A Balanced Scorecard Approach to Performance Measurement in the Government Sector

*Terry R. Collins, Ph.D., P.E., Manuel D. Rossetti, Ph.D., and Julie A. Watson, M.S.
Industrial Engineering Department
University of Arkansas

Abstract

This research study presents the utilization of the balanced scorecard methodology as a means to measure performance for the Defense Logistics Agency Distribution Centers (DLA-DDC). The most critical performance metrics were selected and were placed into one of four scorecard categories of Cycle-Time, Quality, Financial, and Resource. A performance metric ballot has been constructed to select specific performance metrics critical to the mission of the Depot.

Keywords
Performance Metrics, Balanced Scorecard, Scorecard Ballot

1. Introduction
This paper presents partial results of a research study which have selected critical performance metrics, and positioned these metrics into associated categories in a Balanced Scorecard (BSC). This research study was conducted at the University of Arkansas under the direction of The Logistics Institute (TLI), and the United States Defense Logistics Agency (DLA). The DLA-Defense Distribution Center (DDC) is a branch of the Department of Defense. The DLA-DDC serves the materiel needs for United States military installations around the world. The DLA-DDC has 24 depot facilities strategically positioned across the US and abroad to receive, stow and issue over 4 million National Stock Numbers (NSN's). The combined value of the NSN's at all depots exceeds 8 billion dollars.

The Logistics Institute is an Industry/University Cooperative Research Center co-jointly administered by the University of Arkansas and Georgia Tech Industrial Engineering Departments. The program began in 1982 as the Material Handling Research Center (MHRC) at Georgia Tech. In 1991, The Logistics Institute and the University of Arkansas joined the MHRC, and in 1994 the MHRC changed officially to The Logistic Institute partnering Georgia Tech and the University of Arkansas.

Since there are several interpretations for the terms "performance metric" and "balanced scorecard" it is appropriate to provide clear definitions used in the context of this study. A Performance metric consists of a set of analytical tools that take measurements, display recordable results, and the ability to initiate actions based on the measurement results [8, 10]. Performance measurement is essentially comprised of several criteria consisting of: effectiveness, efficiency, quality, productivity, quality of work life, innovation, and profitability [2,9].

The balanced scorecard approach was developed by Kaplan and Norton [5] in 1995. Since then this decision analysis tool has been used in many different applications with the expected result being a systematic approach to measure performance of an organization. Therefore, by definition, the balanced scorecard approach is a planning and goal-setting management process that enables organizations to focus on long-term objectives.

The remainder of this paper will discuss how the performance metrics are selected and the methodology of grouping the various metrics into related categories in the balanced scorecard. The performance metric ballot and balloting process will also be presented. Finally, this paper will discuss continuing research efforts for this project with the utilization of a multiple attribute utility theory (MAUT) to consider tradeoff characteristics when using a balanced scorecard approach.

2. Project Goals and Objectives

There were several initiatives for this project regarding the development of performance metrics for use in the strategic decision making process for DLA. The first and foremost goal of the project was to develop a core set of appropriate, balanced, and robust performance metrics for the DLA depots. In close association to developing the set of performance metrics, a secondary goal was to identify the necessary information technology architecture/analysis tools that were available. This would enable DLA to categorize the most important performance metrics such that the complexity of monitoring and tracking the metrics would be manageable.

Once the goals of the project were identified, the next phase was to set clearly defined objectives to meet the expectations of the study. The project objectives are presented below:

1. analysis of the current DLA performance metrics with industry and the DLA mission to identify critical metrics relevant to the strategic decision making process,

2. accumulate and categorize a core set of critical performance metrics for DLA use,

3. identify and select the most appropriate methodology for integrating the performance measures for use in the strategic decision making process, and

4. analysis of current DLA information systems will be required to determine the level of data support for the proposed methodology.

3. Methodology

In order to satisfy the aforementioned objectives a thorough literature review was needed to identify the most appropriate method for collecting, categorizing, and presenting the results of the selected performance metrics. Choosing which metrics to monitor and track is typically one of the most difficult tasks in developing a performance measurement system. Previous TLI research by Watson, Malstrom, and Landers explored and identified all known possible performance metrics that could possibly be used by DLA-DDC [10]. TLI researchers gathered preliminary information on performance metrics through direct contact with DLA personnel and a TLI sponsored workshop with industry leaders.

The study generated a total of 247 performance measures that were considered metrics. This best practices logistics performance evaluation study reviewed earlier research work which provided insight on the methodologies used to administer a best practices studies [1,3,6,7]. Most of this research had a logistics focus, however, all of the research applied to the general strategic decision making mission for the DLA.

Utilizing all 247-performance metrics would be well beyond the capabilities of any performance metric evaluation system. Therefore, the best practices study consolidated duplicate metrics and organized the metrics into four common categories. These categories were cycle time, quality, financial, and resource. Captured within each category was the matching performance metrics. In particular, the cycle time grouping of metrics focused on how responsive DLA was to meeting customer needs. The quality metrics specifically dealt with service quality. Monitoring both, short and long-term profitability was the purpose of the financial metrics. Finally, using resource metrics in the performance measurement system helped address the delicate issues of depot capacity and facility utilization.

Once the performance metrics were categorized, a methodology for collectively evaluating the metrics was needed. In review of the research literature there were two performance measurement evaluation models which provided the best fit with the objectives of the research study. These two models were Kaplan and Norton's Balanced Scorecard approach [5], and the Oregon Productivity Matrix Model developed by the Oregon Productivity Center at Oregon State University [4].

The Balanced Scorecard provided the best fit with the ability to incorporate company missions and strategic management into a tool that monitored and tracked performance against set goals. The best and most distinguishing characteristic about the Balanced Scorecard model was it evaluated past performance as outcome measures and

integrated these measures with long-range strategic management metrics which tend to drive future performance. Therefore, the Balanced Scorecard was said to be balanced in the sense that companies looked at the short and long term goals of the organization. Figure 1 illustrates how the Balanced Scorecard maps the strategic mission of the organization to the Balanced Scorecard categories.

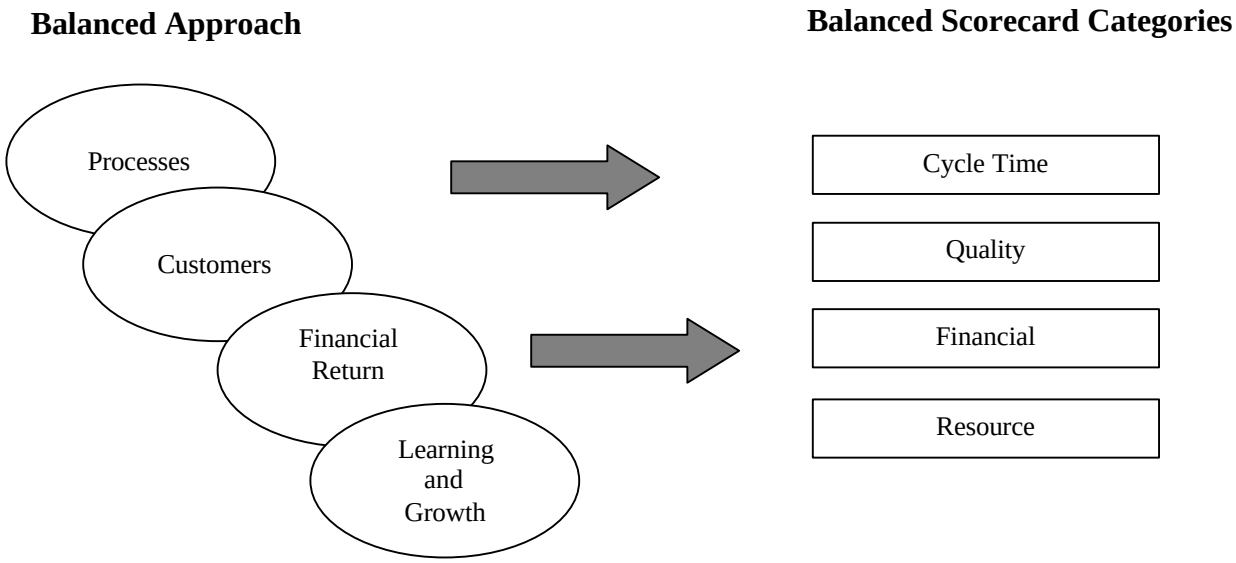

Figure 1. Using the Balanced Approach to determine the Balanced Scorecard Categories

4. Performance Metrics Ballot

The selection of the Balanced Scorecard methodology as the performance measurement system was very appropriate as it systematically organized the performance metrics into groupings by category. The next step of the research process for the project was to develop a survey instrument for identifying a few critical performance metrics for each category of the balanced scorecard. The 97-performance metrics needed to be reduced to approximately 10-15 in each of the four Balanced Scorecard categories. Once these metrics are pared down they will be placed in the appropriate category of the Scorecard and tracked on a set interval by the DLA Depot facility. Tables 1-4 show the selected metrics from the four categories of the balanced scorecard approach. A complete set of metrics can be referenced in the TLI Best Practices Logistics Performance Evaluation Study Final Report [10].

Table 1. Cycle Time Group – Metrics

SUBGROUP	METRIC	UNITS	VOTE ✓
Distribution/ Filling	Cycle sub-time-distribution/filling	Distribution/filling time	
	Stock-to-non-stock ratio	%Material shipped by regular stock	
	Cycle time-total (full stream)	Elapsed time order entry->material available & visible in computer	
	Number of locations/dates touched	Listed locations with most recent touch reason and date	
	Date item last touched	Cumulative time since most recent receipt, shipment, inventory, etc.	
Souring	Point-of-use deliveries	#Shop deliveries/total deliveries	
	Supplier direct deliveries	#Deliveries from supplier/total deliveries	
	Throughput rate	WIP/cycle time	
	Cycle sub-time-transportation	Transit time	
Transportation	Expedite ratio	#Shipments expedited/total shipments	
	Off line shipments	#Off-line shpts/total shpts	

The survey instrument selected for the study was a ballot-style questionnaire. The intent of the ballot was to have Depot Commanders check in each category the performance metrics that were most critical to the mission of the DLA Depot. Particular attention was given to the performance measurement issues of customer satisfaction, service, responsiveness, type of customer (civilian/military), and the characteristics of the items shipped. The initial components of the ballot were developed in an earlier TLI study with DLA-DDC [10].

Table 2. Quality Group - Metrics

SUBGROUP	METRIC	UNITS	VOTE ✓
Defect-free	Claims ratio	#Shipment claims/total shipments	
	Complete orders	#Complete orders/total orders	
	Correct destination	#Orders delivered to correct dest./total orders	
	Damage free (concealed)	#Orders with no concealed damage/total orders	
Information Integrity	Forecast accuracy	MAD,MSE, bias	
	Inventory accuracy	%Stock pts same contents quantity/items as records	
	Record accuracy	#Erroneous records/#records	
	Tracking accuracy	#Entities in known status/total entities	
	On-time delivery	#On-time deliveries/total orders	
On-Time	On-time loading	#On-time loaded orders/total orders	
	Absenteeism	Lost time/total time	
	Accident rate	Accidents/(unit:miles,employee,empl.days worked)	
Partnering	Associate/employee retention	#Associates or employees retained/total employed over a time pd	
	Diversity	%Minority/disabled	
	Supplier partnership	Supplier rating index	

Table 3. Financial Group - Metrics

SUBGROUP	METRIC	UNITS	VOTE ✓
Cost	Annual cost of maintenance by operator	Cost/operator-year	
	Cost per operation	Cost/activity	
	Cost per piece	Cost/piece	
	Cost per transaction	Cost/transaction	
	Cost per unit of throughput	Cost per unit of throughput	

Table 4. Resource Group – Metrics

SUBGROUP	METRIC	UNITS	VOTE ✓
Capacity	Trailer/tractor ratio	Trailer/tractor ratio	
Productivity	Asset turnover	Sales/assets	
	Pack rate	Orders packed/person-hr	
	Productivity - on road	Miles/truck/day	
	Revenue or profit per square foot	Revenue or profit/sq. ft.	
	Shipments per associate or employee	#Shipments/associate or employee	
Utilization	Asset utilization	Capacity used/capacity available	
	Cube utilization (load factor)	Cube used/cube available	
	Downtime	Downtime/(Operating time+idle time+downtime)	
	Empty miles	Empty miles/total miles	
	Empty trailers/containers	(Empty trailers/containers)/(total trailers/containers)	

5. Balloting Process

The collection of ballots is still in process. The results of the ballots will be formalized into the Balanced Scorecard for the DLA-DDC. Once the balloting portion of the study is complete the fourth objective of the study will be addressed. DLA information systems will be evaluated to determine whether the selected performance metrics can be provided for analysis using the Balanced Scorecard method. The final results of the study will be presented in a future research paper.

6. Conclusions

This paper presents the current findings of a research study being conducted for the DLA-DDC. The results presented in this paper have identified the best performance measurement system to be used at the DLA Depots. Using the Balanced Scorecard method, the DLA-DDC will be able to track strategic performance metrics and make decisions that reflect the mission of the Depot.

A ballot-type questionnaire has been developed to determine the most important performance metrics for the Depots. The selected metrics from each category will be used in the Balanced Scorecard. One unique feature on the Balanced Scorecard performance measurement system is that current metrics can be easily replaced with new metrics as the mission of DLA changes in time. This enables the system to be flexible and attentive to the time-based strategic management objectives of the DLA-DDC.

Thus far, this study has been very successful in developing a performance measurement system for the DLA-DDC. The four objectives mentioned in the introduction section of this paper are in the process of being satisfied. The first three objectives are near completion, and the fourth objective will be met once the performance metrics for the Balanced Scorecard has been identified.

7. Extended Research Activities

In addition to the completion of the balloting process, continued research is underway to consider the tradeoff aspects when using the balanced scorecard approach. The decision analysis method of Multiple Attribute Utility Theory (MAUT) will be used to wage "what-if" situations between performance levels within the balanced scorecard system. The continued research activities will be administered by a TLI research team located at the University of Arkansas.

The use of MAUT as a decision tool to evaluate tradeoffs will be an important extension of the current research. MAUT will assist in identifying and modeling the interdependencies and interactions between a selected set of performance metrics and the strategic goals that are site specific to the individual DLA Depot locations. Tradeoffs can be analyzed considering the dependencies between metrics as they relate to strategic decisions of the Depot.

These strategic decisions will be based on several scenarios. The first scenario is the differing strategic missions between Depots that serve specific branches of service. A second scenario is the variation in strategic decisions between peace-and wartime situations where the strategy to supply the customer will change dramatically. And, the third scenario will be to evaluate internal tradeoffs within the same Balanced Scorecard based on a peacetime strategy to improve depot operating performance.

For example, in a peacetime situation more weighting may be placed on the performance metrics in the financial group category of the BSC. The use of MAUT would also evaluate the tradeoffs between different metrics within the same Balanced Scorecard category. In a wartime situation, higher weightings may be placed on the cycle time and quality group categories of the Balanced Scorecard. Tradeoffs for each of these scenarios would be beneficial to the overall performance of the DLA-DDC initiative, and support the dynamic strategies of supplying materiel to DLA-DDC customers around the world.

References

1. A.T. Kearney, Inc., 1990, *Improving Quality and Productivity in the Logistics Process-Achieving Customer Satisfaction Breakthroughs*, Oak Brook:CLM
2. Brinkerhoff , R. and Dressler, D., 1990, "Productivity Measurement: A guide for Managers and Evaluators," *Sage Publications*, Newbury Park.

3. Cohen, M., Zheng, Y., and Agrawal, V., 1997, "Service Parts Logistics: A Benchmark Analysis," *IIE Transactions*, N29, 627-39.
4. Felix, G., and Riggs, J., Autumn,1983, "Productivity Measurement by Objectives," *National Productivity Review*, 386-93.
5. Kaplan, R. and Norton, D., 1995, *Translating Strategy into Action: The Balanced Scorecard*, Boston, MA, Harvard Business School Press.
6. Michigan State University,(1989, *Leading Edge Logistics Competitive Positioning for the 1990s*, Oak Brook: CLM.
7. Michigan State University, 1997, *World Class Logistics: The Challenge of Managing Continuous Chan*ge, Oak Brook: CLM.
8. Rose, K.,1995, "A Performance Measurement Model," *Quality Progress*, V28, N2, 63-6.
9. Sink, S., Tuttle, T. and DeVries, S., Summer,1984, " Productivity Measurement and Evaluation: What is Available," *National Productivity Review*, 265-87.
10. Watson, J., Malstrom, E, Landers, T.,, Dhodpaka, S., Smith,V., and Harris, R., 1999, "Best Practices: Logistics Performance Evaluation," *The Logistics Institute Final Report*, TLI-MHRC-98-01