as arriving 8 days before the expected date of arrival to 91 days after it.

Figure 13 – South Asia: Distribution of Airfreight Arrivals



As demonstrated by the data, the issue noted by personnel in the End of Mission Report *was,* in fact, widespread in this operation, which affected over half of the airfreight shipments.

Understanding the true magnitude of the issues noted in operations may help to present a stronger case for change in the organization when an operation is being reviewed. By quantifying the issues of an operation through data analysis, the organization can then truly begin to build on its "lessons learned" in order to improve performance in the future.

5.4 Common Themes Discovered between the Two Operations

Based on the analysis performed in Sections 5.2 and 5.3, the following have emerged as the most important lessons to learn in order for a framework of metrics to be successfully implemented and for performance to subsequently improve:

1. There is a clear need at the outset of an operation to establish priorities of item groups as well as delivery and financial efficiency targets. Although it will take more time on the "front-end" of an operation, it will save time over the course of the operation so that the logistics department does not have to communicate individually with donors what the priorities are at various points in time. This will also ensure that the organization will be able to meet their goals of how much they desired of each item group to deliver in an operation.
2. The need for complete and accurate data must be stressed at a high level of any non-profit humanitarian organization. If the scorecards are to be interpreted accurately by the logistics department, the budget prices for *all* items listed in the appeal must be input into the system. Additionally, donors should enter transportation cost data completely and in a timely manner. Finally, the importance of entering the date of delivery (the GRN date) in a timely manner should be stressed to the field personnel.
3. Organizations should also consider regularly using an analysis similar to Figures 11 and 12 to help them understand if they are really delivering the items that they have committed to providing, based on the identified priorities.

6 Implementing the Framework

6.1 Applying the Framework to Other Organizations

As described in Section 4.1.2, the proposed framework has been tailored specifically to the needs and business processes of the IFRC. However, since other non-profit humanitarian organizations work toward a similar strategic goal of distributing goods to beneficiaries after disasters as quickly and effectively as possible, this framework can be adapted and applied to them as well.

The IFRC works primarily by having field personnel assess the needs of the population after a disaster and report these needs back to headquarters, in the form of specific items and quantities required. The logistics department at the IFRC headquarters then posts the list of required items to donors, which is what constitutes the "appeals list." Oftentimes the specific goods are donated as in-kind donations; otherwise, money is donated directly to the IFRC, which then procures the goods from suppliers. The goods are then shipped to the destination country.

Based on discussions with individuals at the Fritz Institute who work with a multitude of non-profit humanitarian organizations, this process is performed somewhat differently elsewhere. Typically, needs are assessed in the field and reported back to the organization's headquarters. Management at headquarters then determines a total amount of money they wish to raise from donors to contribute towards the operation. This dollar amount of money is then communicated to donors. Donors respond by contributing money directly to the organization,

which procures the goods and ships them to the destination country. The basic difference is that a higher percentage of monetary donations occur in other organizations, rather than in-kind donations.

To account for this additional step in the relief chain of raising funds before procurement can occur, an additional metric needs to be added to the proposed framework as one of the performance indicators of appeal coverage. This metric will allow the organizations to see how quickly they are reaching their fundraising goal:

$$\frac{\textit{Pledges Received (in Dollars)}}{\textit{Total Amount Requested (in Dollars)}}$$

The second metric of appeal coverage, shown as *Percent of Appeal Coverage* (in quantity of items) on the IFRC framework, should then change to *Percent of Items Procured* (in quantity of items). This metric would allow the organization to have visibility of how quickly they are able to find suppliers for the goods they wish to purchase after donors for the money have been found. The third metric of appeal coverage from the IFRC framework, Percent *of Items Delivered* (in quantity of items), will remain the same in the universal framework.

The metric of *Donation-to-Delivery Time* also needs to be altered in the universal framework. Since the items are not donated as in-kind donations, but rather the organization is procuring the items from suppliers, the length of delivery time should begin to be measured after the order has been placed with the supplier. Hence, this metric appears on the universal framework as *Procurement-to-Delivery Time*. It is possible that there may be a delay, however, from when an organization receives money and when they actually place an order with a supplier. In this case, the *Procurement-to-Delivery Time* metric will not accurately capture these inefficiencies. However, it is impractical to begin measuring the lead time when the money is received because not all monies received have been specified at that time to purchase certain

goods. Despite the limitation of not measuring those potential inefficiencies, *Procurement-to-Delivery Time* will capture a significant portion of an organization's total procurement process.

The adjusted framework is shown in Figure 14, which should be able to better closely match the business processes of other non-profit humanitarian organizations.

Figure 14 – Proposed Framework for Application in Other Organizations

Operation Name Appeal Date				
Status Update: Final ***Current Date:***	**Operation Total (Weighted)**	**Priority 1 Item Group**	**Priority 2 Item Group**	**Total Op Target**
Pledges Received / Total Amount Requested (in $)				
After 1 week	%	N/A	N/A	%
After 2 weeks	%	N/A	N/A	%
After 1 month	%	N/A	N/A	%
After 2 months	%	N/A	N/A	%
After 3 months	%	N/A	N/A	%
Percent of Items Procured (in quantity of items)				
After 1 week	%	%	%	%
After 2 weeks	%	%	%	%
After 1 month	%	%	%	%
After 2 months	%	%	%	%
After 3 months	%	%	%	%
Percent of Items Delivered (in quantity of items)				
After 1 week	%	%	%	%
After 2 weeks	%	%	%	%
After 1 month	%	%	%	%
After 2 months	%	%	%	%
After 3 months	%	%	%	%
Procurement-to-Delivery Time				
Mean (# days)	# days	# days	# days	# days
Median (# days)	# days	# days	# days	# days
Financial Efficiency				
Actual Cost - Budget Cost) / Budget Cost	%	%	%	%
Actual $ Spent - Budget $	$	$	$	$
Transportation Cost / Total Product Cost	%	%	%	%
Assessment Accuracy: Revised Budget / Original Budget				
After 2 weeks	%	%	%	%
After 1 month	%	%	%	%
After 2 months	%	%	%	%
After 3 months	%	%	%	%

Note: In the framework above, the first metric added to the Appeal Coverage indicator has been marked "N/A" in the "Item Group Priority" columns. This is because an organization typically raises funding for the operation as a whole, not for specific types of items. Targets of how

quickly the organization would like to raise the total funds for an operation should go in the "Total Op Target" column.

6.2 Feasibility Issues of Implementing the Framework

Implementing a framework of metrics to be used in the logistics department of a non-profit humanitarian organization implies much more than simply gathering data in the form of a scorecard. There is an entire cultural shift that will need to occur within an organization in order for a system of metrics to be successfully implemented. First, because non-profit organizations are not used to measuring performance, people should be well-educated about any such system before one is implemented. There may be resistance from either personnel within the organization or from donors that additional time is being spent on metrics that will take away from the time spent delivering aid to beneficiaries. It should be stressed to personnel *and* to donors that metrics will emphasize that time, money, and efforts should be spent on the "back end" of an operation so that the "front end" of an operation—the part where the final beneficiaries receive aid—is performed faster, more reliably, and more efficiently.

Another cultural shift that will have to take place for metrics to be successfully implemented is increased communication between the various departments of a non-profit organization. Even though the purpose of this proposed framework is to measure the performance of the logistics function, the logistics department on its own cannot successfully implement any such system as a stand-alone effort. The logistics department should be more involved in the front-end planning of an operation so that they are made aware from the outset what the overall organizational goals relating to an operation are and what types of goods are the highest priority to mobilize from donors and deliver. Furthermore, since the scorecards rely on timely and accurate information obtained from various departments, the support of these

departments is needed in order for the framework to be successfully implemented. For the IFRC, the critical links of communication appear to be between the Logistics Department, the Regional Operations department, the Operations Support department, and the Finance Department.

7 Opportunities for Future Research

7.1 Expansion Possibilities Given Additional Data

The four key indicators described in this research will be more valuable in the future when certain additional data can be captured by the information system. For example, computerized logistics systems such as HLS should be designed to capture changes to the appeals list on a transactional level, so that one can see exactly how the appeal changed over time and not just what the quantities requested were at certain points in time. Second, as mentioned in Chapter 4, transportation costs are currently captured in HLS only on a "one donor per operation" basis. If these costs could be tracked at the shipment level and reported to the centralized logistics function during the operation, this would greatly increase the value of the financial efficiency metrics. Finally, using the computerized system to track delivery data to the final beneficiaries would also add value to the performance measurement system. The scorecards would then be able span the entire length of the supply chain, from procurement to final distribution.

There are many opportunities for future research to increase the value of the proposed framework of metrics. The universal framework presented in Chapter 6 can be used as a starting point for other organizations to build their own framework of metrics, based on their organization's goals and business processes. Future opportunities can be explored in shaping and

manipulating the scorecard itself to make it specific to whatever purposes the organization would like to use it for. For example, comparing results across different operations may be very important to an organization, and there may be better ways to present the scorecard so that information is more easily comparable across operations. One such way to make the metrics more comparable across operations is to eventually shift from metrics that are "absolute" to metrics that are more "relative." The *Donation-to-Delivery Time* metric, for example, currently measures the length of delivery lead time in the number of days (an absolute metric). As organizations gain visibility over their supply chain and the delivery process becomes more standardized over time, the scorecard could progress to measuring the standard deviation from the expected lead time (a relative metric), which would be easier to compare across multiple operations.

There are also opportunities of future research related to how organizations should set their performance goals. As mentioned in Chapter 3, an organization should determine what their baseline performance levels are and if these baselines vary between operations based on external factors (such as disaster type and geographic region). There also may be better ways for an organization to emphasize metrics or item families that are particularly relevant to their specific goals. For example, the IFRC seeks to be a world leader in providing housing and shelter items just as Oxfam International seeks to be a world leader in providing water and sanitation items (Oxfam, 2006). These two organizations could potentially alter their scorecards to include a specific metric which would address such goals in all of their operations.

Finally, there may also be opportunities for research relating to change management and business processes in non-profit humanitarian organizations and how to better foster the "cultural shifts" mentioned in Chapter 6, that will orient organizations towards measuring themselves in

the future. The change management aspect of implementing any system of measurement is absolutely critical, because after all, it is people that make up non-profit organizations. The success of implementing both a framework of metrics and fostering a cultural shift towards measuring performance is directly correlated with the amount of the support such an initiative receives from the people within the organization, and particularly from top management of an organization.

7.2 Conclusion

The purpose of this research was twofold: to identify the best practices of performance measurement systems from military and commercial logistics and to apply them in the development of a framework of metrics to be used by non-profit humanitarian organizations in their disaster relief operations. Although these organizations will continue to be challenged by several factors which are inherent to providing aid after disasters, there is reason to look to the future with optimism. The development of information technology systems that can provide visibility to the multiple actors in the supply chain of relief items is a huge step forward for the sector as a whole. As more organizations begin to adopt and implement these systems and visibility is established throughout their supply chains, the use of key performance indicators will then be essential in further improving these supply chains. By clearly defining operational targets and then measuring actual performance to these targets, organizations will be able to better retain the lessons they learn from each operation and provide a higher level of service to their beneficiaries in the future.

Bibliography

About Oxfam International. (2006). Retrieved on April 27, 2006, from http://www.oxfam.org/en/about/what/.

Beamon, B.M. (2004). Humanitarian Relief Chains: Issues and Challenges. *The 34th Annual International Conference on Computers and Industrial Engineering*, November 14-15. San Francisco, CA.

Brauner, M., & Lackey, A. (2003). CWT and RWT Metrics Measure the Performance of the Army's Logistics Chain for Spare Parts. Retrieved December 5, 2005, from http://www.rand.org/publications/RB/RB3035/index.html.

Caplice, C., & Sheffi, Y. (1994). A Review and Evaluation of Logistics Metrics. *The International Journal of Logistics Management, 5* (2), 11-28.

Caplice, C., & Sheffi, Y. (1995). A Review and Evaluation of Logistics Performance Measurement Systems. *The International Journal of Logistics Management, 6* (1), 61-74.

Dumond, J. (2000). Define-Measure-Improve: The Change Methodology that has Propelled the Army's Successful Velocity Management Initiative. Retrieved December 5, 2005, from http://www.rand.org/publications/RB/RB3020/index.html.

Earthquake Engineering Research Institute – EERI (2005). *Pakistan: A summary report on Muzaffarabad earthquake*. Retrieved on February 24, 2006, from http://www.reliefweb.int/rw/RWB.NSF/db900SID/KKEE-6HWRYR?OpenDocument.

Edmonson, R. G. (2005). Special Tactics: Disaster-relief Logistics Borrows from Military and Commercial Experience. *Journal of Commerce*. September 19, 2005, 1.

Guyoton, S., & Muirhead, B. (2002). Building a Balanced Scorecard for the International Federation of Red Cross and Red Crescent (IFRC). Unpublished masters thesis, INSEAD, Fontainebleau, France.

IFRC (2005). *Pakistan: Earthquake Emergency Appeal - Appeal no. 05EA022 revised on October 25, 2005*. Retrieved on April 13, 2006, from http://www.ifrc.org/cgi/pdf_appeals.pl?05/05EA022-revised.pdf.

IFRC (2005, July 22). *Burkina Faso, Mali, Mauritania and Niger: Food Insecurity Emergency Appeal - Appeal no. 05EA015*. Retrieved on April 13, 2006, from http://www.ifrc.org/.

IFRC (2005, July 22). *Red Cross / Red Crescent to Launch Appeal in Response to Niger Food Crisis*. Retrieved on April 13, 2006, from http://www.ifrc.org/docs/News/pr05/4305.asp.

IFRC (2006). *Pakistan: Earthquake Operations Update - Appeal no. 05EA022* revised on March 28, 2006. Retrieved on April 29, 2006, from http://www.ifrc.org/cgi/pdf_appeals.pl?05/05EA02224.pdf/.

IFRC (2006). *End of Mission Report: South Asia Earthquake Operation.* Briefing Note - International Federation of the Red Cross and Red Crescent Societies.

Johnson, H., & Kaplan, R. (1987). *Relevance Lost: The Rise and Fall of Management Accounting*. Boston, Massachusetts: Harvard Business School Press.

Kaplan, R., & Norton, D. (1992). The Balanced Scorecard: Measures that Drive Performance. *Harvard Business Review, 70* (1), 71-79.

Lambert, D. (2001). Supply Chain Metrics. *The International Journal of Logistics Management, 12* (1), 1-20.

Meyer, M. (2005). Can Performance Studies Create Actionable Knowledge if We Can't Measure the Performance of the Firm? *Journal of Management Inquiry. 14* (3), 287-291.

OCHA (2005). *Humanitarian Response Review – Commissioned by the United Nations Emergency Relief Coordinator and Under-Secretary-General for Humanitarian Affairs.* Retrieved on November 9, 2005, from http://www.reliefweb.int/rw/rwb.nsf/db900SID/EVOD-6FUDDK?OpenDocument.

Ozsoy, A., & Nelson, A. (2006). Pakistan: Isolated quake survivors eager to rebuild in mountain villages. Retrieved April 13, 2006, from http://www.ifrc.org/Docs/News/06/06040701/index.asp.

Russell, T. (2005). The Humanitarian Relief Supply Chain: Analysis of the 2004 South East Asia Earthquake and Tsunami. Unpublished masters thesis, Massachusetts Institute of Technology, Cambridge, MA.

SCC, Inc. (2004). *Supply Chain Operations Reference Model, Version 6.1.* Supply-Chain Council, Inc.

Sheffi, Y. (2005). *The Resilient Enterprise: Overcoming Vulnerability for Competitive Advantage*. Cambridge, Massachusetts: MIT Press.

Thomas, A., & Ramalingam, V. (2005). Lessons from the Tsunami: Top Line Findings. Retrieved September 28, 2005, from http://www.fritzinstitute.org/PDFs/Programs/Findings_Sept2605.pdf.

Appendix - Abbreviations and Definitions

CWT – Customer Wait Time

GRN – Goods Received Note

HLS – Humanitarian Logistics Software

IFRC – International Federation of the Red Cross and Red Crescent Societies

RWT – Requisition Wait Time

SCOR Model – Supply Chain Operations Reference Model

SSA – Supply Source Activity

UNOCHA – United Nations Office for the Coordination of Humanitarian Affairs