

The Balanced ScoreCard for Public Education

Introduction

In this paper, we'll start with a broad premise then move into progressively more narrow issues. These include the importance of the Mission for organizations, how we measure progress in accomplishing that Mission, the selection of a suite of indicators, and then a brief example of one approach: The Balanced ScoreCard.

Let's begin at the beginning with a fundamental premise: We create organizations to accomplish that which we cannot accomplish on our own. (If I had in mind to do something, and I could do it myself, I would likely find it faster and easier to just do it myself and have done with it!) When we create organizations, we do so because (1) we have in mind what we want to accomplish, some sort of Mission or Purpose however nebulously framed, and (2) we can't do it alone.

The Importance of Mission

Oddly, as we add people to an organization that Mission sometimes gets lost. Thus, in a mature organization, getting agreement on a shared Mission can pose a surprisingly difficult task. When we pose the question, What change do we want to see in the world? we may find ourselves discouraged by the variety and inconsistency of the responses.

Do we find it important that people actually agree? Yes, we do. As organizational theorist John Carver (Carver 1997) writes,

“Mission, defined as the briefest, broadest Ends statement, does not determine everything about an organization's intended results, but it tells us the range within which all further results will occur. A mission defines the arena and answers these simple questions: ‘What is this organization for? How will the world be different as a result of our being in business?’”

He defines Mission as that which we set out to accomplish, our promise to the rest of the world. Thus the rest of the world can fairly hold us accountable for the Mission and we accept that accountability. We have explicitly or implicitly negotiated that agreement. Think about this – if we have no agreement on Mission / Purpose / Ends¹, then why would the rest of the world (ROTW) give us any resources, any permissions to act, any forbearance for missteps? Why would the ROTW give us anything if it has no reasonable understanding or expectation of what it will receive in return? After all, the ROTW does give our organizations something. Frequently it gives money:

- shareholders give GE money (buying shares) in order to gain income.
- citizens give MnDOT money to keep the roads repaired.
- donors give the United Way money in order to take on social challenges.

Sometimes the ROTW gives something other than money:

- hours and / or expertise as volunteers in schools, churches, fraternal organizations, nonprofits, or coaching youth athletics.

¹ This document will use *Ends*, *Mission*, and *Purpose* (capitalized) interchangeably.

- people’s good names when they endorse causes, sign petitions, or host / attend political fund-raisers.

The preceding discussion looked at organizations in general. Let’s turn to schools in particular. Why do we have public schools? What Mission does public education have? Take Minnesota as a case in point:

**ARTICLE XIII
MISCELLANEOUS SUBJECTS**

Section 1. **UNIFORM SYSTEM OF PUBLIC SCHOOLS.** The stability of a republican form of government depending mainly upon the intelligence of the people, it is the duty of the legislature to establish a general and uniform system of public schools. The legislature shall make such provisions by taxation or otherwise as will secure a thorough and efficient system of public schools throughout the state.²

Note that the constitution calls for the funding of public schools (a Means) in the context of fostering ‘the intelligence of the people’ (the Ends). Let’s examine that: the state will fund schools (give them resources) in order to help them accomplish a worthy Mission – developing intelligent people. By virtue of accepting the funding, public schools implicitly accept this as their overall Mission. (BTW, the Minnesota constitution says only this on the topic!) Of course, individual schools may refine or narrow that Mission (i.e., we might consider how our understanding of ‘intelligent people’ varies in grade six vis-à-vis intelligent people in grades eight or twelve) but it does not seem reasonable that schools can ignore this Mission or radically alter or expand it.

For this paper, let’s leave aside *prescriptive* comments on what intelligent people ought to know, a topic on which thoughtful people can legitimately disagree. Let’s also leave aside the notion of *multiple* intelligences (visual, verbal, musical, etc.). Instead, let’s look at *descriptive* comments on what intelligent people do. For instance, let’s try this as an operational definition:

intelligent people meet challenges, capitalize on opportunities, and continue their own and their community’s development in a fashion consistent with prevailing values ... they succeed in achieving useful goals fairly.

And for the sake of argument, let’s take a broad and long view of *success* – not mere financial success and not just this week. (Without these constraints, we might permit defining illicit drug dealers as ‘successful’.) Let’s include the notion of ‘for ourselves and our children and our children’s children.’ This eliminates, say, our entirely using up some natural resource for *our* generation and leaving the account empty for the *next* generation. Let’s talk about success in terms of physical health, psychological health, political health, and also economic health.

² Text found at the website <http://www.house.leg.state.mn.us/cco/rules/mncon/mncon.htm> and accessed on 2 Dec 05.

So now we have the idea that ‘intelligent people’ succeed at what they attempt to do within the broad bounds of doing good for themselves and the community over the long term. We might put these ideas together: we fund public education in order to develop intelligent people, and we evaluate their intelligence based on our assessment of the degree of success as we have defined it: multiple dimensions for multiple generations.

Now we can move on to the next step: When and how do we measure that success? The first part of that question focuses on timing. The second part asks us to establish reasonable indicators that someone has actually succeeded.

Measuring Progress Toward Accomplishment of the Mission - Indicators

Let’s take the timing aspect first. We can only know that someone has succeeded – marketing an invention, giving a concert performance, earning \$100,000 last year, having resolved a disagreement without resorting to physical or psychological force, doing well on a test – after the fact. Only in looking backward can we truly evaluate success. In fact, using the multiple-generation criterion for success, we can only evaluate someone’s success two generations after that person dies! When Harold still walked among us, did he undertake actions that left the world in good stead for his children and grandchildren? Let’s say he did. Then we call him a success! Note this critical point: we can only measure his success after the fact. And note this important corollary: No longer can Harold do anything about it! Once Harold has died, he can do nothing to alter his degree of success. He did what he did; we cannot rewrite the record.

We thus characterize *success* as a *lagging indicator*. We can only measure it after the fact. Did we get to the doctor on time? Did we perform well at the concert? Looking backward. Must we wait until the students in our schools graduate, live their lives, and eventually die to have a sense of whether our schools did a good job? Of course not!

And in fact we don’t do that now. We monitor lots of things during the school year – at the level of the individual student, the teacher, the school, the school district. We monitor those things – test scores, daily attendance, number of discipline referrals, funding per student, student-teacher ratios – because we believe that they somehow relate to the eventual success of public education. That is, we have mental models or maps that indicate that (a) these serve as suitable proxies, (b) we can measure them now, and (c) they somehow tie to the eventual success of the schools (in developing intelligent people) and ultimately to the success of those people.

Steven Covey’s second habit for highly effective people (*The Seven Habits of Highly Effective People - Advanced Applications* 1996) admonishes us to Begin with the End in Mind:

“While chasing short-term objectives, people often lose sight of the long-term vision or the ‘big picture’. For example, people start projects, relationships, or careers without ever stopping to ask, ‘What is the long-term purpose? What meaning does this have for me?’ Habit 2 suggests that long-term vision is necessary to focus short-term objectives.”

Whatever we choose to monitor in the short term, we should inform our choice by our clear understanding of how it supports the long term and the ‘big picture’. And we do

need to monitor *something* in the short term or we will lack the ability detect whether we still remain on track.

For instance, Bryk and Schneider (Bryk and Schneider 2002) offer evidence to support their claim that the students in schools with high levels of relational trust demonstrate superior academic achievement. Thus, their mental map shows a connection between trust and student achievement. Interestingly for our purposes, we can measure the level of trust within a building right now – we don't have to wait until later. If we measure the level of trust and find it 'too low', we can take action later today to boost it up. And we would want to do this because we believe that the level of trust (measurable today) ties to the level of student achievement (only measurable after the students have taken the tests or delivered the portfolios we use to assess their achievement).

Just as success or student achievement serves as a *lagging* indicator, measurable only after the fact, trust serves as a *leading* indicator, measurable right now. Additionally, we can not only measure a leading indicator *earlier* in the process, we can also measure it *more often*. That is, we can measure it, assess the reading, take some action, then measure again to see if the action had the desired effect. We create a higher velocity feedback loop. It's only worth measuring a leading indicator, however – and we stress this observation – if we *genuinely believe* that it has a causal relationship (or at least shows a high correlation) with a critical lagging indicator tied to our Mission.

An important element of the documentation of these connections involves taking what we know or believe and making it explicit in the form of a visual map that can guide our actions. We can use this approach to close the gap between what we know and what we do – thus, if we truly believe in the validity of Bryk and Schneider's research, then how ought we to act, what ought we to do? Well, we ought to measure the level of trust in our schools, we ought to have in mind some sort of minimum standard or goal for the level of trust, and when we find a school where the level of trust falls below that level, we ought to take immediate steps to improve it. Not just because we generally favor a trusting environment but because we believe that the level of trust serves as a valid leading indicator for performance on the lagging indicator of student achievement.

Adopting Business Models?

Our experience includes work in three domains: the for-profit world of big business, the nonprofit world of small service agencies, and the public sector world of urban public education. As we move from one domain to another, we try to bring along the best of each. We note especially one feature of businesses as publicly held companies: their availability for study and the clarity of key measurements. Jim Collins spent years trying to determine why some business plod along making a profit while others reach a point where things really take off. He called the former business 'good' and the latter 'great.' When organizations in the nonprofit or public sectors – what he calls the 'social sector' – dismissed his ideas because they came from the business world, he issued a warning (Collins 2005):

“The critical distinction is not between business and social, but between great and good. We need to reject the naïve imposition of the 'language of business' on the social sectors, and instead jointly embrace a language of greatness.” (p 2)

That is, public education can not just randomly reject or adopt business *models*, it must adopt a business-like *focus* on accountability, strategy, and measurements in support of greatness ... and do this in the face of changing government mandates and political pressure.

In terms of choosing what to measure, what to use as an indicator, he goes on to write “All indicators are flawed, whether qualitative or quantitative. ... What matters is not finding the perfect indicator, but settling upon a consistent and intelligent method of assessing your output results, and then tracking your trajectory with rigor. What do you mean by great performance? Have you established a baseline? Are you improving? If not, why not? How can you improve even faster toward your audacious goals?” (p 7 – 8)

A focus on results using thoughtfully selected indicators over time provides essential feedback for evaluating performance improvement.

Collins later suggests that while profit might serve as the primary indicator in the business world, it makes no sense in the social sectors. Some readers might claim that public education bears no resemblance to business. It lacks the simple-minded focus on a single variable – profit – that characterizes business. Collins suggests instead a strong focus on a single alternative. Since people typically found social organizations to achieve some specific Mission, the achievement of that Mission must remain at the forefront. Thus,

“ ... a key link in the social sectors is brand reputation – built upon tangible results and emotional share of heart – so that potential supporters believe not only in your mission, but in your capacity to deliver on that mission.” (p 25)

Collins plans to do further research to test this claim, but offers ‘brand reputation’ now as his best guess as a unifying focal point in lieu of profit.

Furthermore, public education has become entangled in a nasty web of governmental regulations, union contracts, funding unpredictability ... Again some readers might insist that we really have to fix the whole system, not just an individual school. To that complaint, Collins responds this way:

“It might take decades to change the entire systemic context, and you might be retired or dead by the time those changes come. In the meantime, what are you going to do *now*? This is where the Stockdale Paradox³ comes into play: You must retain faith that you can prevail to greatness in the end, while retaining the discipline to confront the brutal facts of your current reality. What can you do *today* to create a pocket of greatness, despite the brutal facts of your environment?” (p 30)

To recap: Collins points out that public sector organizations must

- adopt a focus on greatness (letting go of ‘good enough is good enough’),
- identify a baseline for consistent and intelligent assessment of outputs, and
- work toward creating a pocket of greatness wherever you find yourself now.

³ Named after Admiral James Stockdale, a survivor of seven years in a Vietcong POW camp. He held on to two contradictory beliefs: his life couldn’t be worse at the moment, and his life would someday be better than ever. (Collins 2001)

The Balanced ScoreCard – Choosing Indicators

In the spirit of bringing not just business *models*, but rather a business-like *focus* across to the public sector, let's turn to the Balanced ScoreCard introduced by Kaplan and Norton in their book, *The Strategy Focused Organization* (Kaplan and Norton 2001).

They begin by suggesting several principles, including these two:

“*Principle 1: Translate the Strategy to Operational Terms ...*

Principle 2: Align the Organization to the Strategy ...” (p 10-11)

Again, notice the fundamental emphasis on developing a strategy for success and using it to bring coherence to the organization.

At its most elemental, the Balanced ScoreCard (BSC) stresses that the strategy must include attention to both lagging *and* leading indicators. That is, it must address how we will measure success (a lagging indicator) and what we can measure now which we believe will bring about that eventual success (leading indicators). They suggest that effective BSCs include four distinct layers: Financial, Customer, Internal Business Processes, and Learning and Growth. That is, a for-profit organization has the ultimate purpose to generate profit. Hence we measure success financially. If we shift our focus from a for-profit organization to a social sector organization, we need to return to the ultimate purpose of those organizations – to achieve their Missions. Thus we substitute *Purposefulness* for *Financial* as ultimate indicator: What do we claim as our Mission? What level of success do we show in achieving it? Do people believe in our Mission and our ability to deliver on it?

Kaplan and Norton suggest that the BSC must include a prudent balance between leading and lagging indicators. Further, they suggest that

“Measuring organizational performance requires such causal chains of value creation. Stand-alone measures cannot capture the means by which improvements in intangible assets and internal processes lead to increased performance in outcome measures. The linkages in strategy maps provide the recipes for such transformations and value creation.” (p 100)

That is, not only must the BSC show balance, it must also make the interdependencies among the indicators clear for all to appreciate. Returning to the example of Trust noted earlier, the BSC must draw a clear linkage between the (current, leading) Level of Trust and (eventual, lagging) Student Achievement.

“The BSC strategic management system works best when used to communicate vision and strategy, not to control the actions of subordinates. This use is paradoxical to those who think that measurement is a control tool, not a communication tool. Excellent leaders recognize that the biggest challenge they face in implementing change and new strategies is getting alignment throughout the organization.” (p 352)

Fundamentally we use the BSC to communicate what an organization proposes to measure as it attempts to achieve its overarching goals. The focus does not rest on the Means (e.g., programs, projects, initiatives, activities) but on the Ends, the achievement of the intended outputs.

Bossidy and Charan strongly echo this message in their book, *Execution* (Bossidy and Charan 2002).

“To understand execution, you have to keep three key points in mind:

- Execution is a discipline, and integral to strategy.

- Execution is the major job of the business leader.
- Execution must be a core element of an organization's culture." (p 21)

The culture of an effective organization emphasizes greatness and the organization has a clear strategy to achieve it. They believe in clarity and simplicity in the communications:

"Along with having clear goals, you should strive for simplicity in general. One thing you'll notice about leaders who execute is that they speak simply and directly. They talk plainly and forthrightly about what's on their minds. They know how to simplify things so that others can understand them, evaluate them, and act on them, so that what they say becomes common sense." (p 70)

Common sense serves as the key idea here – *common* meaning broadly shared, and *sense* meaning that people understand it.

To recap Kaplan and Norton and combine their thoughts with Bossidy and Charan: Effective organizations create strategies to help them achieve greatness. Those organizations communicate those strategies broadly throughout the organization. They have in mind a sense of their goals and how improving their performance on *this* indicator will eventually result in superior performance on *that* indicator. And those organizations stress simplicity.

Layers in the Suite of Indicators

Actually, indicators do not typically fall into two neat categories called *leading* and *lagging*. Rather we note that certain indicators show more of a leading orientation while others show more of a lagging orientation. Kaplan and Norton put forth these four layers:

"The BSC provided a framework to look at the strategy used for value creation from four different perspectives:

1. *Financial*. The strategy for growth, profitability, and risk viewed from the perspective of the shareholder.
2. *Customer*. The strategy for creating value and differentiation from the perspective of the customer.
3. *Internal business processes*. The strategic priorities for various business processes, which create customer and shareholder satisfaction.
4. *Learning and growth*. The priorities to create a climate that supports organizational change, innovation, and growth." (p 23)

Financial indicators show the most lagging orientation, while *Customer* indicators seem slightly less so. *Internal business processes* serve as leading indicators while *Learning and growth* seem even more so. If we carry on with the practice of substituting *Purposefulness* for *Financial*, then that would serve as the social sector's most lagging indicator. We further suggest recognizing that social sector organizations do not have a nice clean relationship to one dominant stakeholder called a *Customer*. Thus we also suggest substituting *Stakeholders* for *Customer*.

We want to make two additional comments to help explain the choice of indicators in the four layers identified above. Firstly, regarding the layers or categories themselves: we do not choose them arbitrarily; rather they call attention to what we need to measure in four dimensions. We expand on that by including two words for each layer to clarify that intent. Thus we talk about the categories this way:

- Purposefulness Perspective – Delivering Results
What can we measure to show that we have achieved our Mission?
- Stakeholder Perspective – Developing Relationships
Whom do we need to focus on satisfying and what constitutes evidence of their satisfaction?
- Internal Process Perspective – Developing Routines
What must we do well, day-in and day-out, with predictable high quality?
- Learning and Growth Perspective – Developing Potential
What skills or competencies must we focus on developing today in order to increase our probability of success later?

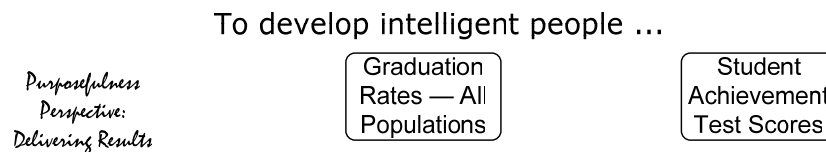
The ultimate category emphasizes delivery of *results* while the others emphasize legitimate developmental elements along the way.

Secondly, we want to note that we view the Purposefulness Perspective as essentially an outward-looking set of indicators, those that the public at large ought to show interest in. On the other hand, we view the Learning and Growth Perspective as essentially an inward-looking set of indicators, those that the leadership of the organization must attend to today.

A Brief Example

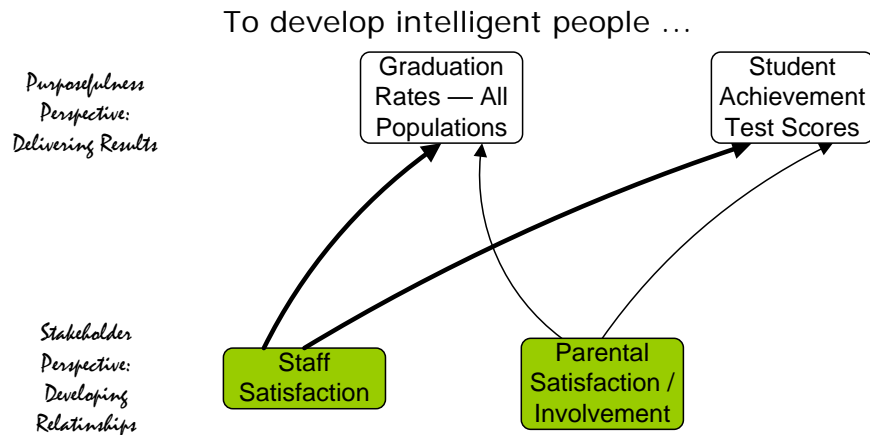
Let's turn at last to an example. We'll use a hypothetical high school that has adopted as its Mission 'the development of intelligent people.' As a result of adopting this Mission, the staff at the school has settled on two lagging indicators at the *Purposefulness* layer. They talked it over and decided that if the students actually stayed in school and graduated, that would indicate their having more intelligence (and a greater likelihood of succeeding in their lives after high school). Such factors as the commitment to staying in school, an appreciation of the long term benefits of a better education, etc. influenced this decision. They thus chose *Graduation Rates – All Populations* as one key lagging indicator. They also agreed that simply attending classes reliably and graduating did not by itself indicate intelligence. That is, not only attendance but also demonstration of learning mattered in assessing intelligence. Since the state already required the use of a variety of standardized tests, they decided to use the results of those tests as a convenient measure of *Student Achievement*. (They also elected to supplement the basic test results, but that falls beyond the scope of this paper!)

When they drew this on paper to create the visual map of how things work, they started like this:



As they continued the conversation they concluded that at the next layer down, *Stakeholders*, they also needed two indicators. They reasoned that *Staff Satisfaction* critically affected both the *Graduation Rate* and the *Student Achievement Test Scores*

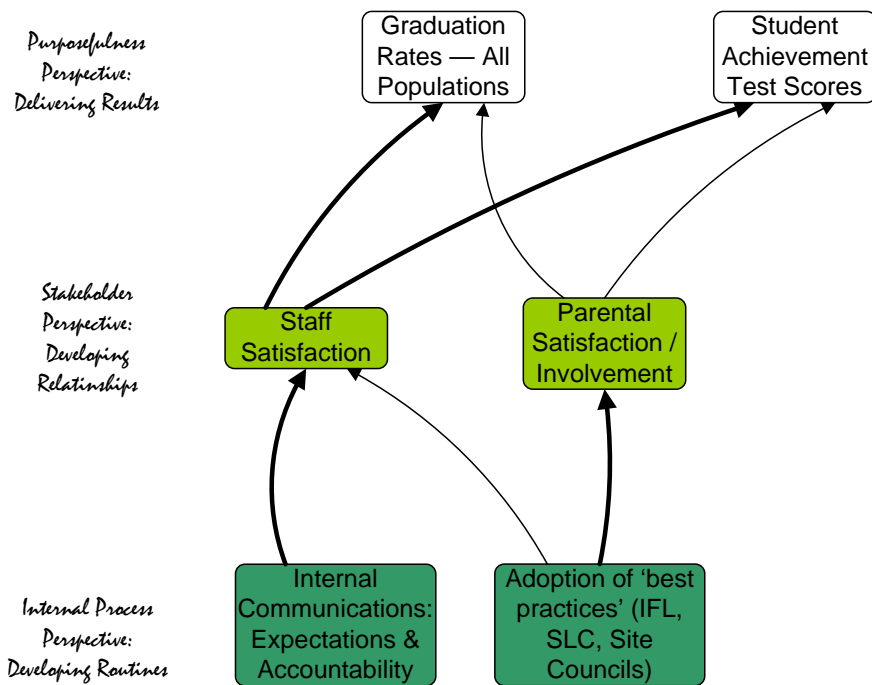
due to the staff's ability to make the learning environment attractive and challenging. They also decided to include *Parental Satisfaction / Involvement* as crucial, both in terms of volunteers in the building supporting the teachers and students and also in terms of community support for referenda to fully-fund essential school activities. Their next iteration of their drawing looked like this:



Note that they used a darker arrow to indicate a stronger interdependency. Thus they saw Staff Satisfaction as a stronger cause of eventual success.

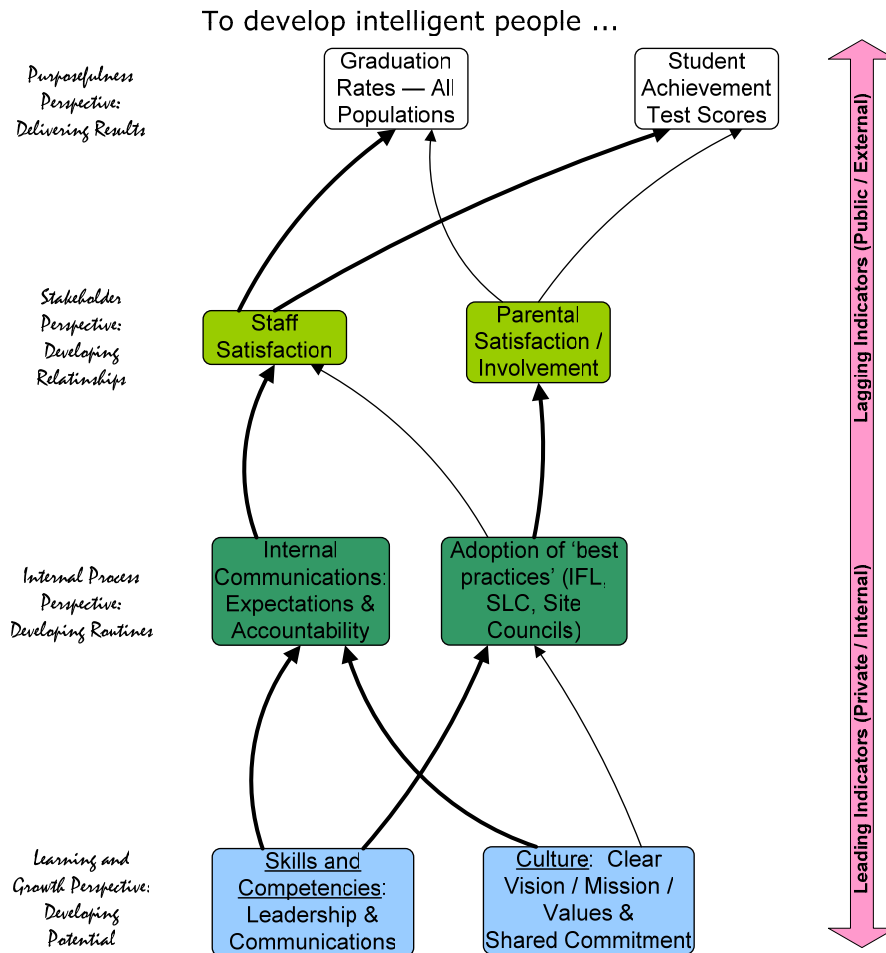
They continued their conversation shifting their focus to the *Internal Processes* layer. In asking themselves what led to staff satisfaction, they noted the importance of the staff feeling fully involved in setting the school's directions and policies. Thus, when the administration did not make any major decisions leaving the staff out of the discussion, the staff shared clarity about accountability for various actions and results. In addition, they noted that their reputations in the community as professionals depended on the extent to which they adopted and incorporated into their day-to-day practices research-based conclusions about what worked best. That led them to add two indicators at this layer:

To develop intelligent people ...



Note that they tied *Internal Communications* strongly to *Staff Satisfaction* and *Best Practices* strongly to *Parental Satisfaction*. They saw the connection between *Best Practices* and *Staff Satisfaction* as weaker but still worth noting.

Finally, the staff shifted its concentration to the bottom-most layer. What must they learn and how must they grow in order to generate better internal communications and adoption of best practices? They decided that an effective organization needs strong and effective leadership across the board not just from those in top positions. After much soul-searching, they concluded that they must learn to communicate better under all circumstances – emergencies and ordinary situations, high- and low-stress environments, internally and externally, up and down the hierarchy, across stakeholder groups. They also decided they needed to strengthen their culture of greatness and renew their commitment to the school’s values and mission. With this in mind, they amplified their model to look like this:



Again, note the varying strengths they found for the interdependencies. Further, note that they added as a reminder the vertical arrow on the right. While the public will generally show the most interest in the *lagging* indicators near the top, the *leaders* within the organization must also stay focused on the *leading* indicators near the bottom. That is, the indicators near the top have greater external appeal because of their immediate relationship to the Ends; the indicators near the bottom function more like internally-oriented Means.

Warnings!

The example we've just walked through remains incomplete. Even to get to this incomplete state, it took a significant amount of time and effort. It reflects the thinking of those involved; the leaders must fully explore this with the rest of the staff in order to gain broad understanding and acceptance. A fully-drawn BSC ought to include two-to-four indicators on each of the four layers, for a total of twelve-to-fifteen overall indicators. Each of the indicators requires an 'operational definition' clear enough to support effective use of the BSC.

Note also that another school, even another high school in the same city, cannot simply adopt this BSC as its own. Just because "they're in the same line of work" does not

mean that the circumstances and priorities will carry over. The process of developing and agreeing to a model forms part of the dialog that characterizes well-led and effective organizations. As Bossidy and Charan note:

“Only the leader can set the tone of the dialogue in the organization. Dialogue is the core of culture and the basic unit of work. How people talk to each other absolutely determines how well the organization will function.” (p 25)

Simply adopting someone else’s model will not gain the sort of examination of critical topics and ultimate commitment that result from serious dialogue.

Of course, now that an organization knows *what* indicators to measure, it must determine *how* to measure them. That brings us to the next topic, the Principal’s Dashboard.

Conclusion

In this paper, we’ve tried to establish several ideas:

1. The organization’s Mission really does matter
2. Commitment to achieving the Mission implies measuring progress
3. Thoughtful leaders can adapt ideas and concepts from the business domain to the public education domain
4. Choosing a useful suite of measurements requires attention to both lagging *and* leading indicators
5. Creating a visual representation of the interdependencies helps with the communication process
6. Using dialogue during the development process increases the likelihood of achieving a truly shared understanding of how we see the connections between these things.

We hope you find the ideas and the example presented here useful and challenging. You can find a host of resources related to Balanced ScoreCards and their use in various organizations via the Internet.

If you’d like our support in bringing these ideas into your organization, please contact us:

The Commonwealth Practice, Ltd.
info@TheCommonwealthPractice.com
Michael Ayers 612.308.0501 (mobile) mbayers@earthlink.net
Mike Marois 651.398.4695 (mobile) mjmarois@comcast.net

References

- Bossidy, Larry, and Ram Charan. 2002. *Execution: The Discipline of Getting Things Done*. New York: Crown Business.
- Bryk, Anthony S., and Barbara Schneider. 2002. *Trust in Schools*. New York: Russell Sage Foundation.
- Carver, John. 1997. *Boards That Make a Difference*. San Francisco: Jossey-Bass.
- Collins, James. 2001. Level 5 Leadership. *Harvard Business Review*, January 2001.
- Collins, Jim. 2005. *Good to Great and the Social Sectors*. Boulder CO: Jim Collins.
- Kaplan, Robert S., and David P. Norton. 2001. *The Strategy-Focused Organization*. Boston: Harvard Business School Press.
- The Seven Habits of Highly Effective People - Advanced Applications*. 1996. Covey Leadership Center.