

USING IMPERFECT METRICS WELL: TRACKING PROGRESS AND DRIVING CHANGE

Jerry L. Tally and Eugene H. Fram

The strategic plan is done. The objectives are clear. The time frame is set. The Board has done its work . . . until someone utters the word *metrics*. “How are we going to measure the outcomes?” comes the call. “Don’t we have to evaluate our Executive and our organization somehow?” And suddenly the “work of the Board” seems to once again blossom anew.

Strategic plans are most vulnerable not in their development, but in their implementation. And implementation often hinges on some measurable indication of progress. Without those metrics, the plan is a group of intentions always on the verge of greatness. Without hard data on which to anchor organizational outcomes, the organization can wobble off course without a clear warning signal.

But measurement is a daunting field. Decades of work in the sciences, engineering, theory building, and psychological testing have generated rules and models that require statistical sophistication and research to implement. Except for large national nonprofit groups, most nonprofit budgets just cannot afford such luxuries when scarce resources are needed to deliver services. Yet governmental agencies, accrediting bodies, foundations, and individual donors want some (even

imperfect) attempts at assessments of outcomes. It is better to try to assess outcomes than to approach these organizations empty-handed.

Not being able to afford the time and money to develop excellent metrics, nonprofits often have to glean whatever value they can from using imperfect metrics. To be more precise about the term *imperfect*, we mean metrics that are anecdotal, subjective, interpretive, or qualitative. Or perhaps the metric relies on a small sample, uncontrolled situational factors, or cannot be precisely replicated. For most nonprofits, it is nevertheless a great leap forward from doing nothing to using even seriously flawed but reasonably relevant measures for their critical goals. Aside from technical requirements, the most critical requirement is that both the board evaluator and the operating manager agree that the process is reasonable and that the outcomes from it constitute fair and trustworthy information. With that goal in mind, we can explore how to use an imperfect metric well.

What Should Be Measured?

We see metrics at a fragile point conceptually. They are partly defined by the strategic objectives of the non-

Strategic plans are most vulnerable not in their development, but in their implementation.

profit organization; that is how you decide *what* to measure. It would be easier to measure organizational activities, but the nonprofit board's proper focus is on outcomes, not organizational efforts. The frequent temptation, however, is to look into the operational level of the organization, where potential metrics abound. It is the tension between "what we *should* measure" and "what we *can* measure."

But all the really important things seem almost impossible to measure. An organization can create reasonable indicators of finances, membership, clients served, attendance, and other operational measures. But how does it measure actual results in the world outside, such as enhanced quality of life, elevated artistic sensitivity, community commitment, successful advocacy, or any of the other honorable but inherently vague goals that not-for-profits frequently adopt?

Metrics are equally constrained by the technical requirements of good measurement. There are standards, we are told, for a "good measure." So there is a substantial pressure to develop more precise metrics, regardless of whether they are strategic or operational in focus.

If the nonprofit pushes for technically correct metrics, it often means months of tedious board debate and volunteer time, but it has been our observation that an organization frequently ends up with good measures of peripheral events, such as changes in attendance at the annual dinner. Decades of blindly implementing

"Management by Objectives" have made such practices routine. As a result, nonprofits are left with precious little that tracks the relevant outcomes defined in the strategic plan. They focus on what they *can* measure instead of what they *should* measure.

Nonprofits need not choose between having "no measures" or the high cost of developing perfect measures. The better answer is to learn how to use imperfect but relevant metrics well.

But how could an imperfect metric be useful? Wouldn't it contaminate the whole process? An example may help to clarify the benefits.

As part of its accreditation process, the prestigious American Assembly of Collegiate Schools of Business allows accredited schools to utilize local business executives to conduct mock interviews for assessing graduating students' communications and presentation skills. The purpose is to obtain the executives' estimates of the skill levels the students have acquired during their undergraduate years. The insights garnered from these sessions can be interpretative, subjective, and anecdotal, and based on the experiences of the evaluator. Consequently, their comments reflect the viewpoint of each interviewer as much as the actual achievement of the interviewees. In short, it is a very imperfect measure.

Nonetheless, the process can allow the schools to:

*The tension is between
"what we should measure"
and "what we can
measure."*

“Wouldn’t an imperfect metric contaminate the whole process?”

- Obtain outside perspectives of the communications learning that the students have acquired and better understand minimum business expectations.
- Improve communications between the faculty and the business community.

- Indicate to students that academic content has practical values beyond helping to pass tests.
- Provide insights for curriculum change and for faculty research.

In short, imperfect metrics used well can have positive benefits! The three cases in the sidebar provide more examples of how metrics that fail the rigorous standards of scientific measurement might still serve the needs of a nonprofit organization. (See the sidebar, “How Imperfect Metrics Can Provide Positive Outcomes.”)

Building Better Metrics

We have argued that an organization need not be excessively academic or sophisticated in building outcome metrics. That does not imply that “anything goes.” In fact, we would argue that some steps are absolutely es-

HOW IMPERFECT METRICS CAN PROVIDE POSITIVE OUTCOMES

Case #1

Families Primary is a nonprofit counseling service offering a range of services to improve mental health in the metropolitan community in which it is located. Services range from individual counseling to being legal conservators for elderly clients. The mission of the organization is to reduce mental health problems in the community. Local county health officials have noted a significant increase in inner-city mental health problems. However, the use of the agency’s services by these residents was very modest. The costs of conducting a reasonably comprehensive client attitude and mental health needs assessment study would be too high. Yet not measuring these key outcomes leaves the agency vulnerable to any number of criticisms.

The president/CEO was evaluated by a board assessment committee. Committee members took primary responsibility for establishing board-approved goals and evaluating specific outcomes (for examples, finance, personnel, and fund development) of the operation. The person assigned to client development was asked to create outcome measures for improving the perception of Family Primary among inner-city residents. He and the CEO concluded that a board member needed to interview the executive directors of the five inner-city community centers to obtain a macro-assessment of the agency’s image. Cost constraints prohibited developing more precise outcome metrics.

All five executive directors unanimously reported the local residents were “uncomfortable” with the agency’s staff.

Based on the interviews’ outcomes, the board member and the CEO agreed that in 12 months, the board member would revisit the five executive directors to assess changes in perceptions, based on corrective actions to be instituted by the CEO. It was the responsibility of the CEO to devise the corrective actions that were needed to drive change. Subsequently, a quantitative goal would be set for recruitment of inner-city clients.

The next year, the executive directors reported some improvement in perceptions. However, it took a second year of corrective actions before the board member and the CEO agreed it was time to establish quantitative outcomes to evaluate performance. Incremental achievement had taken place driven by an *imperfect* outcome measurement.

The metrics were admittedly qualitative, subjective, and vulnerable to unpredictable distortions; in short, they were imperfect. But they provided a focus on relevant issues. They linked players into productive conversations with each other about where to spend resources and when to revisit their progress. There was a process robust enough to benefit from poor metrics.

Case #2

A preservation advocacy group in a major city was hopeful that they could mobilize the preservation community to take action on specific projects. Typically, the organization would "take a stand" when someone threatened to tear down or "modernize" a historical treasure. Rather than just be one more voice in the political melee, their strategic objective was to direct and energize the larger cluster of organizations with interests in preservation, everyone from local chapters of national preservation organizations down to neighborhood associations and even contractors. But how do you measure political influence? How do you measure whether the organization motivated or directed another organization?

The group settled on two or three imperfect metrics that nonetheless provided a major step forward. The first was to track how many of the board members had a spouse, friend, or business partner on the boards or staffs of other preservation organizations. Mere membership clearly does not guarantee influence (hence, it's a poor metric), but having no connections to other organizations probably does guarantee the absence of influence. And merely collecting the data drew attention to how well board members were building bridges to other organizations. It also highlighted the organizations for which they had no connections, and motivated board members to explore new possibilities.

The second metric was only slightly better. The group decided to count the number of other organizations that were willing to "stand with" them on any particular project. Deciding whether another organization was "standing with" them rather than just "standing next" to them was obviously debatable. Another organization may simply have come to the same conclusion rather than deciding to join forces. But that debate was exactly the question that needed to be highlighted. Deciding whether another organization could be considered as a deliberate ally vs. an accidental one drew people into the very issues they wished to understand.

Case #3

A local Jewish community center had a strategic objective to become the "nexus of Jewish life" in the region. While board members felt the objective was absolutely central, they had a difficult time giving it a precise, *measurable* meaning. It had a slightly different nuance for each board member. And the openness of the concept was part of its appeal. They did not want to nail it down precisely; it was important that it be allowed to evoke different images in the staff and membership.

The first measure used was whether program participants had found a new friend through their involvement with the center. This is clearly a dubious metric, but it did focus attention on whether the center was merely a service provider (which people just attended) or whether it was actually a catalyst for community members (who might talk to each other enough to find a new friend).

The second measure was whether people felt more connected to the Jewish community as a result of their involvement in the center. It is clearly a very subjective measure, and it was reliant on all the vagaries of self-selection in a survey tool. But watching it trend up or down gave the staff a good reason to re-examine how they were engaging community members and meeting their needs.

Building the relationship for assessment is as important as the assessment tool itself.

ential, even if the resulting metric fails to achieve high standards of rigor.

We have also argued that building the relationship for assessment is as important as the assessment tool itself. For that reason—not surprisingly—our suggestions for how to build a better metric addresses two different challenges: the technical challenge and the relationship challenge.

The Technical Challenge

Each of the examples in the sidebar followed a five-step process:

1. *Agree on relevant outcomes:* Metrics should be used to reflect organizational outcomes or impact, not activities or efforts. For example, *one general outcome* could be “To enable the student to refine and evaluate his/her occupational goal.” Or it might be “Build community support for preservation initiatives.” Or “Be the catalyst for stronger community ties.” Outcomes should focus on a desired change in the nonprofit’s universe rather than a set of process activities.
2. *Agree on measurement approaches:* There are many possibilities for measurement. These include personal interviews, mail questionnaires, sampling data in client records, Internet surveys, comparisons with other agencies, peer or outside consultant visits, and comparing the organization’s imperfect data with

similar types of national data. Boards often prefer methods that are more quantitative because they are easier to manage; often the richest data outcomes can be developed from insights generated by more qualitative methods, based on small samples.

3. *Agree on specific indicators:* Develop behavioral outcomes desired. For example, one of a number of specific goals might be “Some students find, as a result of cooperative experiences, that they made poor occupational choices.” Or, mentions in the local newspaper can be used as an indicator of public presence. There will often be a temptation to add in other indicators simply because they are available, or because they “would be interesting to look at.” Keep the focus on the indicators of agreed-upon outcomes!
4. *Agree on judgment rules:* Board and management need to agree *at the outset* upon the outcome metric numbers the organization would like to achieve for each specific indicator that contributes to the desired strategic objective. The rules can also specify values that are “too high” as well as “too low.”
5. *Compare measurement outcome with judgment rules:* Determine how many of the specific objectives have been achieved to assess whether or not the strategic objective has been achieved.

The Relationship Challenge

Meeting the technical challenge described above will not get the organization very far unless the board and executive develop a good working relationship in implementing the metrics that have been agreed upon. In our experience, three steps are key to achieving a positive working relationship:

1. *Link an imperfect metric to a good process.* Even an excellent metric will be caustic if applied by people who have little communication or trust in each other. How the metric is used to track progress and drive change will be as important as how it was defined. Trust will prove as important as the technical requirements of measurement.

In the evaluation process, it is particularly important that the board chairperson and the senior management executive trust each other. The chairperson must view the senior manager as a competent executive, not an expert in direct service who needs help with management activities. Whether or not the top executive came up the direct service route is unimportant. As the top executive, his or her first job is to manage. The Board holds that person accountable to do the job, using precise and/or imperfect metrics. In return, the members of the board must distance themselves from operations and let him or her do the job. The two then can be ready to evaluate outcomes in a fair and constructive manner.

2. *Let experience drive improvement of the metric.* It would be easy to debate for months over the subtleties of measurements. What data will be collected? How will the data be collected? How will it be displayed? Who should see it? What should we do if it is too high or too low? While these questions will need to be answered eventually, it is better to start using empirical feedback with a developmental attitude than to insist on a complete design before you collect your first sample.
3. *Attend as much to the developing relationships as to the technical act of measurement.* Over time, the metric will likely improve. But it is equally important that the relationship among those measuring and those being measured also improve. The purpose of gathering data and reviewing it is to provoke and inform explorations of how to shift operations of the organization. If the relationships of those involved will not support change and creativity, then even the most precise metric will be of little value.

A poor metric can be modified with experience.

The process for using the metrics needs to engender trust and rapport.

While Boards should not reward mere effort without results, they *should* be respectful of good intentions and honest work.

Conclusion

Without some way of measuring their impact on the community, nonprofit boards can easily degenerate into monitoring staff activities, mistaking efforts for outcomes. That danger is much greater than the danger of using imperfect metrics. A poor metric can be modified with experience, but a board that begins to meddle in tangential routine operational affairs does not necessarily learn from its mistakes.

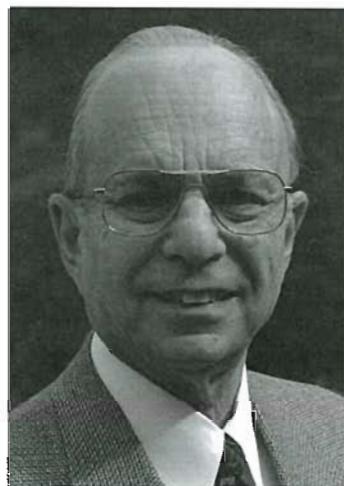
Using metrics, no matter how sophisticated, should never be divorced from the working relationship of the players. Their level of shared understanding, interpersonal trust, and willingness to be vulnerable are as important as the measurement tools they employ.

Organizationally, the senior manager (executive director, CEO, president, etc.) plays a key role in this approach to measurement. He or she can take the initiative in calling for installation of some metrics (no matter how imperfect) rather than tolerating no metrics at all. They can also push for metrics that target relevant outcomes rather than mere activities. And, lastly, senior managers can leverage their more intimate involvement with operations to suggest more subtle measures about which board members may be unaware.

Board members also play an equally important role by redoubling the emphasis on outcomes rather than agency operations. Perhaps more important, they can exercise prudence and good sense in using imperfect measures, especially when they reflect on the compensation or even dismissal of the agency executive director. The board needs to carefully consider the appropriate confidentiality around the measurements it collects, and to ensure that all board members help to support those agreements. And, since it is the party with the power, the board must be especially diligent in building the right relationships with the CEO during the assessment process.



Jerry L. Talley is a veteran of three careers. For 18 years, he taught on the Stanford faculty in the Department of Sociology. During that time, he also held a private practice as a licensed marriage and family therapist. And for the last 25 years he has consulted to companies in virtually every segment of the economy. He has a particular focus on nonprofit governance issues. His current practice (www.JLTalley.com) is built around an advanced problem-solving model, a new perspective on decision making, and helping nonprofit boards learn new models of leadership. He also serves on the board of the Family & Children Services of Santa Clara County.



Eugene H. Fram is emeritus professor in the E. Philip Saunders College of Business at the Rochester Institute of Technology. For decades, he has been a nonprofit author, consultant, and volunteer board member. "Policy vs. Paper Clips—Selling the Corporate Model to Your Nonprofit Board" (2nd edition, 1995) is his widely used governance book, which focuses on building internal board-staff trust relationships and on more clearly separating managerial duties from board policy and strategy duties. He is frequently contacted by the media for his views on marketing, corporate governance, and nonprofit management, usually logging about 100 national and local media placements a year.

