Evidence Based Library and Information Practice

Building Scorecards in Academic Research Libraries: Performance Measurement and Organizational Issues

Vivian Lewis
Acting University Librarian
McMaster University
Hamilton, Ontario, Canada
Email: lewisvm@mcmaster.ca

Steve Hiller
Director of Assessment and Planning
University of Washington Libraries
Seattle, Washington, United States of America
Email: hiller@u.washington.edu

Elizabeth Mengel
Associate Director, Scholarly Resources and Special Collections
The Sheridan Libraries
Johns Hopkins University
Baltimore, Maryland, United States of America
Email: emengel@jhu.edu

Donna Tolson
Director of Strategic Assessment Services
University of Virginia Library
Charlottesville, Virginia, United States of America
Email: djt5k@virginia.edu

Abstract

Objective – This paper describes the experiences of four prominent North American research libraries as they implemented Balanced Scorecards as part of a one-year

initiative facilitated by the Association of Research Libraries (ARL). The Balanced Scorecard is a widely accepted organizational performance model that ties strategy to performance in four areas: finance, learning and growth, customers, and internal processes.

Methods – Four universities participated in the initiative: Johns Hopkins University, McMaster University, the University of Virginia, and the University of Washington. Each university sent a small group of librarians to develop their Scorecard initiatives and identified a lead member. The four teams met with a consultant and the ARL lead twice for face-to-face training in using the Scorecard. Participants came together during monthly phone calls to review progress and discuss next steps. Additional face-to-face meetings were held throughout the year in conjunction with major library conferences.

Results – The process of developing the Scorecards included the following steps: defining a purpose statement, identifying strategic objectives, creating a strategy map, identifying measures, selecting appropriate measures, and setting targets. Many commonalities were evident in the four libraries' slates of strategic objectives. There were also many commonalities among measures, although the number chosen by each institution varied significantly, from 26 to 48.

Conclusion – The yearlong ARL initiative met its initial objectives. The four local implementations are still a work in progress, but the leads are fully trained and infrastructure is in place. Data is being collected, and the leadership teams are starting to see their first deliverables from the process. The high level of commonality between measures proposed at the four sites suggests that a standardized slate of measures is viable.

Introduction

> *A strategy without measures is just a wish and measures that are not aligned with strategy are a waste of time.* (Matthews, 2008)

The Balanced Scorecard is a widely accepted organizational performance model that ties strategy to performance in four critical areas: finance, learning and growth, customers, and internal processes. While originally designed for the for-profit sector, the Scorecard has been adopted by non-profit and government organizations, including some libraries. This paper focuses on the experiences of four prominent North American research libraries (Johns Hopkins University, McMaster University, the University of Virginia and the University of Washington) as they developed and implemented scorecards as part of a one-year initiative facilitated by the Association of Research Libraries (ARL).

This paper is divided into four major sections: an introduction to the Balanced Scorecard and its key components; an overview of the ARL initiative and the process used to develop scorecards at each library; an exploration of the concept of a standardized suite of measures for ARL libraries based on a commonality of key objectives; and a review of organizational challenges faced by the pilot sites during their implementations. The authors hope that the lessons learned and strategies employed at their institutions will assist other academic libraries choosing to implement the Balanced Scorecard.

What is the Balanced Scorecard?

The Balanced Scorecard was developed by Harvard Business School professors Robert S. Kaplan and David P. Norton in the early 1990's as a reaction to the industrial age emphasis on financial measures as the sole indicator of success. In their groundbreaking book, *The Balanced Scorecard: Translating Strategy into Action*, Kaplan and Norton argue that the economic realities of the information age require a more well-rounded set of measures to evaluate and drive an organization's performance:

> "The Balanced Scorecard is a new framework for integrating measures derived from strategy. While retaining financial measures of past performance, the Balanced Scorecard introduces the drivers of future financial performance. The drivers, encompassing customer, internal business process, and learning and growth perspectives, are derived from an explicit and rigorous translation of the organization's strategy into tangible objectives and measures." (Kaplan & Norton, 1996, p. 18)

The Balanced Scorecard model is premised on strong and very direct linkages between key planning elements. Each measure is directly aligned to one or more strategic objectives. Success in meeting targets is a clear indication that the organization is moving its mission forward. Linkages between measures (both within and across the four perspectives) help ensure that the organization maintains a truly "balanced" approach. In the same way, strategic initiatives are directly linked to the measures: only projects that improve an organization's success in meeting its targets are eligible for linkage to the scorecard (Kaplan and Norton, 1992).

Who uses the Balanced Scorecard?

While originally designed for the commercial sector, non-profit organizations have also been attracted to the model. Kaplan and Norton (1996) note that, "while the initial focus and application of the Balanced Scorecard has been in the for-profit (private) sector, the opportunity for the scorecard to improve the management of governmental and not-for-profit enterprises is, if anything, even greater" (p. 179). The Balanced Scorecard was first recommended for adoption by U.S. federal government procurement agencies during the Clinton administration. The City of Charlotte, North Carolina, and the United Way of Southeastern New England were also early adopters (Kaplan & Norton, 2001).

As the concept has matured, the pool of non-profit organizations exploring the use of the Balanced Scorecard has grown along with specialized expertise in the use of the model in specific settings. Ascendant Strategy Management Group, the consulting firm used for the pilot, has helped government and non-profit organizations like the U.S. Federal Bureau of Investigation, the U.S. Securities and Exchange Commission, the Atlanta Public Schools, and the Catholic Charities Archdiocese of Boston apply the Scoreboard to achieve change. Many of these organizations have had to quickly adapt to game-changing events such as 9/11 or the recent mortgage meltdown, and have turned to the Balanced Scorecard to promote successful organizational change.

Although the total number of libraries adopting the Balanced Scorecard is unknown, it is likely still only a few handfuls worldwide. In *Scorecards for Results: A Guide For Developing a Library Balanced Scorecard* (Matthews, 2008), examples of libraries with experience using the Balanced Scorecard include the Singapore Public Library and the University of Virginia Library. Aside from Virginia, which developed their scorecard in 2001 (Self, 2003), only a small number of academic libraries are known to have adopted this approach.

Libraries, Measures, and the Balanced Scorecard

While relatively few have adopted the Balanced Scorecard, libraries have a long tradition of collecting statistical and other measures related to organizational performance. For the most part, libraries collect *input measures*, the amounts of resources invested or put into the development and delivery of collections and services. Input measures traditionally deal with such categories as collections, facilities, staffing, budget, and more recently, technology. They count things such as the number of volumes, user seats, librarians, dollars, or computers. They form the basis of many of the regional or national statistical surveys where comparisons between libraries can be made. For example, the Association of Research Libraries Membership Index tracked the variables of number of volumes held, the number of volumes added during the year, number of current serials received, total operating expenditures, and total number of professional and support staff.

While input measures track the investment in library collections and services over time, they do not indicate if these resources and services are actually being used or how effective they were in meeting user needs. The use factor can be handled with **output measures** that count uses or transactions associated with library activities. These might include number of items loaned, number of reference transactions, instruction sessions, gate counts, computer log-ins, and Web site visits. Output measures are often used as surrogates for library effectiveness (i.e., an effective library is one that is heavily used). While these metrics do incorporate the user, they do not actually measure the impact specific services or resources had on that user. They also are not necessarily tied to any strategy or set of objectives.

Process measures, also used extensively in libraries, measure the activities related to turning inputs into outputs. Sometimes they are called efficiency measures as they calculate the amount of time per activity or the cost of that activity (for example, the average length of time to catalog a book or the cost of staffing a service point). Process measures can also have a customer component such as the average time it takes to order a book or answer a question.

Finally, *outcome measures* represent the effect or impact of a particular service or resource on the customer or what that service or resource enables the customer to do. Successful outcome measures are usually linked to objectives and goals, which may not be solely defined by the library. For example, if there is a learning objective for students to cite information correctly in their term papers, an outcome measure might be that 95% of citations are accurate. Durgan, Hernon, and Nitecki note that another goal of performance measurement is, "How well does the library serve the institutional mission and serve as an effective partner or collaborator?" (2009, p.38).

Brophy (2008) considers measuring library performance to have two basic goals: "How good is this library? How much good does this library do?" (p. 7). That said, performance measures in themselves are not sufficient to achieve these goals if they are not tied directly to overall organizational strategy and objectives. While an increasing number of libraries are developing and using measures that tie directly to achievement of strategic objectives (Franklin (2011) estimates at least 10% of ARL libraries are including metrics in their strategic plans), they are usually applied to specific areas and are neither balanced nor integrated. The Balanced Scorecard provides an opportunity, not only to integrate these library performance measures within a more structured planning process, but one that also connects to synergistic organizational performance.

At the international level, the potential impact of the Balanced Scorecard as an organizational performance model for libraries can be seen in Poll's and Boekhorst's second revised edition of *Measuring Quality: Performance Measurement in Libraries*. The authors' selection of 40 indicators

is based on 4 criteria, 1 of which is: "To cover the different aspects of the service quality as described in the Balanced Scorecard, including indicators for the aspect of development and potentials" (Poll & Boekhorst, 2007, p. 9). Poll and Boekhorst use the term "indicators" rather than performance measures and, citing ISO 11620, note that good indicators are informative, reliable, valid, appropriate, practical, and comparable.

The Association of Research Libraries Initiative

The Association of Research Libraries (ARL) advances the interests of the major research libraries in the United States and Canada. It has established a strong and multifaceted assessment program to enhance understanding of current and future trends in academic libraries and to assist member institutions in meeting their strategic objectives. ARL places a strong focus on evidence-based decision making, and creating a culture of assessment. It has facilitated the introduction and use of many tools for building this capacity, including LibQUAL+® and MINES for Libraries®.

With the difficult economic climate and increased requirements for accountability throughout the higher education sector, the need to enhance member libraries' capacity for driving change has become even more crucial. ARL decided to explore the Balanced Scorecard as a key tool for measuring performance and leading change within member institutions. As noted in Kyrillidou (2010), ARL intended to accomplish two tasks: "to assist, train and facilitate the use of the Scorecard in a small number of ARL libraries; and to test the value of a collaborative model for learning about and implementing the new tool" (p. 33).

In late 2008, ARL put out a call to its members for expressions of interest in participating in a one-year exploration of the Balanced Scorecard. The initiative was described as "an investment in helping libraries make a stronger case for the

value they deliver by developing metrics that are tied to strategy" (Association for Research Libraries, 2009).

The initial November 2008 meeting ultimately produced four universities keen to participate: Johns Hopkins University, McMaster University, the University of Virginia and the University of Washington. The four institutions brought a wide spectrum of experiences to the table. The University of Virginia Library had used the Balanced Scorecard for a number of years, but was interested in refreshing their implementation and providing assistance to the new sites. The University of Washington had a strong assessment program, but no experience with the Scorecard. Johns Hopkins and McMaster had developing assessment programs and no past experience with the Scorecard.

Each university sent a small group of librarians to develop their Scorecard initiatives, and identified a lead member. The four teams met with the consultant and ARL lead twice for face-to-face training in using the Scorecard. Participants came together during monthly phone calls to review progress and discuss next steps. Additional face-to-face meetings were held throughout the year in conjunction with major library conferences.

Overview of the Balanced Scorecard Process

As with many other prominent performance management models, the Balanced Scorecard process appears relatively straightforward. Participants are directed to:

1. Identify the organization's strategic objectives. Categorize these objectives into four perspectives (financial, customer, internal process, learning and growth);
2. Render these objectives to a "strategy map," a one-page representation of the organization's strategic objectives;
3. Construct metric(s) to measure progress on each objective;

4. Set ambitious but reachable targets for each metric;
5. Identify strategic initiatives to improve the chance of meeting targets;
6. Communicate Scorecard results regularly – both to staff and stakeholders;
7. Review and adjust the full complement of objectives, measures, targets and initiatives on a regular basis.

Easily stated but, as each library discovered, the Balanced Scorecard is not a simple or quick undertaking. The process demands a significant investment of time and intellectual labor. To be successful, the model also requires strong commitment from executive leadership and mid-level managers to champion the process to staff, customers and other stakeholders. And the impact on the organization can be equally significant. The Scorecard forces an organization to have new – sometimes challenging – conversations, and to analyze aspects of its current and future state that may have otherwise gone unexamined. Ultimately, the Scorecard may substantially shift an organization's strategic direction or dramatically change how its human capital and other resources are allocated. The Scorecard is, by its very nature, a change driver. And the change is relentless. The model commits the organization to continuous and regular reflection and to communicate the results of those reflections with a new level of discipline and precision.

Getting Started - Defining a Purpose Statement

Once committed to the process, the four libraries began immersing their teams in the language and key concepts associated with the Balanced Scorecard. ARL brought the participants together and facilitated the conversations. The consultant provided the training, homework, and content for learning the process.

The planning teams began by creating "purpose statements." A purpose statement defines the extent of an organization's business in one single statement. It articulates why an organization exists, the scope of its work, and the advantage it brings. The statement differentiates one organization from its peers and helps to put a fence around the more lofty and grander vision and mission statements. These purpose statements were not created for public consumption but, for some sites, proved to be useful internal tools when working on strategy.

Identifying Strategic Objectives

Prior to entering the pilot, all four libraries had strategic plans with defined mission, vision, and value statements. All four sites had concerns about the value of these plans to drive their organizations forward. All sites were maintaining formal lists of goals or objectives, but recognized that the links between these goals and their overall mission were sometimes fairly tenuous. Each site engaged in ambitious slates of projects, but the alignment between these projects and the organization's overall mission, goals and objectives was often weak. In addition, the teams discovered that their current slates of objectives were focused more on what happened last year than on what they needed to do in the coming years to achieve their missions.

The Scorecard process forced the four teams to re-examine their current slates of objectives in light of a new "balanced" four-perspective framework. Did their objectives adequately address the four perspectives or did they put too much emphasis on one or two? Did the objectives drive change or just describe and justify the current landscape? Did the objectives sync with the priorities of the larger university? What story did the current strategies tell and what story did they want them to express? How can an organization tell if it is achieving its mission when the concepts are so intangible?

Unlike in for-profit organizations, the teams discovered that their current slates of objectives tended to focus primarily on the customer (the users) and internal processes (administrative efficiencies) – with relatively little attention

being paid to the staff learning and growth and financial health perspectives. Interestingly, the changes that were happening in the overall economy in 2008 and 2009, forced a new and sharper focus on financial issues.

In some cases, existing objectives were mapped into the framework, while in other cases, new directions were required. The groups were encouraged to aim for a maximum of 15 objectives (preferably 2 or 3 per perspective), each framed using an active verb. The consultants strongly discouraged the teams from mistaking "projects" for objectives. Given that libraries do not like to stop doing anything and consistently strive to be all things to all people, narrowing down the past goals and initiatives into this smaller, more defined subset caused some angst at all sites.

Creating a "Strategy Map"

Participants were encouraged to render their slates of objectives into a "strategy map." The map is a one-page visual representation of an organization's strategic objectives. The maps were expected to very clearly show the balance and interrelationship between the four perspectives. If done well, a staff member should be able to recognize their organization's map because it accurately reflects what that organization is all about. Leaders are known to carry their strategy maps with them at all times – to help tell their organization's story to others.

Some organizations with well-developed Scorecards and access to graphic artists have devised very clever renditions of their strategy maps. (A search of Google for "strategy maps" turns up very interesting results.) Even very basic strategy maps can be extremely powerful if they effectively capture an organization's strategic future.

With the assistance of the consultants, the four ARL sites crafted very simple strategy maps

early in the process – and then returned to rework them many times during the implementation period. In some cases, discussions with stakeholders revealed the need for fairly significant overhauls. In other cases, the changes were more minute (e.g., reworking the wording to improve clarity).

On occasion, the limitations of the initial strategy maps were not revealed until later stages – when organizations were trying to identify specific measures and targets. The teams soon discovered that the choice of words was pivotal: if the word appears in an objective statement, it should be measured. For example if an objective is framed to "hire, retain, train, and develop highly motivated, productive, technologically fluent, diverse staff" then, ultimately, that organization will need to measure its hiring, retaining, training, developing, processes. In addition the organization may need to measure the motivation, productivity, tech fluency, and diversity of its staff. Eight measures could be required to fully evaluate a single overly-wordy objective. This is one of the clear focusing mechanisms of the Balanced Scorecard: it forces an organization to reconsider their lovely, lofty, lyric objectives in favor of more precise statements.

Commonalities Between Strategy Maps

Many commonalities are evident in the four libraries' slates of strategic objectives (Table 1). While the exact wording on the strategy maps may be slightly different, the intentions are strikingly similar. Overlap is evident in each perspective, but is most noticeable in the Customer and Financial perspectives.

The following analysis of objectives and measures is a snapshot of what each library recorded at the time this paper was written. Because this is a change process, objectives, measures, and even the look of the strategy maps changed fairly frequently.

Table 1
Number of Strategic Objectives by Library and Perspective

Organization	Customer	Financial	Learning & Growth	Internal Processes	Total # of Objectives
Johns Hopkins	4	3	2	2	**11**
McMaster	4	1	3	2	**10**
U. of Virginia*	3	3	3	9	**13**
U. of Washington	3	3	3	5	**13**
TOTAL - ALL LIBRARIES	**14**	**10**	**11**	**18**	**53**

* University of Virginia numbers based on 2007/9 scorecard

Common Objectives Across the Perspectives

The *Financial Perspective* provided the strongest commonalities in objectives. Of the 10 objectives, Johns Hopkins, Virginia, and Washington had 3 each, and McMaster had 1. The themes in this perspective were clear – secure funding for operational needs (4), align resources strategically (2), and measuring and improving the impact of resources and services (2).

In the *Customer Perspective* there were a total of 14 objectives across the 4 libraries, 4 each from McMaster and Johns Hopkins and 3 each from Virginia and Washington. Commonalities included the following:

- Providing productive and user centered spaces, both virtual and physical (4)
- Discovery, access, and preservation of collections for current and future scholars (3)
- Providing access to library tools and services (3)
- Becoming a world class teaching, research, and learning library (3)
- Integrating into the university's teaching and research mission (3)

The *Learning and Growth Perspective* also displayed many commonalities. Certain words appeared frequently to describe staff including "collaborative," "innovative," "dynamic," "diverse" and "healthy." Of the 11 objectives

logged, Johns Hopkins accounted for 2 while McMaster, Virginia, and Washington each had 3. Common themes are as follows:

- Developing workforces that are productive, motivated, and engaged (4)
- Developing workforces that are based on alignment with their strategic plan (2)
- Supporting diversity (2)

The unique objectives under this perspective are indicative of the local environment and organizational culture. From embedding flexibility into everyone's job description to providing clear paths and processes to carry innovation into production, clearly these libraries are reexamining the type of staffing they will need in the upcoming years.

The *Internal Processes* perspective displayed the most divergence in content. There were a total of 18 objectives – 9 at Virginia, 5 at Washington, and 2 each at Johns Hopkins and McMaster. The wide variation in sheer number of objectives is attributable to local preference: Some locations chose to position traditional internal process objectives within the customer service perspective given the focus on users. Common objectives included the following:

- Promoting the libraries resources, services, and value (3)

- Advocating for scholarly communications (2)
- Optimizing performance through efficiencies and effectiveness of programs (this is similar to objectives in the financial perspective from two of the libraries) (2)
- Integration of resources into user environments, increasing access to content, and improving search and discovery (2)

The unique objectives under this perspective include:

- Identifying unique, rare, and valuable collections
- Developing a world class preservation program
- Building collection management strategies for materials needed by scholars
- Retooling and expanding collection storage space.

Identifying Measures

Once the slates of strategic objectives were set, the four teams moved on to develop measures. As noted earlier, all four sites had been collecting vast amounts of data for many years. Two of the libraries had more robust assessment programs in place and so were more quickly able to map the measures they had to their respective objectives.

For those libraries with less advanced assessment programs, the consultant provided an exercise to facilitate measurement development. Given that the objectives themselves are often large and intangible, groups were advised to ask a very simple question – *if we want to achieve that then what do we have to do well?* Once an organization understands what it needs to do well, developing a measure is somewhat easier.

Selecting Appropriate Measures

While there are a number of considerations in choosing measures for the Scorecard, five critical ones became readily apparent to the four teams:

1. Does the metric directly measure performance to achieve the objective?
2. What data is needed for the measure?
3. How often should the data be collected and used?
4. How many measures are needed for each objective?
5. How should the measurement be presented?

Does the metric directly measure performance in achieving the objective?

Most metrics operate as surrogates or indicators of performance measurement for objectives. If the objective is narrowly written and framed using quantitative data, then it should be possible to find direct measures for it. For example, an objective to "increase the amount of gift and endowment revenue" could be linked to a measure of current revenue against a baseline. Objectives at a broader level, such as "create world-class teaching and learning spaces," would most likely use performance indicators such as user satisfaction with space or number of instructional spaces. Essentially, metrics should be able to measure or indicate an organization's progress in achieving its objectives.

What data is needed for the measure?

As noted earlier, all four libraries were already collecting vast amounts of data for purposes of campus and professional association accountability. Data that is already being collected should be reviewed first for use as measures. However, Matthews adds a cautionary note: "There is a tendency among libraries to consider only the measures that are currently being collected or that would be easy to collect" (2008, p. 67). Yet, the time and costs

involved in beginning new data collection processes can be substantial and should not be underestimated. The primary focus of time and effort should be on achieving the objective rather than coming up with the best method of measuring it. Above all, the data should be practical – obtainable with a reasonable amount of effort and easy to use and understand.

How often should the data be collected and used?

The issue of data frequency was a regular point of discussion. Much of the Balanced Scorecard literature calls for frequent reporting of results from measures, many on a quarterly basis. The frequency may depend on how readily available the data is. Data can be extracted from automated systems on demand, but survey data or other assessments may have a longer reporting cycle. Academic libraries also need to consider the academic calendar: monthly or quarterly tracking will be less useful than term-to-term or year-to-year comparisons of the same academic term periods.

How many measures are needed for each objective?

The four groups struggled to determine the correct number of measures for each objective. The preferred number depends on the objective and the data. Narrowly defined objectives generally require fewer measures than those broadly defined. For example, an organization with an objective around enhancing teaching and learning activities might consider tracking the number of instructional sessions and participants, session evaluations, number of academic programs reached, evidence in student work, survey responses, and faculty evaluations of usefulness. Data availability and frequency may also have an impact on the number of measures. Some data, such as satisfaction surveys, may be available only once every 2-3 years while other data is collected on an ongoing basis. Matthews notes that, "It is better to have fewer measures than too many" (2008, p.87). The number of measures per perspective should also be limited. Finally, the measures should be

looked at holistically within the entire Scorecard to ensure that they provide a balanced measure of overall performance and are not reliant on the same data sources.

How should the results be presented?

Choosing the best way to present the data was also a significant consideration. The four teams spent much time visualizing how specific measures would be displayed so as to capture what was most meaningful. Understanding what each particular chart type can provide can clarify what you want to show even more.

- Bar charts compare the performance of different projects at points in time
- Pie charts show the composition of a metric and are helpful for showing ratios
- Stacked bar charts show the accumulation of a measure over time
- Line charts show the performance of a metric over time

Developing a standardized slate of measures for ARL libraries

Commonalities between pilot slates

We have seen some commonalities among the pilot sites. As a result a key question is whether ARL can offer a menu of potential themes and associated metrics that libraries can target to excel in achieving their organizational mission and vision. While each library in this study is to some extent in the initial phase of measure and target development, some general observations can be made about the areas of overlap that can give insights as to whether an ARL menu of metrics can be constructed.

Table 2 identifies the measures per objective for the four pilot sites. The table indicates fairly wide variation in the number of measures – from a low of 26 measures logged by McMaster and Washington to a high of 48 measures logged by Johns Hopkins. The average number of

Table 2
Average Number of Measures Per Objective, by Library

Organization	# of Objectives	# of Measures	Average # Measure/Objective
Johns Hopkins	11	48	4.3
McMaster	10	26	2.6
University of Virginia	13	36	2.7
University of Washington	13	26	2

Table 3
Number of Measures Per Perspective and Library

Perspective	Institution	Number of Measures
Customer 47 total measures	JHU	19
	McMaster	11
	University of Virginia	9
	University of Washington	8
Financial 24 total measures	JHU	13
	McMaster	3
	University of Virginia	5
	University of Washington	3
Learning and Growth 27 total measures	JHU	7
	McMaster	6
	University of Virginia	5
	University of Washington	9
Internal 38 total measures	JHU	9
	McMaster	6
	University of Virginia	10
	University of Washington	13

measures per objective also varies significantly, from two to 4.3 measures.

An analysis of common measures across the four perspectives (Table 3) uncovers many trends. It is important to note that the libraries may have the same measure but align it with a different perspective.

In the *Customer Perspective* there were a total of 47 individual measures identified by the four libraries. Common measures include the following:

- Providing productive user-centered space (4)
- Customer satisfaction survey either home grown or LibQUAL+® (4)
- Instruction (3)
- Turnaround time of interlibrary loan (ILL) or other delivery methods (3)

- Collection, preservation, and discovery of collections, tools and services both currently and in the future (3)
- Integrate the library into the university's teaching and research (3)
- Growth and use of the Institutional Repository (2)

In the *Financial Perspective* the following were common measures:

- Generation of funds from either development activities or other sources of revenue generation (4)
- Cost and/or unit cost of e-resources (2)
- Ranking in the ARL Investment index (2)
- Library allocation as a percentage of the overall university budget (2)

Unique financial measures include amount of grant funding, unit cost of specific functions such as ILL, and measures surrounding how the library contributes to faculty research or how many journals the library holds based on citations by faculty authors.

In the *Learning and Growth* perspective the following were common measures:

- Employee satisfaction (three, with two specifically mentioning use of ClimateQUAL)
- Diversity (2)
- Completion or creation of training and development plans (2)
- Retention rates (2)
- Alignment with their strategic plans (2)

Finally, common measures in the *Internal Processes* perspective include the following:

- Promoting the library and communicating its value (3)
- Assessment of specific services (3)
- Effectiveness of their liaison services (2)
- Scholarly Communications issues (2)

- Resolution of information technology (IT) related problems (2)
- Circulation of new monographs in the past two years as a measure (2)

The unique measures in the internal perspective often deal with process improvements unique to each library.

Creating a Standardized Slate

The high level of commonality between measures being proposed at the four pilot sites suggests that a standardized slate is viable. Participants benefited greatly from sharing lists of measures. Reviewing a peer's slate sometimes suggested new areas for exploration. Discussions around measures often saved considerable time: partners benefitted from the successes and the failures at the other institutions.

The concept of reviewing the ARL statistics in light of common measurements also appears worthwhile. Such a strategy would standardize the definitions being used across member institutions and allow benchmarking between peers.

But more work needs to be done. The four pilot sites have established their preliminary set of measures and have, in some cases, started to collect data. Early experience suggests that libraries need to go through a full cycle of collecting and analyzing the data, then actually attempt to use it for discussion and decision making before determining if the framing is right. In some cases, a measure might seem to be useful until the first set of numbers appear. It is not until the collector tries to render the first set of charts or the analyst first puts it under the microscope that the true nature of the measure emerges. And sometimes the true picture is not really known until the data comes before the library's leadership group for the first time. Even the simplest measures turn out to be more complex than originally expected.

Targets

Once the measures were identified, the pilot sites began the very challenging task of setting targets. Ultimately, measures have little context without clear expectations or targets. Targets articulate the level of success needed in achieving the objective. Targets are quantitative and should be attainable. They can be based on overall mission, benchmarked practices, historical performance, and baseline data. For many measures, some form of baseline data will already be available that can be used to set targets. In other cases, a best "guesstimate" will be needed at the beginning. Targets should not be set so low that they are easily achieved without much effort. A higher target should be reachable with effort in a reasonable amount of time. Setting targets too high may lead to staff frustration and a perception by those outside the organization that the libraries are not succeeding in meeting their mission and objectives. Targets can be revised, especially if the initial effort was set without sufficient data.

The University of Virginia Library was in a unique position, having used a Balanced Scorecard approach for almost a decade. They had long used a "two target" approach (high target = full success, low target = partial success, no target = no success). Low targets were usually set at a point slightly better than current performance, while high targets were set to encourage substantially improved performance. When possible, the value for current performance was based on historical data, but occasionally targets were based on the educated guesses of responsible staff as to current performance on a certain measure. Virginia has analyzed their measures annually, noting their level of success for each measure (Self, 2004).

Organizational issues

The participating libraries faced a host of organizational issues that required considerable time and effort to address. In many cases, the issues were not well-covered in the Balanced

Scorecard literature: the literature assumes that senior leadership will wholeheartedly champion the decision to implement the Scorecard, that a senior team with authority to make decisions will oversee the process, and that staff throughout the organization will naturally understand and follow. The reality in a large academic library with a history of a more cautious approach to change, a strong emphasis on consensus and a suspicion of non-academic approaches, is often very different.

Getting the Senior Leadership Team's Attention

Not surprisingly, discussions of the Balanced Scorecard battled for attention with the immediate – the operational imperatives that suck the time out of typical leadership meetings. Teams had to convince senior leaders that strategy needed to drive operations and that the Scorecard presented a healthy mix of what the organizations have needed for some time. Scorecard team members needed to find champions within the leadership group to support what amounted to an institutional leap of faith – that the Scorecard would apply a level of discipline that would (ultimately!) simplify operational decisions, reduce waste and provide greater clarity around priorities.

All four teams reported success in engaging their leadership teams relatively early in the process. The leadership groups came to realize that the Scorecard could raise the level of discussion at executive meetings, simplify decision making, and help steer budgetary decisions. The site with two Associate University Librarians on the Scorecard team experienced the least difficulty in moving the initiative forward. This parallels the leadership involvement at the University of Virginia Library when they first adopted the Balanced Scorecard in 2001. Teams with a more distant relationship to the senior leadership group encountered more difficulty during the early days in getting the leadership's attention, securing time on leadership group agendas, and ultimately capturing interest in the Scorecard.

Overcoming Resistance – The Human Dynamic

Participating libraries found that some of their staff colleagues viewed the Balanced Scorecard with a degree of suspicion – in some cases, even cynicism. This response is not unique to libraries – with its emphasis on performance measurement, change, and accountability, employees of any organization are likely to offer resistance to the Scorecard process, and countless articles and book chapters offer strategies on how to address this reaction. One presenter at a recent conference for mission driven organizations drew laughter from the audience when he spoke about how he handled "malicious compliance" within his organization.

The tension between strategy and operations was experienced by all teams. Staff expressed concern about not seeing their specific work assignments explicitly linked to the strategy map, metrics and initiatives. The four teams conveyed a similar message back to their organizations. In some cases, the work being done by a particular unit is extremely important to supporting institutional priorities, but in itself, is not strategic at the organizational level.

In addition, the ARL initiative was underway during a particularly dire economic crisis when library budgets were stretched to the limit and layoffs were a distinct reality at many institutions. At the individual level, the Balanced Scorecard can be threatening – though not generally directly tied to job evaluations, it can be seen as a form of public performance management, and often is used to focus attention on strategic goals at the expense of ongoing, perhaps outmoded, operations. Even in flusher financial times, Virginia experienced staff reluctance to set "stretch" targets in their areas for fear of failure.

It also may be that certain elements common to the organizational culture of academic libraries may contribute additional resistance to the Balanced Scorecard.

First, libraries and library staff are not widely known as change agents. On the contrary, the collective focus has traditionally been on preservation and stability. Even key innovations have often been focused on maintaining continuity and access to historical material, albeit in new ways. Academic libraries, in particular, like the colleges and universities of which they are a part, are just beginning to substantially change their basic physical and organizational structures. Only in the last decade have libraries begun to prioritize digital over physical collections and hire programmers as they once hired bibliographers. In many ways, libraries still operate very similarly to their counterparts a century ago. Libraries are working at change, but unlike their counterparts in the fast-paced commercial world, still change relatively slowly.

Second, academic libraries and library staff are not predisposed to adopt business tools. Often mirroring the views of the academic community they serve, libraries tend to think that the scholarly nature of their work precludes the successful use of business-based strategies. However, there are signs that this attitude is beginning to change, with outcome-based budgeting and return-on-investment models gaining traction at institutions of higher education.

Finally, libraries tend to operate within silos. This too is reinforced by the larger structure of the college or university, where individual departments typically retain a great deal of autonomy. Despite the contemporary focus on collaboration and interdisciplinary scholarship, academics still tend to work on their own. Library staff are not used to coming together to talk about the organization as a whole. Staff members tend to focus on their own areas of specialization (e.g., cataloguing, reference, etc.) and do not typically create forums to facilitate high-level discussions about the future of the entire organization.

By implementing the Balanced Scorecard, planning teams are asking staff and campus stakeholders to make several leaps of faith – not only that tracking progress will increase the probability that the library will achieve its collective goals, but also that change is necessary and good, that a solution developed by and for the business world may have value in the academic environment, and that by working together, library staff can achieve more than by working independently.

Making Decisions/Authority

The Balanced Scorecard, and the strategy that underlies it, is a compendium of choices or decisions – many of them hard ones. The Strategy Map forces the organization to choose one priority or direction over another. The final choice of metrics reflects a collective decision about what truly matters and is worth counting. The specific projects or initiatives linked to the Scorecard reflect hard decisions about where the organization will invest its time and limited resources.

Some participants reported issues associated with governance and authority structures. In some cases, decision-making structures overlapped, thus getting in the way of setting clear priorities. In other cases, the decision-making structure was not clear.

More often though, the issue was behavioral rather than structural. Participants encountered hesitancy to commit to one plan over another and a reluctance to be the one to make the final decision. Staff often reported that they did not have enough information to provide an opinion on a given tactic. Groups tended to revisit the same issues over and over again and bring closure to the issues at hand.

This reluctance might be associated with the historic focus on consensus as a decision-making style within many academic libraries (and within the academy as a whole). Achieving

100% consensus on a given issue can take time and sometimes results in weaker solutions.

The four teams used a variety of techniques to arrive at decisions. All sites used a blending of staff committees to work on various aspects of the Scorecard. At the end of the day, in all four sites, final decisions were made by the senior leadership team. Some sites had success with expressing the continuous nature of the review process. The strategies were framed as hypotheses, the best choice of action at the given time and with the given information. Mid-year adjustments and regular review were part of the process.

Integrating the Balanced Scorecard into the strategic planning process

The four groups recognized that incorporating the Balanced Scorecard into their libraries' existing assessment program was relatively easy, but if that was the full extent of the integration, the implementation would be only partially successful. The Scorecard is not, as might appear at first, simply a container for assessment data. Rather, the scorecard is a management and change process first, and a metrics process second. The teams recognized that the Scorecard required a robust planning and decision-making cycle. To be effective, the senior leadership group needed to be reviewing the metrics and the strategic initiatives on a regular basis. The review meetings needed to be deep and focused on achieving success. The Scorecard becomes the catalyst for rich conversations and sometimes difficult decisions – not just another cluster of data to shelve from quarter to quarter.

Communicating Progress with Staff

Given the complexity and intense integration of the Scorecard into the organizational fabric, regular communication of progress with staff proved to be essential – but not always easy.

The strategy maps provided a good graphical representation of the libraries' strategic objectives – but the interrelationships between the objectives, the metrics and the initiatives were hard to explain. Participants struggled to find the right visual to bring all the pieces together and some pieces were undoubtedly lost in translation. The pilot leads struggled with providing the right information at the right time – without unnecessarily confusing their colleagues. In many cases, chunks of time past without noticeable progress – and the initiative moved to the back of people's consciousness. And of course, the participants themselves were learning as they went along.

The participants tried a variety of approaches to share their stories with their colleagues and their campus communities. Most teams sent out regular communiqués to library staff and held a variety of face-to-face sessions (e.g., presentations, hands-on workshops, etc.). Some sites reported the best progress when blending the balanced scorecard information with other broader events.

Conclusion

The yearlong ARL initiative has met its initial objectives. The four local implementations are still a work in progress, but the leads are fully trained and the infrastructure is in place. The sites continue to refine their measures, set their targets, and occasionally circle back to their original objective statements. Data is being collected. The leadership teams are starting to see their first deliverables out of the Balanced Scorecard process.

Although still early in the game, the concept of identifying standard suites of objectives and measures that ARL libraries can select or start from appears to hold merit. Strong commonalities are evident in the four sites' work. As well, the ARL objective to test a collaborative approach to assessment has been fruitful. The opportunities to discuss concepts and wording with peers helped reduce barriers.

The community of practice around the Scorecard process has helped make each implementation richer.

The study has identified the challenges but also the tremendous opportunities for implementing the Balanced Scorecard in an academic library. The process requires a substantial allocation of time and intellectual effort. The process requires a significant and ongoing commitment from senior leadership to be successful. The strength of the scorecard is its linkages. The process, if done effectively, can help solidify the bond between the organization's strategic objectives and the specific initiatives it elects to undertake. The Balanced Scorecard forces an organization to have new, sometimes challenging, conversations and to analyze aspects of its current and future state that may have otherwise gone unexamined. Ultimately, the Scorecard may substantially shift an organization's strategic direction or dramatically change how its human capital and other resources are allocated. The Scorecard is, by its very nature, a change driver. The model commits the organization to continuous and regular reflection and to communicate the results of those reflections with a new level of discipline and precision.

References

Association of Research Libraries. (2009). *E-News for ARL Directors*. Retrieved 30 May 2013 from http://old.arl.org/news/enews/enews-15jan09.shtml

Brophy, P. (2008). Telling the story: Qualitative approaches to measuring the performance of emerging library services. *Performance Measurement and Metrics, 9*(1), 7-17. doi: 10.1108/14678040810869387

Durgan, R. E., Hernon, P., & Nitecki, D. A. (2009). *Viewing library metrics from*

different perspectives: Inputs, outputs and outcomes. Santa Barbara, CA: Libraries Unlimited.

Franklin, B. (2011). Surviving to thriving: Advancing the institutional mission. (2011). *University of Connecticut Libraries Published Works*, paper 36. Retrieved 30 May 2013 from http://digitalcommons.uconn.edu/libr_pubs/36

Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard: measures that drive performance. *Harvard Business Review, 70* (1), 71-79.

Kaplan, R. S., & Norton, D. P. (1996). *The Balanced Scorecard: translating strategy into action.* Boston: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. (2001). *The strategy-focused organization: how Balanced Scorecard companies thrive in the new business environment.* Boston: Harvard Business School Press.

Kyrillidou, M. (2010). The ARL Library scorecard pilot: Using the Balanced Scorecard in research libraries. *RLI, 271*, 33-35. Retrieved 28 May 2013 from http://publications.arl.org/10opit.pdf

Matthews, J. R. (2008). *Scorecards for results: A guide for developing a library Balanced Scorecard.* Westport, CT: Libraries Unlimited.

Poll R., & te Boekhorst, P. (2007). *Measuring quality: Performance measurement in libraries.*.Munich: K.G. Saur.

Self, J. (2003). From values to metrics: implementation of the balanced scorecard at a university library. *Performance Measurement and Metrics, 4*(2), 57-63. doi: 10.1108/14678040310486891

Self, J. (2004). Metrics and management: applying the results of the balanced scorecard. *Performance Measurement and Metrics, 5*(3), 101-105. doi: 10.1108/14678040410570111