ONLINE BENCHMARKING FOR TRANSPORTATION PROVIDERS

Manuel D. Rossetti, Ph. D., P.E.
Tat-Yuen Sang
Terry Collins, Ph. D., P.E.
Department of Industrial Engineering
University of Arkansas
4207 Bell Engineering Center
Fayetteville, AR 72701
Phone: (501) 575-6756
Fax: (501) 575-8431
rossetti@uark.edu

SUMMARY

This paper discusses a system for automating the performance benchmarking of transportation service providers. The system allows for the identification, evaluation, collection, and analysis of performance metric information in a secure on-line data-driven environment. The current system has a total of seventy-nine metrics arranged in a balanced scorecard classification. In addition, the system allows transportation service providers to perform a multi-attribute decision technique to investigate best-in-class performance. This paper details the metrics and the system functionality.

INTRODUCTION

In order to remain competitive, transportation service providers, such as trucking firms, must constantly evaluate their operations, management structures, information systems, and customer relations with respect to their competitors. One of the standard approaches used by industry to evaluate performance is known as benchmarking. Benchmarking is a process by which companies determine best practices that lead to exceptional performance; however, benchmarking can be a complicated and time-consuming process. First, the data necessary to support a benchmarking analysis can be difficult and time-consuming to obtain. For example, companies may be reluctant to share the data necessary to determine best practices because of proprietary or competitive concerns. Second, performance metrics required to perform the benchmarking can be difficult to determine and difficult to interpret their relevance to best practices. Third, finding the best practice in a subject area can be a lengthy task that requires expertise and analysis techniques.

This research examines the development of an on-line benchmarking system to facilitate the benchmarking of transportation providers. We have implemented the system in prototype form as an "on-line benchmarking system (OBS)". The system allows transportation providers to develop performance metrics to be used during a benchmarking analysis. In particular, the system allows transportation providers to recommend performance metrics, rate their importance, classify the metrics into service categories, and develop other specifications for a performance measurement system. In addition, the system allows for the on-line collection of performance metric values through a on-line survey process. The data is collected and stored in a database for later benchmarking analysis against other participating benchmarking parties in an anonymous fashion. Finally, the system facilitates best practice analysis through a multi-criteria process based on a balanced score card approach. The metrics are categorized according to the four categories in the balance scorecard approach to performance measurement.

In this paper, we present the details of the collected performance metrics and the structure of the online benchmarking system in the context of trucking carriers. Finally, methods that we use to prioritize the performance of carriers are presented. We begin with a review of literature relevant to the benchmarking of transportation service providers.

BACKGROUND LITERATURE

According to Zivan (14)(15), then Xerox's vice president for logistics and distribution, a pioneer in developing benchmarking processes, benchmarking is the heart of the planning process for any company that holds customer satisfaction as its highest priority and benchmarking is an integral part of any quality improvement process. Performance metrics are used to indicate the performance of an organization within a benchmarking analysis and within performance measurement systems. According to Watson et al. (5), performance metrics can be defined as the analytical tools in the performance measurement process that take measurements, display results, and determine subsequent actions. The specific value of a metric indicates the performance of a specific area in an organization.

In general, an enterprise will have hundreds of potential performance metrics to be incorporated into their performance measurement systems. Because of the large number of potential performance metrics, it is often very use to classify the metrics into subject areas. For example, Watson et al. (5) proposed a comprehensive logistics performance framework and a best practice template in their work. A total of one hundreds and twelve metrics were identified and categorized into four groups, which were "cycle time", "quality", "financial", and "resource". In addition, a list of eighty-two best practices was complied from two hundred and sixty six best practices that matched the appropriate metrics. In addition, their work "identifies benchmarking that links qualitative values (best practices) with quantitative measurement (performance metrics, across the value chain". (5)

In a general classification hierarchy, metrics can be classified into two main categories: operational performance metrics and financial performance metrics. Wouters et al. (3)

discusses performance metrics identification and classification, especially for non-financial performance metrics in the transportation and distribution areas. They argue that the strategy-selection approach was often not practical for companies who wish to develop non-financial measures but did not yet have a clear strategy. A strategy-selection approach uses non-financial measures to translate strategy into more specific objectives that provide guidelines for operational action for middle and lower management, and to measure whether the organization is in fact successfully implementing the desired strategy. (3) As a result, Wouters et al. recommended that benchmarking be used as a tool to facilitate the selection of non-financial performance measures. In this approach, important non-financial measures were selected by analyzing statistical data from a benchmarking study. In addition, the benchmarking analysis also indicated which non-financial measures were most associated with financial performance.

Wouters et al. produced two models, which involved both the transportation and distribution aspects of a logistics system. In these models, they identified the relationships that linked the operational measures and financial measures. In addition, the models presented mathematical relationships between the financial and operational measures. Then, they used a one-factor correlation analysis to identify the critical operational measures by comparing the average values for best-performers and discussing with experts. We have incorporated a number of the metrics identified by Wouters et al. into our performance metric framework.

In a related article, Donselaar et al. (2) discuss similar performance issues in the trucking industry. The article addressed the importance of using performance indicators other than financial because of the growing trend of quality improvement in many industries. A research project known as BRAVO of Eindhoven University in The Netherlands concentrated on logistics performance from the transportation providers' point of view. The objective of the project was to develop a list of performance metrics that was critical to the operations of transportation service providers. The metrics were categorized into transportation and distribution groups. Both groups were further broken down into short distance and long distance groups. The results in the short distance group showed that both average wages per driver and average load in a truck were the most important metrics that highly correlated to the performance of companies. Improvement made to these metrics can result in better operational performance; therefore, the metrics were classified as a "Critical Success Factors (CSF)". On the other hand, the results in the long distance group showed that "the percentage of Km driven empty" and "distance traveled per trip" were considered as CSFs. Therefore, improvement made to these metrics can directly improve operational performance. CSF and some other metrics that were presented in this study were included in our performance metric list.

Caplice et. al.(8) identified eight criteria to consider when selecting performance metrics for logistics and business purposes, which were: 1)Validity, 2)Robustness, 3)Usefulness, 4)Integration, 5)Economy, 6)Compatibility, 7) Level of detail, 8)Behavioral soundness. In addition, they suggested three forms of measurement that can be used to capture the performance of business activities, which are 1) Utilization, 2) Productivity, and 3)

Effectiveness. (8) Another study by Tang et. al.(7) applied Analytical Hierarchy Process to measure the performance of outbound logistic processes within Taiwan's notebook computer industry. By conducting a survey, they identified seven key performance metric areas that closely related to logistic activities: 1) information system, 2) customer service, 3) order process, 4) distribution process, 5) assembly process, 6) inventory control, and 7) warehouse management.

Boyd and Cox (9) used a techniques known as the "negative branch" to evaluate the cause and effect relationships of performance metrics within an organization. According to them, the construction of "Negative branch" is a four step process: "1) Write down the positive effects that are expected to result from the action, 2) Write down the negative effects that might result from the action, 3) Connect the proposed solution with your suspected positive and negative effects by cause-and-effect relationships, and 4) Read the negative branches from bottom up using if-then logic, scrutinizing every statement and logical connection along the way, and make necessary corrections." (9) For example, if on-time delivery is a metric, then drivers may be forced to exceed the time allowed for consecutive driving hours in order to consistently meet on-time delivery. This process helps to identify trade-offs between the metrics.

Kaplan and Norton (11) created a performance metrics framework known as the balanced scorecard. The balanced scorecard is a "multifaceted tool for communications, alignment, improvement, and control that integrates critical non-financial performance measures into the basic management structure of the organization". (13) The framework was divided into four perspectives: customer, internal business, learning and growth, and financial. The learning and growth perspective includes employee training and corporate cultural attitudes related to both individual and corporate self-improvement. (12) The internal business perspective refers to internal business processes. Performance metrics that based on this perspective enable the managers to know the status of their business. Furthermore, the metrics tell whether the products and services conform to customer requirements. (12) On the other hand, the customer perspective is based on the importance of customer satisfaction in all businesses. (12) Finally, the financial perspective is based on timely and accurate financial data from the management. According to Brewer and Speh (10), this framework *"balances the inclination to overemphasize financial performance by incorporating metrics related to business process measures, innovation and learning measures, and customer satisfaction measures"*. They developed a method that applied the balance scorecard to measure the performance of supply chain

The performance metrics identified for transportation providers within the on-line benchmarking system are classified into the balanced scorecard framework in order to facilitate the understanding and dissemination of the metrics. In the next section, we present the performance metrics that we identified for use within the on-line benchmarking system.

PERFORMANCE METRICS

In order to facilitate online benchmarking, we need a comprehensive list of performance metrics in all areas of interest for transportation providers. In order to develop a comprehensive list of potential metrics, we examined literature related to trucking service providers and discussed metrics with experts within the industry. After identifying potential metrics, we categorized the metrics into the four categories associated with the balance scorecard approach to performance measurement. The grouping of metrics was subjectively based on their functionalities and properties. For this research, we collected a total of seventy-nine metrics. By using the balanced scorecard groupings, we attempted to cover all the major operations within a company. Table 1 presents the collected metrics.

Table 1: Performance Metrics

Internal Business Process	Financial Measures	Learning and growth	Customer
Loading size	Direct cost per truck	Absentee rate	Claims settling rate
Average length of haul	Maintenance cost /mile	Average # of workdays	% of active customer
Average speed	Driver's gross pay	Driver's weekly work hours	Delivery accuracy
Operating hour/Truck/day	Operating Revenue/wages	Average nights a driver's home/week	% Correct order
%Distance driven empty	Revenue/driver	Training received	% Order returned
Trailer loading rate	Revenue/(driving time)	Recognition	Delivery date compliance
Trailer unloading rate	Revenue/Trip	Accident rate	Re-purchase rate
Load Factor	Revenue/truck	Availability rate	% Order on time
Miles/Truck/week	Operating Revenue/distance	Education	Service termination rate
Average loaded miles	Revenue/hour	Turnover rate	Claims rate
Average empty miles	Maintenance Cost	# Grievances/year	Complete order rate
# Stops per route	Logistics operating expenses	Recruitment rate	Damage free rate
% Routes repeated	% Return on asset	truck operator's experience	On-time delivery
Average MPG	% Return on investment	# Truck operators	On-time order entry
Trailer, tractor ratio	Revenue growth rate	Employment opportunities/year	On-time loading
% Driven not empty	% Market share	% Offers accepted	Correct destination rate
Time driven between 2 stops	Rental Expenses	Promotion rate	Data Entry Accuracy
Tracking accuracy	Insurance expenses		
Inventory record accuracy	Annual inventory usage		
Pick rate	Inventory asset value		
Fill rate/quantity	Operating revenue/year		
Material handling rate			
Inventory Turnover			
% Complete sales			
Loads/trailer/year			
Average # of dispatches/load			

As indicated in Table 1, the metrics are classified according to the categories in the balance scorecard. The metrics were mainly collected by reviewing related literature and by using

the suggestions of experts within industry. The metric table is not rigid. In fact, one can add more metrics into the table so that the collection can be more comprehensive. The metrics shown in the table were evaluated by the transportation and logistics experts to ensure their validity and accuracy.

For each metric, we develop a detailed metric description that includes a name, a textual description of the metric, a formula for computing the metric if applicable, the measurement units for the metric, the purpose of the metric, and a place for comments or discussion about the metric. The following is an example of such a description:

Name:	Loading Capacity
Definition:	Maximum load per truck in tons
Formula:	N/A
Units:	Tons
Purpose:	Used to identify the loading capacity of trailers.
Discussion:	This metrics is needed to derive other metrics, such as "Average loading truck".
Collection Frequency:	2/year
Collection Method:	Written report

In the following section, we discuss the structure of our online system.

ON-LINE BENCHMARKING SYSTEM

The objectives of the on-line benchmarking system are to facilitate performance metric identification, automate the collection of performance metric evaluation information, automate the collection of performance metric data, and provide analysis tools for the comparison of performance between participating benchmarking partners. The system is data-driven and can be assessed by using a web browser. For example, a company that searches for a few useful metrics can visit our website and perform a query search on the metrics database. In addition, they can complete surveys to provide feedback on the metrics. Finally, they can complete surveys to collect metric values in order to compare their performance with other participating companies. We designed the supporting database using the unified modeling language (UML) in order to meet the aforementioned objectives. The on-line benchmarking system was constructed by using VBScript, an HTML scripting language and active server pages (ASP). The on-line benchmarking system hosts the performance data of transportation providers, performance metrics, surveys, SMART rating application, and links to other benchmarking information. In addition, the system supports data security so that all the information provided is confidential. Figure 1 presents the home page for the on-line benchmarking system application.

Figure 1: The On-line Benchmarking System

After the metrics were collected, we organized them into a categorized table (as shown in Table 1). Instead of using a complex hierarchy of metrics categorization, we apply the simple and well-known grouping of the balanced scorecard. All of the metrics can be viewed online and they can be searched according to specific criteria. For instance, a manager may search for all the metrics that are in "customer" groups and with the collection frequency of less than a year. The online system supports a number of queries to facilitate search. Specifically, one may search the metrics database by category, collection frequency, key words or name of the metrics. The result of the search can be viewed online with full details.

PERFORMANCE METRIC IDENTIFICATION AND EVALUATION

In order to facilitate the evaluation process of the metrics, we created a web interface to support online metric evaluation so that experts in industry or any knowledgeable individual can review and comment on the metrics. As one may notice, the interface features a set of questions that incorporate Caplice et. al.'s (8) eight metric evaluation criteria. The following is an example of the metric evaluation questions:

Name: Loading size
Definition: The physical load size of a loaded trailer
Unit: tons

1) Please rank the importance of this metric to your company's operation:
   Least 1 2 3 4 5 most important
2) Does your company collect the information in this metric?
   Yes_____ No_____
3) Do you understand the definition of the metric? Yes____ No____

4) How often do you collect this metric? _____
5) What is the method of collection for the metrics?_____
6) Validity: Does this metric capture the events and activities being measured? Yes____ No____
7) Robustness: Is this metric comparable across time, location, & organization and is repeatable? Yes____ No_____
8) Usefulness: Is this metrics understandable and provide a guide for strategic action? Yes____ No_____
9) Integration: Does this metric include all related aspect of a process and encourage the coordination across functions? Yes____ No_____
10) Economy: Do the profits obtain from using this metric out-weight the cost of metric collection? Yes____ No_____
11) Compatibility: Is this metric compatible with the existing information, material, cash flows and systems in the organization? Yes____ No_____
12) Level of Detail: Does this metric provide a sufficient degree of detail to the user? Yes____ No_____
13) Behavioral Soundness: Does this metric minimizes incentives for counter-productive acts or game playing, and is presented in a useful form? Yes____ No_____

The first five question deal with the general information of the metrics, while the remaining questions captured Caplice's eight metric evaluation criteria. In addition, as suggested by Caplice et. al., it is often not practical to develop metrics that encompass the eight comprehensive criteria. There are always some trade-offs between the criteria. (8) Failing to recognize the trade-offs may produce inaccurate result and contribute to strategy/management errors. In our system, we created an interface to include the trade-offs between the criteria. Specifically, we provide a web interface so that any possible side effect associated with the use of a metric can be input or contributed by a system user. The inputs include the name of the affected metric and the description of the "side effect". This incorporates a functionality that supports Boyd and Cox's negative-branch process.

PERFORMANCE METRIC DATA COLLECTION

In order to collect the data needed for the online analysis tools, we created online surveys so that performance data can be collected. The database was structured so that respondents are not identified. The questionnaire was designed by referring to the performance metrics list. The questions were mainly designed to obtain values for all the metrics in the list. The surveys are divided into five categories, which are "Carrier Categorization", "Customer Perspective", "Internal Business Perspective", "Learning and Growth Perspective", and "Financial perspective". Note that in addition to the four balance scorecard groups, "Carrier Categorization" was added to the survey so that an accurate type of carrier can be recognized. For example, a trucking company may be categorized as a less than truckload and regular-route carrier, who delivers hazardous material only. The following is a question taken from the Internal Business Process Survey:

Q12.What is your total mileage operated in intercity service annually?

[ ] Loaded Miles

[ ] Empty Miles

We wish to stress that the website is dynamically data-driven. In other words, it is supported by a database and as users interact with it. The results of the surveys are collected into a survey database. Queries can be performed so that information can be extracted from the database. Query creation is facilitated through the use of drop down list boxes. For example, one may create a query to extract the information on the number of empty miles of all irregular route carriers in the database. Then, the user must select empty miles on "metrics list menu" and carrier type from "carrier list menu" and click on the "search" button. The following diagram show the important concepts and their relationships of the database:

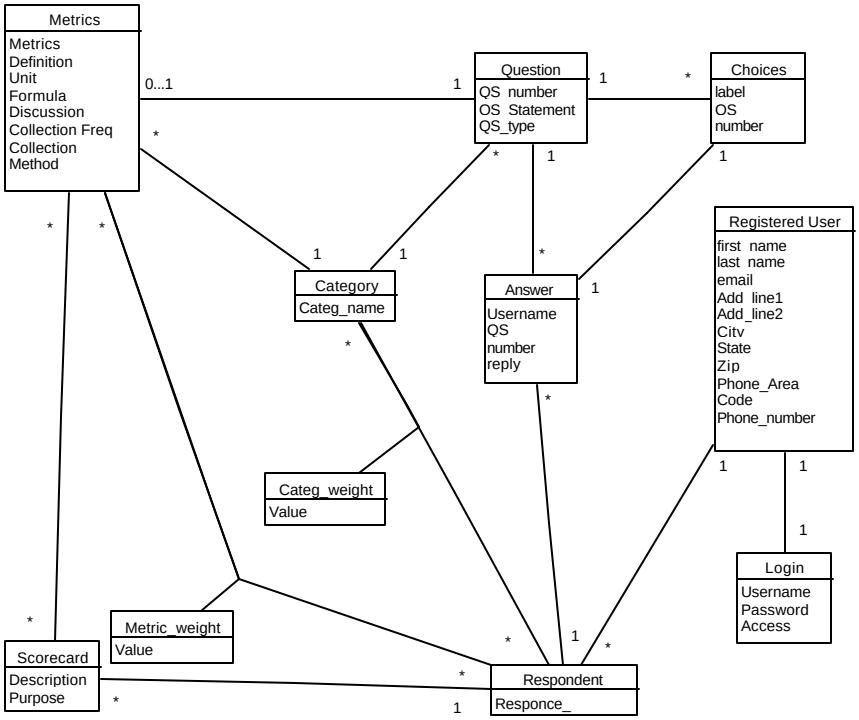

Figure 2: Database Design Diagram

Associations are indicated by an adorned line between entities. Each end of an association is adorned with the multiplicity of the relationship. Multiplicity indicates the number of instances of one entity type that may relate to a single instance of an associated entity type. For instance, the association between the "Question" and "Answer" entities indicates that a question can have more than one answer; however, each answer belongs only to one question. Each entity consisted a number of attributes, which are represented with a labeled name when mapped to a table. The value of attributes can be collected online and can be dynamically change or edited via a web interface. A user has to register online in order to use the system and the user has to choose a username and a password. All users who respond to the survey will be added to the respondent table. The survey questions are stored in the question table. Question types include multiple choice (choices are stored in "choice" table), open ended, and single choice. All questions are related to one category. The "category" table records all the category types. A response to a question is called an answer

and it will be stored in the "Answer" table. Metrics are stored in the "metrics" table and each metric is related to one question. Since the SMART method is used for performance analysis, a user must assign weight to a selected list of performance metrics and categories. The weights will be stored in the database as well. Then, the SMART calculations are performed via server-side scripts and the results are generated and displayed online.

MULTI-ATTRIBUTE ANALYSIS

Transportation providers may be interested in comparing their performance to other providers using the system. The online system provides an application that allows a trucking service provider to compare itself against others in terms of overall performance. The method of comparison is known as the SMART method or "Simple Multi-Attribute Rating Techniques". SMART is a rating techniques that is capable of handling more than one attribute. The method requires the user to assign weights to attributes and calculations are performed to determine the overall "aggregate benefit" of a metric. The higher the value of aggregate benefit, the better the performance of a company. The SMART method consisted of the following two steps: 1) Identification of crucial performance indicators and 2) Multi-criteria rating and ranking.

After all the necessary metrics are collected, we have enough attributes to perform SMART analysis. Note that *not all* metrics need to be used in the analysis. There are two reasons why we want to use a limited number of metrics:

a) Practicality: It may not be practical to use all the metrics because some companies simply do not collect the metrics.
b) Complexity: Having too many metrics increases the complexity of the SMART analysis by making more difficult to assign and interpret weights.

During the process the user provides performance data for the metrics and rates the metrics Due to the diversity of trucking companies (e.g. truckload, less than truckload, regular route carrier, private carrier, etc.), some companies may consider different performance metrics as crucial in indicating their performance. Therefore, the metrics can be dynamically changed according to a company's preference or situation. In the website, we allow the user to pick his/her own metrics as the attributes. Then, the user must enter the value of the metrics so that "step 2" can be carried out.

After the performance metrics have been determined, the usual SMART calculation can be performed, namely,

a) *Determine the value function of each attribute*: In this prototype, all the value functions were estimated by assuming a linear relationship. A linear function was used for simplicity and because it does not imply a specific preference structure.
b) *Determine the aggregate benefits of each alternative:* Aggregate benefits of trucking companies (i.e. alternatives) can be computed at this step. The largest aggregate benefits value indicates the best overall performer.

The results of the SMART analysis can be represented via a "Graphical Dashboard". The dashboard, like the instrumental dashboard of a motor vehicle, indicates the performance of a company graphically based on the results obtained from SMART, see Figure 2.

Figure 2. Graphical Dashboard

CONCLUSIONS

The on-line benchmarking system for transportation service providers exploits Internet technology to provide more efficient benchmarking techniques through the storage of performance metric data, automation of data collection and sharing, and the provision of innovative tools for the comparison of benchmarking companies. The major purpose of the system is to facilitate the benchmarking process for transportation providers so that they may improve their operation performance. This research not only provides a system for benchmarking transportation service providers but also provides a comprehensive database of well-documented performance metrics for transportation companies. These metrics can be used as a starting point in conducting successful benchmarking projects. Transport service providers who participate in this system can share information in a confidential manner and compare their performance with other participants. The initial prototype of the system is focused on the trucking industry; however, the system is flexible enough to provide benchmarking analysis to other transportation companies. Future work on the system will involve developing more sophisticated methods for comparing system performance based on other multi-objective analysis techniques such as the analytical hierarchy process (AHP).

ACKNOWLEDGEMENTS

The work was supported through grant, MBTC 2015, through the Mack-Blackwell Rural Transportation Center at the University of Arkansas. The views expressed in this paper are solely those of the authors. We also like to thank those companies and their representatives who participated within our metric identification and evaluation process.

REFERENCES

1) C. A. Voss, P. Ahlstrom, and K. Blackmon, "Benchmarking and operational performance: some empirical results." *International Journal of Operations & Production Management*. Vol. 17, No.10, 1997, pp 1046-1058.
2) K. Van Doonselaar, K. Kokke, M. Allessie, "Performance measurement in the transportation and distribution sector." *International Journal of Physical Distribution & Logistics Management*. Vol. 28, No.6, 1998, pp434-450.
3) M. Wouters, K. Kokke, J. Theeuwes, K. Van Doonselaar, "Identification of critical operational performance measures - A research note on a benchmarking study in the transportation and distribution sector." *Management accounting research*, 1999, pp 439 – 452.
4) K. Van Doonselaar, G. Sharman, "An innovative survey in the transportation and distribution sector." *International Journal of Operations & Production Management*. Vol. 17, No.7, 1997, pp 707 - 720.
5) J. A. Watson, E. M. Malstrom, T. L. Landers, S. Dhodapkar, V. Smith, and R. Harris, (1999) "Best Practice: Logistic Performance Evaluation." The Logistics Institute Final Report, TLI-MHRC-98-0
6) L.-L. Tang and C.-C. Hon, "The performance measuring of outbound logistics for Taiwan notebook industry." *Proceedings of the 6th Annual International Conference on Industrial Engineering-Theory, Applications and Practices*. San Francisco, CA, USA, November 18-20, 2001
7) C. Caplice and Y. Sheffi, "A review and evaluation of logistics metrics." *The International Journal of Logistics management*, 1994, pp11-28
8) L. H. Boyd and J. F. Cox, "A cause-and-effect approach to analyzing performance measures." *Production and Inventory Management Journal*, 1997, pp25-31
9) P. C. Brewer and T. W. Speh, "Using the balance scorecard to measure supply chain performance." *Journal of Business Logistics*. Vol21, No.1, 2000. pp.75-92
10) R. S. Kaplan and D. P. Norton, "Putting the balance scorecard to work." *The Harvard review*, 1993, pp.134-147
11) The Balance Scorecard Institute "The Balance Scorecard" (07 November 2001) Retrieved December 26, 2002, from The Balance Scorecard Institute Website: http://www.balancedscorecard.org/bscit/intranet/bsc1.html
13) American Productivity and Quality Center "Balance Scorecards" (1994-2002) Retrieved from December 26, 2002, from the Productivity and Quality Center website: http://www.apqc.org/pm/scorecard.cfm
14) G. Scriven, "Benchmarking: An Overview The Theme for the Nineties." Agility Consulting, (1992, April) Retrieved from Agility consulting website: http://www.agility.com.au/benchmark-frame.htm