EFCOG Guidance Document: Development and Use of Leading Indicators

Table of Contents

Overview

This document provides general guidance in developing and using leading indicators for the Energy Facility Contractors Group (EFCOG). The initiative originated in March 2010 and has evolved into collaboration between Los Alamos National Laboratory (LANL), Sandia National Laboratory (SNL), Lawrence Livermore National Laboratory (LLNL), Savannah River Site (SRS), Hanford Site, Pantex Plant, and Nevada Site Office. The wider EFCOG community has also provided feedback, leading indicator examples, and other guidance to improve the document.

The purpose of this guidance is to describe an approach for developing leading indicators that is broadly applicable across sites and helps managers to manage their operations. As such, it provides a sampling of metric development techniques. Note this is intended to be a living document. Appendices 1 and 2 provide templates for EFCOG facilities to develop leading indicators and to submit examples of leading indicators an EFCOG library.

Leading Indicator Discussion

Leading indicators point to specific outcomes. We want a safer car, a more reliable appliance, a speedier service, a lighter sleeping bag, or a more effective medication, and so we identify and measure factors that we believe are likely to influence or affect those outcomes. In other words, a search for leading indicators is a search for "knobs that we can turn—and real-time feedback for the knobs we do turn."

Figure 1: A Familiar Data Source

One outcome we have all monitored is body weight. Over time we have built up a history of bathroom scale measurements that, while probably not written down, serves as a context for our current measurements. We are not surprised at our weight in January because of measurements we took back in October and November. And of course we factor in all the food eaten over the holidays!

A universally acknowledged leading metric for body weight is caloric intake. As long as we are able to control the daily caloric intake and keep it below a certain target value, then we can achieve our personal target weight.

The weight loss project is clear: limit daily caloric intake (*monitoring* and *acting on* the leading indicator) in order to influence the number of pounds registered on the bathroom scale (*monitoring* the lagging indicator). Over time we are likely to observe a cause-and-effect relationship between the leading and lagging indicator.

Viewed from another perspective, however, caloric intake could itself be seen as a lagging indicator. The following is a partial list of factors that influence or mitigate caloric intake:

- Daily consumption of fast food
- Daily exercise level
- Number of hours worked per week
- Frequency of family stress events (e.g., deaths, births, job changes)

These examples point to a few peculiarities about leading indicators: they may serve as lagging indicators in other metric structures or contexts. And they must be actionable in order to be effective. It does nobody any good to record measurements without taking further action.

To sum up, here are a few basic principles of leading indicators:

- Predictive of and able to influence future performance
- May themselves be lagging indicators in other contexts
- Have a high probability of effecting a particular outcome: "knobs" that we can turn given an organization's existing dynamics
- Can't exist in isolation – decision makers need to use them in order to influence an outcome
- Often attached to process elements
- Leading indicators need only be developed for measures that truly matter

Process Introduction

The process of leading indicator development and implementation described in this EFCOG guidance document consists of five primary steps: (1) setting the stage, (2) selecting indicators, (3) conducting a qualitative review, (4) conducting a quantitative review, and (5) using and refining the indicators. Figure 2 shows the process flow. Each of these primary steps will be described in turn. A real-world example, taken from a recent LANL initiative in developing leading indicators for research and development (R&D) work, will be used to illustrate the steps.

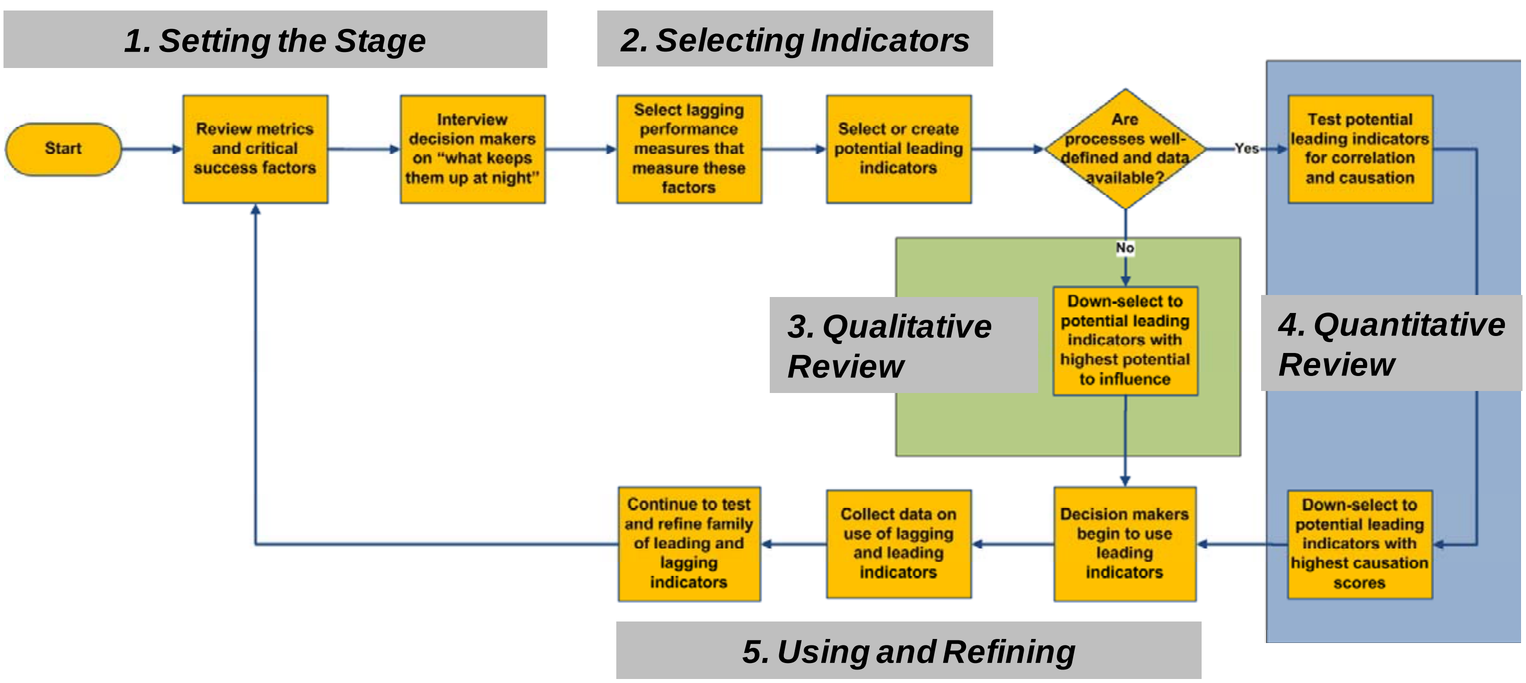


Figure 2: Process Flow for Developing Leading Indicators

Note the process described in this document is not prescriptive. Instead, it is designed to accommodate different site methodologies and conditions for developing leading indicators. For example, sites that employ the lean six sigma method may want to focus on the established DMAIC (Define, Measure, Analyze, Improve, and Control) process for developing indicators. (See Appendix 3, Lean Six Sigma Overview, for further information.) This guidance document should be viewed as a supplement to ensure that best practices, methods, useful hints, and local customization are not overlooked in the indicator development process.

Setting the Stage

The process of developing leading indicators starts with two complementary steps: reviewing the site hierarchy of measures and interviewing decision makers on "what keeps them awake at night." The first step is a logic check; the second is a gut check. Together, these two steps provide a basis for the search for leading indicators and help to ensure that they are relevant for the organization. Figures 3 and 4 illustrate the steps.

Reviewing the Metric Hierarchy

The choice of indicators to monitor can initially seem daunting. A good starting point is the existing site measurement framework, which may take many forms: balanced scorecards, dashboards, enterprise models, status panels, or strategy maps. Whatever the framework, developers should place the search for leading indicators within the existing measure structure.

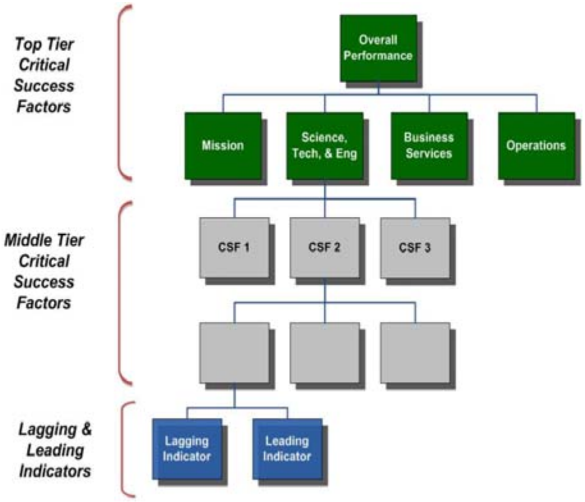


Figure 3: Critical Success Factors

The following questions may be asked: What are the key success factors for the organization? Has the organization's mission and vision been incorporated into the framework? How are these outcomes represented? Have associated goals, key deliverables, and multiyear strategy been incorporated? Have factors for customer and employee satisfaction and loyalty also been considered? How well do the identified success factors translate into the family of metrics? Are risk factors (those elements that could harm the key success factors) represented? Does each tier contain what is necessary and sufficient for the tier above? Where do any potential leading indicators fit into this family of metrics?

Figure 4: The Gut Check

The reasons for this exercise are as follows: it is essential to identify what management considers important. Fitting the search for indicators within the overall set of critical success factors serves as a kind of insurance for an effective search. Also, the existing structure may be a timesaver and provide candidates for leading or lagging indicators.

Interviewing Decision Makers about What Keeps Them Awake at Night

In addition to testing the dashboard structure, developers should directly solicit from the line manager key work, goals, and risk areas and answers to the question "what keeps you awake at night?" Some of this information may be more topical than what is represented in the metric dashboard. For example, the manager may be concerned about recent staff turnover or funding shortfalls that could have an adverse impact on deliverables. Or a new set of environmental requirements may test the ability of the manager and workforce to complete work within regulations. This emerging information may lead to a revised understanding of the critical success factors represented in the metric hierarchy and influence the search for lagging and leading indicators.

Sample Approach "Setting the Stage"

In April 2010 Los Alamos National Laboratory began an initiative focused on improving work control for Research and Development (R&D)-related work after a series of high-severity accidents and near misses involving post-doctoral researchers, students, and recent transfer employees. Figure 5 provides descriptions of a few incidents. One part of the initiative, endorsed by senior leadership, called for developing a set of leading indicators for moderate hazard R&D work.

Description	Subject Matter Expert (SME) Involvement	Changed Work	Work Control	Communications	Student-Mentoring	Work Authorization	Schedule Pressure
National High Magnetic Field Laboratory (NHMFL). May 27, 2009. Student removed a cable connected to the cell safety switch while replacing a pulsed magnet in Cell 3. The safety switch isolated the magnet from the capacitor bank energy source. A worker must be a trained and qualified R&D energized electrical worker to perform maintenance to the capacitor bank system.	Student had not been trained or qualified on the capacitor bank system; original mentor not available to perform that work	A technical staff member (not the mentor) asked the student to replace the magnet in Cell 3; work required knowledge of capacitor bank configuration	Student was not authorized under the proper work control document		Mentor had been reassigned 6 months previously; no replacement was made	Student not authorized to perform capacitor bank repair	
Technical Area 35 Laboratory. July 8, 2009. Operational emergency declared for violent rupture of acid waste container. Addition of acetone to acid waste container resulted because "student did not have adequate technical knowledge for waste determination on his own" and neither mentor explained to student waste disposal practices appropriate for chemical disposal. The work was also not covered by a formal work control document because it had been identified as low hazard.	Lack of SME involvement in explaining to student waste disposal practices appropriate for the chemicals and processes	New hazard introduced of acid-containing solution; not clear if workers properly estimated level of hazard associated with cleaning and disposal for student	Activity identified as low hazard and therefore not requiring a formal work control document	Identified as core issue in the event: mentors did not communicate well with student	Student did not have knowledge and expertise to identify potential chemical reactions and their impact on waste disposal process		
Technical Area 48 Laboratory. January 11, 2010. Explosion of stainless steel pressure vessel in Laboratory. Post-doc had recently changed mentors. Had scaled up chemicals in experiment by a factor of ten without intermediate steps. The use of a General Chemistry work control document applied to the work of chemical synthesis may have contributed to an insufficient level of critical thinking applied to work planning. Unclear how combination of chemicals and heating led to vessel failure.	Post Doc was SME on this type of chemistry; a review likely would not have prevented event		No specific restrictions in work control document would have prevented work; however, use of general work control document "may have prevented critical thinking"	Post doc did not inform mentor or lab owner of intent to use oven		General Chemistry Operations work control document did not permit use of oven	Decision to scale up without intermediate steps "driven by time pressures in competitive area of research"

Figure 5: Sample Approach – Partial List of R&D Occurrences

A subsequent review of Laboratory metrics and critical success factors showed few existing R&D work control measures. And the "what keeps managers awake at night" factor was clearly a series of high-consequence incidents involving R&D work. So this part of the review process was straightforward.

Selecting Indicators

The process of indicator selection falls naturally into two groups, lagging and leading. Lagging metrics characterize the important outcomes we are monitoring for the organization: for example, key product deliverables, vital services, strategic targets, staffing targets, and core capabilities. Leading indicators are those actionable metrics that are expected to influence and predict those outcomes. A key constraint is that leading indicators must be sufficiently early in the process to prevent an undesired outcome.

Selecting Lagging Indicators

One characteristic of lagging indicators is they are often easy to select. They are frequently part of our corporate Dashboards, since the outcomes are obvious and are present in the minds of decision makers. As noted before, the selection of lagging indicators should emerge from the review of the existing metric structure, critical success factors, and customer input, together with decision-maker interviews about "what keeps you up night." Here are a few pitfalls to avoid:

- ***Because-we-have-the-data metrics.*** Many lagging metrics are longstanding and use well-established data sources. However, they may no longer fit the latest set of critical success factors and should be modified to meet the current needs of decision makers.

- ***Spread-the-love metrics.*** Some indicators are in dashboards only because they reliably display a positive aspect of organizational performance, i.e., showing perpetual green status. Be wary of these metrics; they may have too much slop in performance ranges or may not truly represent critical success factors.

- ***No-benchmark-in-sight metrics.*** Metrics sometimes fail because not enough work or research has been done in benchmarking targets. Identifying useful external benchmarks or industry standards set by outside organizations may help in developing a more realistic target or set of performance ranges. Care should be exercised in identifying similar activities or operations for the benchmarks.

- ***Too-many-metrics problem.*** Decision makers may be swamped by a dashboard with too many metrics. Organizations should spend as much effort in prioritizing and deselecting as they do in compiling metrics. Otherwise significant performance may be masked by metric clutter. It may be possible to scale the number of metrics, too, by the judicious use of index metrics.

- ***Everything-and-the-kitchen-sink index.*** This is the flawed counterpart to the too-many-metrics problem. Indexes that combine too many data metrics into a single index can mask performance at lower levels. As much as possible, keep the underlying data metrics to an intelligible set; a good rule of thumb is five or less data metrics for a given index.

The bottom-line question that a lagging indicator should answer is "Is this metric measuring an outcome that my decision makers believe is important?" If the answer is yes, it should be kept. If not, it should be eliminated.

Selecting Leading Indicators

Selecting leading indicators can be a thorny process. It is often hard to identify leading indicators that will have a true cause-and-effect relationship with a particular lagging indicator. Multiple reasons may underlie these difficulties: causes of lagging-indicator performance are poorly understood, underlying processes are not well defined, it is too costly to collect the needed data, or measure ownership is shared across organizational or functional boundaries. Recent studies even suggest that indicator response is more a function of powerful culture and belief systems than proven cause-and-effect linkages.[1]

This section describes a series of analytical and brainstorming activities that can be performed in selecting leading indicators. Note that even after analyses have been performed and metrics selected, the cause-and-effect relationship between the leading and lagging indicators is only assumed. It needs to be tested by acting on the leading indicators and monitoring their effect on the lagging indicators.

The following questions may help guide the search for leading indicators: What is in place that helps reinforce positive outcomes and minimizes the chance of negative outcomes? What is being used to "build a better future"? What are indicators of events and problems of lower significance? What indicators show culture and employee feedback?

Quad Chart Analysis

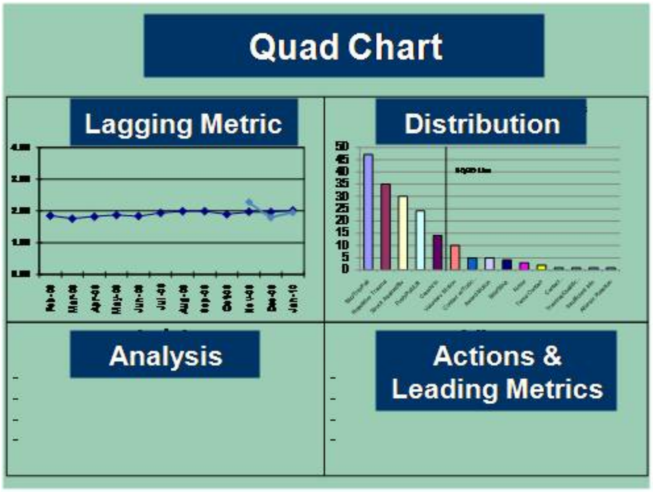


Figure 6: Quad Chart Template

One method of identifying leading indicators is to start by identifying types of events, actions, organizational or personnel distributions, or causes that help to explain the lagging performance result. Figure 6 shows a way of displaying such information – a "quad chart" template. In the template, the lagging metric trend line is displayed in the upper left corner, a Pareto distribution of events is in the upper right, the analysis is set out in the lower left, and proposed actions and leading indicators are in the bottom right.

Note the analysis for a given indicator may be too complex to display in a one-page format. Consequently, a number of drill-downs of the basic quad chart may be required for certain metrics. Also, a Pareto chart does not have to serve as the basis for the analysis. Other analytical techniques (e.g., fishbone diagrams, barrier analysis, etc.) may be more appropriate for the selected indicator and thus incorporated into the template. (See Quantitative or Structured review for further discussion.
The point of the exercise is to arrive at a set of leading indicators and associated actions that flows directly from a causal analysis of the lagging performance result. Thus there is an assumed cause-and-effect relation between the actions/leading indicators and the lagging indicator.

[1] Malina, Mary et al. *Relations among Measures, Climate of Control and Performance Measurement Models.* Contemporary Accounting Research, 2006.

For example, one site may be experiencing a spike of injury-illness incidents involving slips, trips, and falls. A quad chart review indicates that 60% of the total incidents involve falls on broken sidewalks, problem stairwells, or other decaying infrastructure. This may lead to an initiative to target these particular infrastructure items for repair or replacement; decision-makers then develop metrics that monitor the completion of the improvement process across the site. In this way, a quad chart review (or similar causal or analytical process) can be helpful in developing actionable leading indicators.

Process Analysis

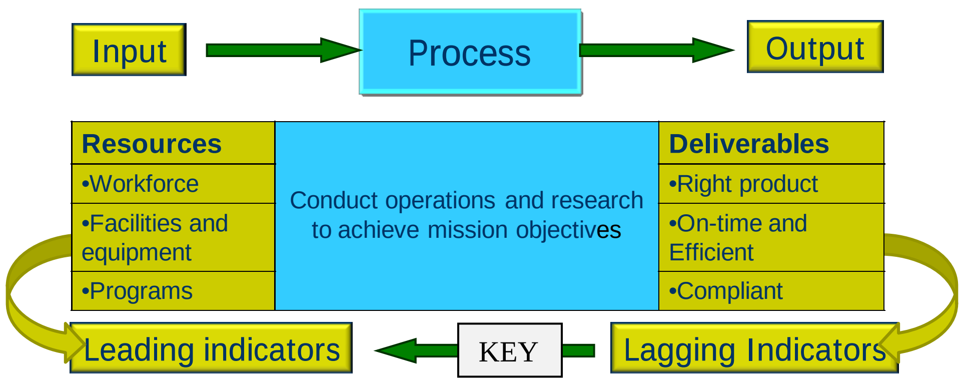


Figure 7: Analysis of Inputs and Outputs

Another method for identifying potential leading indicators is to conduct a process analysis by examining key inputs and outputs. An output is a desired set of deliverables or outcomes. Inputs are those resources that are necessary to achieve the target outcomes. As shown in Figure 7, the outputs may be important products, service results, or regulatory compliance results. Inputs may be related to workforce elements, facilities or equipment, and programs or other key resources.

For example, although it has no current problem producing ten widgets per year, an organization realizes that it must invest significant additional resources to recruit, train, and qualify new staff to continue to meet this goal. Decision-makers therefore develop leading indicators that target this workforce element that puts future deliverables at risk. Resources thus become drivers of the correct products and are a natural source of leading indicators. In this way, a process-oriented analysis is useful because the inputs become a source of leading indicators for the outputs, the lagging indicators.

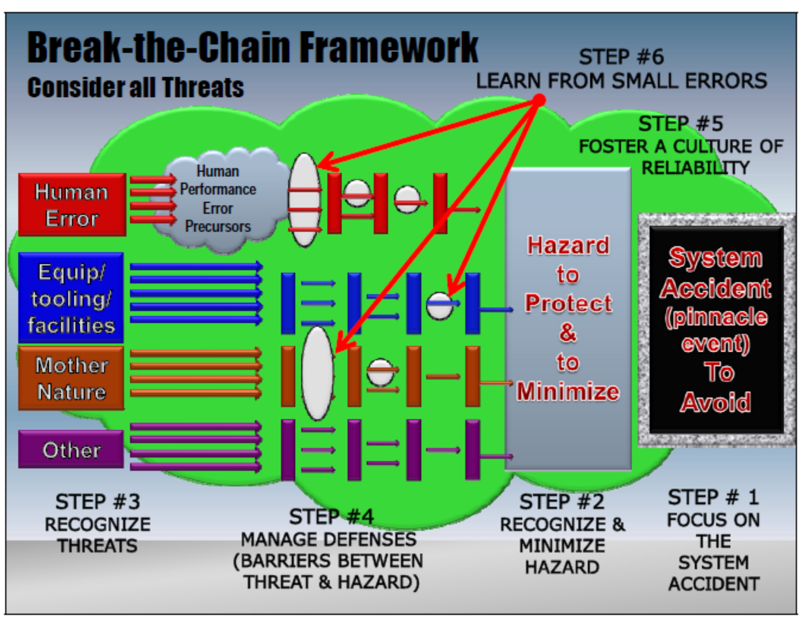


Figure 8: Model of Barrier Analysis

Barrier Analysis

Barrier analysis is a method of managing defenses to prevent unwanted events. Figure 8 shows a model for barrier analysis, developed at Pantex, which reduces the likelihood of system accidents by placing defenses that minimize any unwanted energy flows from threats (i.e., human error, equipment, acts of nature, and other) to the identified hazards. Selection of barriers follows a rigorous evaluation of threats and hazards and observations of work on the shop floor. The type of barriers, number of barriers, and level of effort needed to maintain barriers all depend on the level of consequence and type of hazards associated with the operation.

There is a strong linkage between identification of leading indicators and barrier analysis. Indicators on barrier design, implementation, and maintenance can follow directly from a barrier analysis. They can also lead to organizational learning if the statistics are presented in the proper context. For example, data on barrier failures should be presented as "The first barrier failure is the time to get concerned, not the second or third barrier failure."[2]

Brainstorming

Brainstorming is a key tool for identifying leading indicators. The most obvious rule for successful brainstorming is to enlist the right set of people: invite those subject-matter experts, line managers, and other stakeholders who are knowledgeable of the target activity, process, or organization. The following are additional suggestions:

Figure 9: Brainstorming Activity

- Advertise the session as a brainstorming activity
- Send out any quad charts, process analysis, benchmarking, or other analytical products before the meeting so that participants can "do their homework"
- Ground the session in identified critical success factors, i.e., what management considers important
- Use both a facilitator and a recorder during the meeting
- Write down suggested indicators on a whiteboard, index cards, or flipchart (shown in Figure 8)
- Ask participants "What sorts of resources are necessary to achieve this target . . ."
- Keep the brainstorming separate from metric down-selection

If the session loses momentum, it may be helpful to review the previously developed analytical products and use them as a springboard for discussion. Don't be overly concerned about grounding the discussion in a review of effectiveness, cost, data collection, or other difficulties at this point. The goal of the exercise is to identify possible metrics without too much "editorial mind" being present.

[2] Hartley, Rick, et al. *High Reliability Operations: A Practical Approach to Avoid the System Accident. FY2011 Pantex Barrier Analysis Process.* B&W Pantex, 2010.

Sample Approach "Selecting Indicators"

The team leader for the R&D leading indicator initiative selected a group of line managers and subject-matter experts to evaluate risk in the work control development process and to develop candidate leading indicators. One tool used in the meetings was a risk characterization chart – a hybrid of a quad chart and process analysis. Figure 10 shows the chart.

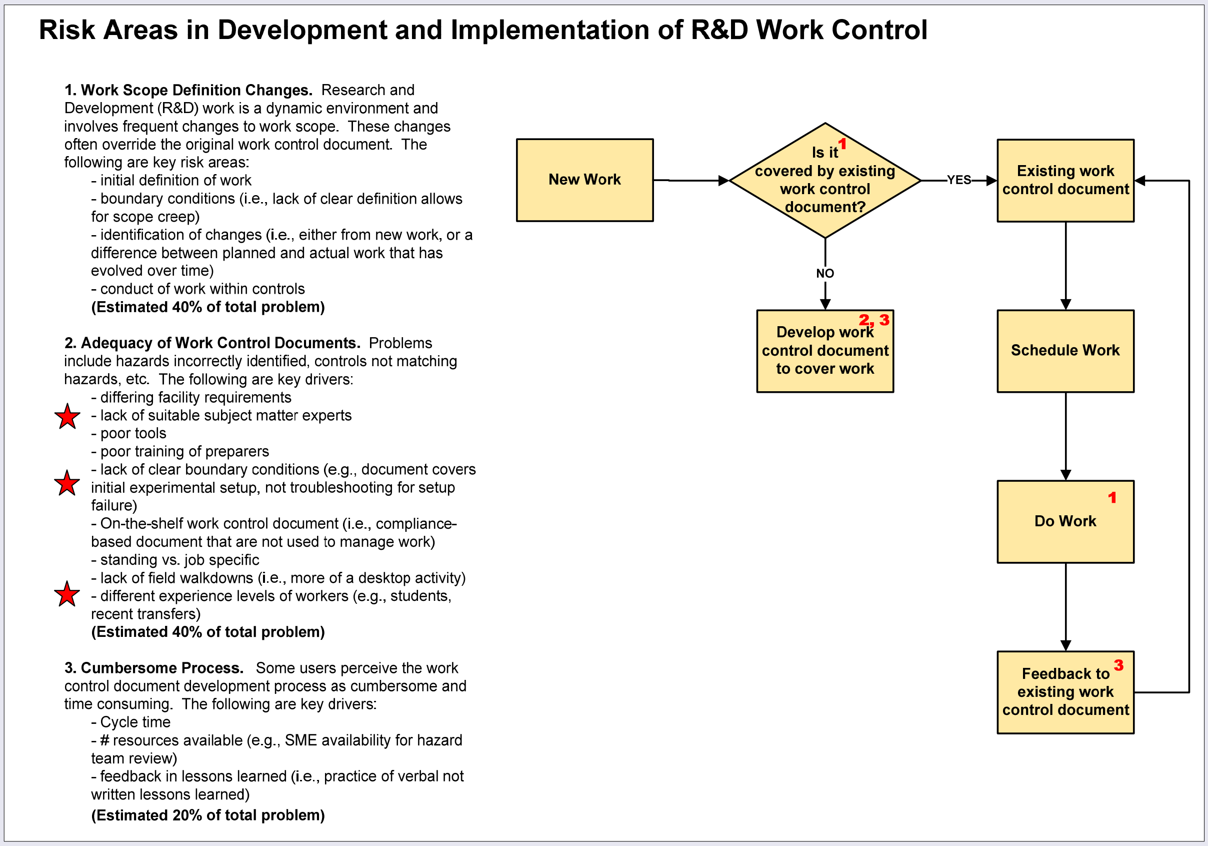


Figure 10: Sample Approach - Risk Area Identification

The team identified the following key weaknesses in R&D work control: (1) work changes in which the initial scope and boundary conditions no longer apply, (2) inadequate work control documents, caused primarily by inadequate subject-matter reviews, poorly defined boundaries, and little provision for different experience levels (Note: the "stars" in the diagram indicate the importance the team assigned to these factors), and (3) a cumbersome work document approval process. On the basis of the risk evaluation, the team brainstormed approximately 20 metrics.

Conducting a Qualitative Review

For many organizations, a qualitative review is the next step in metric selection. The choice of a qualitative or a quantitative/structural review depends on the organizational type, process being reviewed, available data or metrics, and time constraints for metric selection. A qualitative review may be more appropriate for highly complex or rapidly changing organizations, or for organizations without well-defined processes or established metrics.

In general, organizations are more likely to conduct a qualitative than a quantitative review in selecting metrics. Reasons may include the following: the metric owner is more comfortable with an approach that incorporates "management by feel," the metric is new and there is no established data history, or there is a reluctance to employing a more mathematical approach for metric selection.

Figure 11 depicts the basic program of qualitative review. A team of decision makers and subject-matter experts evaluates and selects those metrics that are cost effective and have a high probability of influencing the desired outcome. These metrics are depicted as X's in the upper left quadrant. Note this type of analytical activity is best done separately from the initial metric brainstorming.

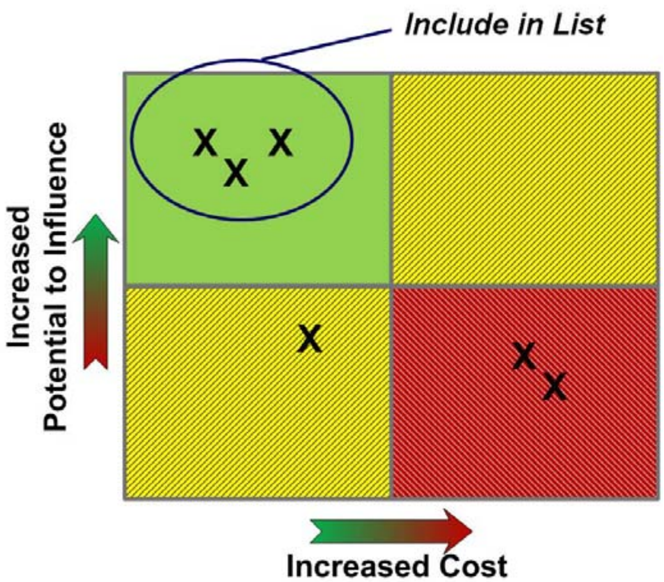


Figure 11: Qualitative Review of Metrics

Factors to consider in the qualitative review are as follows:

- Have the proposed metrics been framed to assist the selection process? An excellent method is to ensure that each proposed metric fulfills SMART (Specific, Measurable, Achievable, Relevant, and Time-Framed) criteria. The process of stepping the metric selection through these elements will often make the strength or weakness of candidate metrics apparent.

- Does the team have sufficient specialized knowledge or expertise to make judgments? Different members from the original metric brainstorming team may be needed to evaluate cost or implementation factors.

- A spreadsheet or other tool to capture the team's qualitative judgments may be useful in structuring the discussion. (See Sample Approach for "Conducting a Qualitative Review.")

Sample Approach "Conducting a Qualitative Review"

After a few brainstorming sessions to identify R&D metrics, the team leader led a qualitative review of the metrics. Each team member evaluated potential metrics in effectiveness, cost and difficulty of implementation, and breadth of application. Metrics received scores from 1 (worst score) to 5 (best score) in each area. Figure 12 shows the overall results for the team.

Table of Potential Indicators		Evaluations of Options				
			Factors			
Option	Description	Total Weighted Performance Score	Effective-ness of Measure (Causation Correlation)	Cost/ Difficulty of Gathering Data	Implemen-tation Time Required	Breadth
		100%	50%	20%	20%	10%
1	Ratio of event critiques to number of reportable ORPS events	3.70	3.0	4.0	5.0	4.0
2	Ratio of human performance improvement (HPI) precursors identified in event critiques to number of critiques	3.90	5.0	2.0	3.0	4.0
3	Cycle time for development of R&D work control documents	3.30	4.0	2.0	3.0	3.0
4	Percent of R&D work control documents w/appropriate boundary conditions	3.30	4.0	2.0	3.0	3.0
5	Percent of R&D work control documents w/appropriate SME review	2.80	3.0	2.0	3.0	3.0
6	Work scheduling process for R&D work	3.50	4.0	3.0	3.0	3.0
7	Percent of R&D work control documents with field or other changes	2.60	3.0	2.0	2.0	3.0
8	Quality of R&D pre-job briefing, detailed work planning discussions	3.30	4.0	3.0	2.0	3.0
9	Percent of field walkdowns completed before start of R&D work	3.00	3.0	3.0	3.0	3.0
10	Ratio of self-reported errors to externally identified errors	4.40	5.0	4.0	3.0	5.0
11	Percent of required walkarounds/work observations completed by managers	3.70	4.0	3.0	3.0	5.0
12	Percent of work and cause codes identified in work control documents for improvement	2.90	2.0	3.0	5.0	3.0
13	Percent of work control documents passing quality review	2.20	1.0	2.0	5.0	3.0
14	Injury/illness rates by organization	3.00	2.0	4.0	4.0	4.0
15	Training completion by organization	2.30	1.0	3.0	4.0	4.0

Figure 12: Sample Approach – Qualitative Review

A few metrics received high scores in potential effectiveness but low scores for cost or implementation. For example, a metric for cycle time in developing R&D work control documents received a high score in potential effectiveness, but a low score in cost because of the difficulty of collecting data from non-centralized sources.

Other metrics, such as training completion rates and injury-illness rates, had high scores in ease of data collection but low scores for effectiveness. The training completion rate metric was viewed as too broad for the specific goal of improving R&D work control. And the injury-illness metric was viewed purely as an outcome metric and not actionable.

The end result was a set of four highly scored metrics: (1) ratio of event critiques to the number of reportable occurrences, (2) ratio of human performance improvement (HPI) precursors identified at event critiques to the number of critiques, (3) self-reported errors versus externally driven reports, and (4) number of required walkarounds and observations done by managers. In justifying their choices, the managers/subject matter experts believed these metrics and their associated actions emphasized workforce engagement, improving work control, reducing fear of mistakes, disseminating lessons learned, and reinforcing learning throughout their organizations.

Because of the absence of process metrics for work control, however, the team chose the cycle time metric even though it had a lower score. Performance in this area, some participants believed, affected the workforce willingness to describe new hazards for and modify work control documents. The senior leader overseeing the process endorsed the selection.

Conducting a Quantitative/Structured Review

Quantitative/structured reviews of metrics may be more applicable to organizations that have well-established data histories and well-defined processes, e.g., manufacturing processes. Such formal reviews may be used to identify mathematical correlation or special types of causation between lagging and leading indicators. A complete discussion of mathematical and causal analyses is beyond the scope of this document. However, the following brief, non-exhaustive survey is offered.

Correlation

Correlation measures the degree of association between two variables. It is not a true measurement of causality: two variables can be highly correlated without being causally linked. An everyday example is alarm clocks ringing and roosters crowing. There is a high degree of correlation between the two data sets, but no causal relationship.

One correlation test, called Pearson's, is the most common measure of correlation and is expressed as a number between -1 and +1. The closer the value is to +1, the more an increase in one variable is associated with an increase in the other variable; the closer the value is to -1, the more an increase in one is associated with a decrease in the other.

For the metric developer, a correlation test provides an indication of whether a leading indicator is even associated with the lagging indicator. Although it doesn't describe causality, the test can provide confidence that the leading and lagging indicator *may* have a causal connection, or provide a justification for further analysis. They may be connected by an undiscovered third variable. The test could lead to a further causal exploration, experiments, or tests of causes for indicator performance.

Aside from the Pearson's test, there are a series of tests that encompass non-linear relationships: Spearman's rank correlation, Kendall's rank correlation, and multi-moment correlation.

Statistical Process Control

Statistical Process Control (SPC) is a method that uses control charts as a key technique for identifying and reducing process variability. The method may also employ techniques such as histograms, Pareto charts, cause-and-effect diagrams, defect diagrams, and scatter diagrams. By examining the behavior of observed data within the control chart, and seeing if it obeys certain rules, the metric developer or analyst can determine if the process is operating within or out of control.

The concept of operating "within control" or "out of control" is the central basis for the management use of SPC. A process which displays "out of control" behavior has results that cannot be predicted from the previous process results. These results include patterns that are significantly higher or lower than previous results, and are not easily explainable as random noise. Especially in the context of leading indicators, "out of control" results represent trends or changing conditions. "Out of control" results (i.e., trends) should be investigated to determine the cause of the change. The results of the investigation may then be used to correct for trends in the adverse direction, and to reinforce trends in the improving direction.

Conversely, if the data are "within control" (i.e., no trends), then SPC theory (Dr. Shewhart's Economic Control of Quality of Manufactured Product) states that no amount of investigation into an individual result will provide information to improve the results of the process. Process results that are "within control" imply that that process is stable and predictable. If improvement is needed, then the entirety of the results for the stable period must be analyzed, causes found, and management action taken to change the process itself.

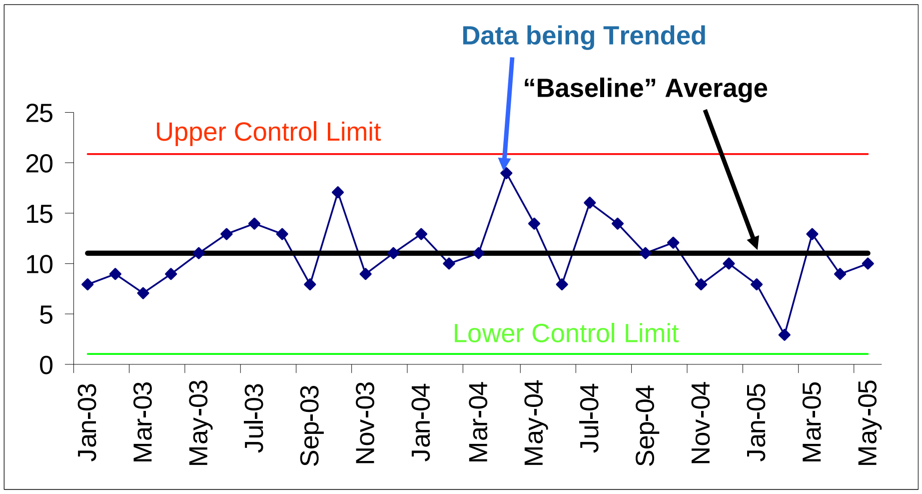


Figure 13: Control Chart

Figure 13 shows several features of a control chart. A typical control chart plots the data itself, a center line at the average of the data, and upper and lower boundaries, called control limits, formed by lines three standard deviations on either side of the data average. The average line is based upon a fixed "baseline" set of data.

The average line should not be adjusted with the addition of new data results. The average line is only adjusted following the detection of an "out of control" set of data which stabilize at a new level. Some authors also establish lines at two standard deviations from the average as "warning limits". The rules for determining if process results are "within control" or "out of control" do vary from author to author. It is important to pick one set of rules and stick with that set, rather than varying the rules depending upon the desirability of detecting a trend or not. (See Appendix 4 for an example of trending analysis using Statistical Process Control.) A common set of rules that could be used is as follows:

- One point outside the three standard deviation control limits
- Two of three points two standard deviations above/below the average line
- Four of five points outside one standard deviation above/below the average line
- Seven points in a row one side of the average line
- Ten of eleven points in a row one side of the average line
- Seven points in a row all increasing/decreasing

As a causal discovery technique, statistical process control is applicable to both leading and lagging indicators and useful for finding performance drivers and data sets. A set of data "out of control" (displaying a trend) on a leading indicator may be the result of a culture change and may provide an important signal of a shift in culture.[3]

Causal Analysis

Causality is the relationship between two events. The first event is known as *the cause* and the second event is known as *the effect* and is presumed to be the consequence of the first event. Causality is not limited to events but can incorporate objects, processes, facts, properties and variables. There are

---

[3] Prevette, Steve. *SPC Trending Primer*. Fluor, 2011.

numerous common methods for identifying causes of events and conditions, some of which are described in Appendix 5.

In addition to these techniques, there are a few causal methods that have more of a quantitative orientation. The following is a brief discussion of five such models: directed acyclic graphs, causal loop diagrams, the Rubin causal model, design of experiments, and Granger analysis.

- **Directed acyclic graphs (DAGs)**, while not necessarily implying causation, have been used extensively in statistical and artificial intelligence applications in structuring time and causation variables. DAGs are not only appealing in that they can be constructed from judgment and knowledge of causal relationships, but they also permit response to external or spontaneous changes.
- **Causal loop diagrams (CLDs)** are diagrams that aid in visualizing how interrelated variables affect one another. The diagram consists of a set of nodes representing the variables connected together. The relationships between these variables, represented by arrows, can be labeled as positive or negative.
- **Rubin causal model (RCM)** is an approach to the statistical analysis of cause and effect based on the framework of potential outcomes.
- **Design of experiments (DoE)** or experimental design is the design of any information-gathering exercises where variation is present, whether under the full control of the experimenter or not. In the design of experiments, the experimenter is often interested in the effect of some process or intervention (the "treatment") on some objects (the "experimental units"), which may be people, parts of people, groups of people, plants, animals, etc. Design of experiments is thus a discipline that has very broad application across all the natural and social sciences.
- **Granger analysis** is a technique for determining whether one time series is useful in forecasting another. Ordinarily, regressions reflect "mere" correlations, but Clive Granger, who won a Nobel Prize in Economics, argued that there is an interpretation of a set of tests as revealing something about causality. A time series X is said to Granger-cause Y if it can be shown, usually through a series of F-tests on lagged values of X (and with lagged values of Y also known), that those X values provide statistically significant information about future values of Y. Granger testing is commonly used to look for causality in comparing economic data.

Recent studies have identified a problem in establishing actual cause-and-effect linkages between leading and lagging indicators in a performance model (e.g., balanced scorecard).[4] If decision makers use performance indicators as part of the decision-making process, then those decisions will affect the future states of both leading and lagging indicators. Malina characterized this problem as, "Managers adapt the firm's actions and the underlying production function to [the model] and other feedback (hence, statistics are unstable)". In other words, the products of management decisions introduce variability that affects the future measurement outcomes. As with the need to control variability in the design and implementation of experiments, it is important to do the same in the analysis of data for management decision making, lest the statistics become unstable.

Causal analysis is still useful for understanding the individual states and trends of leading and lagging indicators in order to provide information useful to decision makers.

---

[4] Malina, Mary, et al. *Relations among Measures, Climate of Control and Performance Measurement Models.* Contemporary Accounting Research, 2006.

Using and Refining

Using and refining the selected indicators is the final stage of the leading indicator development process. In many ways, this stage is the most important because it indicates the degree to which an organization is invested in using the select metrics to monitor operations, develop actions, and improve processes.

Using the Selected Leading Indicator

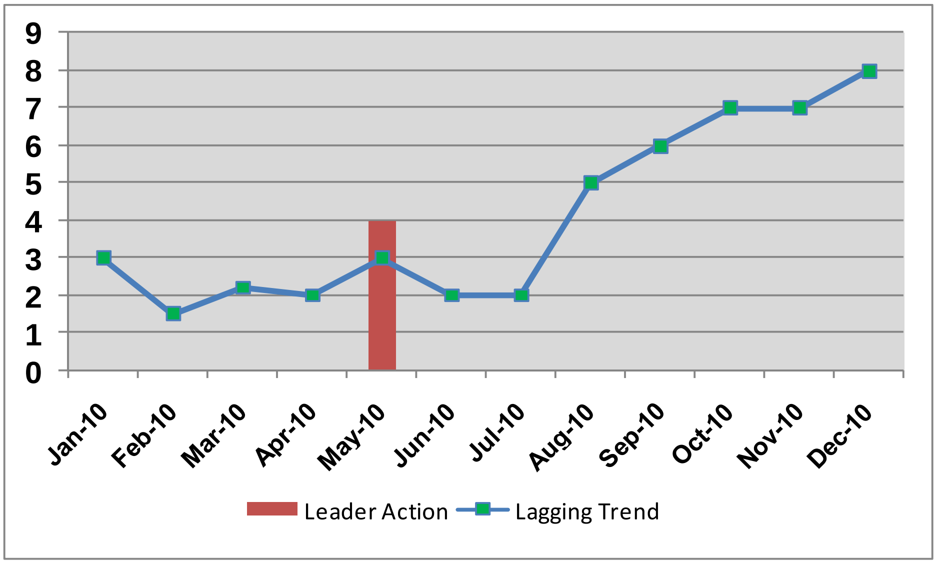


Figure 14: Using the Leading Indicator to Influence Outcome

Once the metric has been selected, the manager monitors the performance and selects actions to influence the desired outcome. If the metric produces the desired outcome, then the metric is retained and the data continues to be monitored. If the metric does not produce the desired outcome, then the metric is eliminated and a new search for leading indicators begins. Care needs to be taken in eliminating leading indicators. Additional factors may be influencing the indicator's behavior, e.g., initial expectations may have been unrealistic, or insufficient time may have been allowed to effect change. Sometimes adding a leading indicator may be a better course of action.

Figure 14 shows an instance of a leading indicator's associated action that provides a distinct improvement in the trend line of the lagging indicator. In reality it can be more difficult to discern the impact of actions associated with a set of leading indicators.

Collecting Data on Leading Indicator Usage/Management

Managers and by extension the workforce should actively use the set of leading and lagging indicators to manage their activities. Data collected on the use of the leading indicators can be important in tracking this engagement. The following are a few suggestions for this type of data collection:

1. Dashboard population data. A metric showing that indicators are being populated on time is one indication that the metric is being used at a basic level. It answers the question: "Is the metric being used in a timely manner?" If not, then it indicates the selected metric is no longer relevant or a true critical success factor.

2. Commentary or action data. The end-game for a leading indicator is its association with organizational actions. Some dashboards provide mechanisms for tracking actions, or for providing commentary on the status of actions. Figure 15 provides one template for evaluating actions. Included are (1) a description of problem status, (2) identification of trends, (3) planned actions and timeline for completions, and (4) a general impact statement.

3. Metric review data. Annual quality reviews of metrics can identify whether metric descriptions, weights, and other characteristics are current and being maintained. Such reviews can become a springboard for metric refinement and process improvement.

4. User data. Data on individual users, user organizations, and other data on usage of online dashboards can be useful in determining how deep metric engagement is within a given organization.

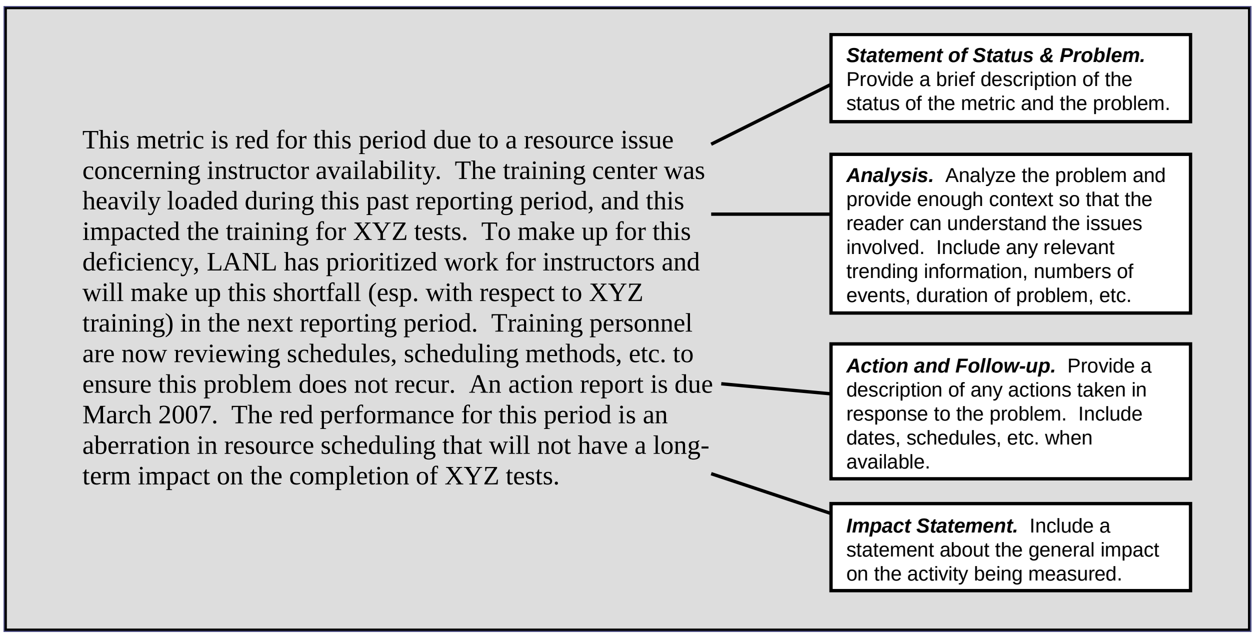


Figure 15: Evaluation of Indicator Commentary

Refining Family of Leading Indicators

Over time, the selected leading indicators should be reevaluated and revised, especially if there have been systemic changes to the working environment where they have been used. The achievement of major organizational or enterprise goals is an example of this type of change. These same changes should motivate decision makers to consider adding new leading indicators to help support future decision making, or eliminating those that no longer add value. While leading indicators remain in use, targets should also be adjusted periodically to reflect improvement goals.

Using and refining the family of leading indicators maximizes their effectiveness as elements of a system that supports decision makers.

Conclusion

To be effective, organizational metrics need sustained care and feeding. Metrics that are not reviewed, that do not consider potential outcomes, that do not represent success factors, or that dry up from user neglect—are not worth having.

So this paper ends with an idea. The idea is that an organization's metrics, whether they are lagging or leading, should become part of routine thought experiments. Does this metric represent what management considers important? Are there better ways of representing success or risk? What type of causal analysis, logic evaluation, or thought process produced the metric? Has the organization brought together those stakeholders who can analyze performance and provide the best insights? These types of considerations can improve the leading and lagging indicators in our metric hierarchies and lead to process improvements.

Contributors

1. Rich Reichelt, Los Alamos National Laboratory
2. Annette MacIntyre, Lawrence Livermore National Laboratory
3. Greg Durfee, Sandia National Laboratory
4. Steve Prevette, Savannah River Site
5. Rick Hartley, Pantex Plant
6. Scott Sterling, Hanford Site
7. Cynthia Bixby, Nevada Site Office
8. Jane Fitzpatrick, Savannah River Site
9. Bob Stuewe, Los Alamos National Laboratory
10. Marc Clay, Los Alamos National Laboratory

Bibliography

- Brown, Mark Graham. *Keeping Score: Using the Right Metrics to Drive World-Class Performance.* New York: Quality Resources, 1996.
- Freedman, David, and Robert Pisani, Roger Purves, and Ani Adhikari. *Statistics.* New York: W.W. Norton & Company, 1991.
- Frost, Bob. *Measuring Performance.* Dallas: Measurement International, 2000.
- Frost, Bob. *Designing Metrics.* Dallas: Measurement International, 2007.
- Harbour, Jerry. The *Basics of Performance Measurement, 2nd Edition.* New York: Productivity Press, 2009.
- Hartley, Rick, et al. *High Reliability Operations: A Practical Approach to Avoid the System Accident. FY2011 Pantex Barrier Analysis Process.* B&W Pantex, 2010.
- Kaplan, Robert, and David Norton. *The Balanced Scorecard: Translating Strategy into Action.* Boston: Harvard Business School Press, 1996.
- Kaplan, Robert, and David Norton. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment.* Boston: Harvard Business School Press, 2001.
- Los Alamos National Laboratory. *Causal Analysis and Corrective Action Development.* Management Systems Procedures, 2010.
- Malina, Mary A., and Hanne S. O. Nørreklit and Frank H. Selto. *Relations among Measures, Climate of Control and Performance Measurement Models.* Contemporary Accounting Research, 2006.
- Manuele, Fred. *Leading & Lagging Indicators.* Professional Safety, 2009.
- Massachusetts Institute of Technology, INCOSE, and PSM. *Systems Engineering Leading Indicators Guide, Version 2.0.* 2010.
- Montgomery, Douglas. *Introduction to Statistical Quality Control.* New York: John Wiley & Sons, 2001.
- Niven, Paul. *Balanced Scorecard Step by Step: Maximizing Performance and Maintaining Results.* New York: John Wiley & Sons, 2002.
- Olson, Dave. *SRR Approach to Safety Management via Leading Indicators.* 2010.
- Prevette, Steve. *SPC Trending Primer.* Fluor, 2011.
- Reason, James. *Managing the Risks of Organizational Accidents.* Hampshire: Ashgate, 1997.
- Spitzer, Dean. *Transforming Performance Measurement.* New York: AMACOM, 2007.

Appendices

Appendix 1: Template for Site Development of Leading Indicators
Appendix 2: Leading Indicator Library Template
Appendix 3: Generic DMAIC Process
Appendix 4: Example of Trend Analysis Using Statistical Process Control
Appendix 5: Table of Some Common Causal Analysis Methods

Appendix 1: Template for Site Development of Leading Indicators

Setting the Stage	
Metric Hierarchy	
Area or Function (e.g. program/operation, supporting operations; assurance function)	
What are the key success factors for the organization?	
Have goals, key deliverables, multiyear strategy, customer input, etc. been incorporated?	
How well do these success factors translate into the family of metrics?	
Are risk factors (those elements that could harm the key success factors) represented?	
Does each tier contain what is necessary and sufficient for the tier above?	
Where do any potential leading indicators fit into this family of metrics?	
Interview managers	
Name or role	
What keeps you up at night?	
Event or condition?	
Contributors or causes?	
Expected performance or behavior?	
Priority?	

Appendix 1: Template for Site Development of Leading Indicators (cont.)

Selecting Indicators	
Potential lagging Indicators	
Indicator for condition?	
Measurable (units)?	
Time frame (periodicity)?	
Owner?	
Data Source?	
Dimensions (customer, organizations, objective, strategies)	
Potential Leading Indicators	
Indicator for lagging indicator?	
Measurable (units)?	
Time frame (periodicity)?	
Owner?	
Data Source?	
Dimensions (customer, organizations, objective, strategies)	
Qualitative Review	
Down Select Potential leading Indicators	
Base evaluations available? (causal or process factors, type distributions, fishbone diagrams, Pareto charts)	
Effectiveness?	
Cost to gather data?	
Time to implement?	
Breadth or Scope?	

Appendix 1: Template for Site Development of Leading Indicators (cont.)

Quantitative Review	
Test for causation and correlation	
Use appropriate correlation tools (Pearson's product-moment correlation, Brownian correlation or popular non-linear rank coefficients (Spearman, Kendal) Tau)	
Use appropriate causality tools (causal loop diagrams, DAGs, Granger causality)	
Down Select Potential leading Indicators	
Select indicators resulting from correlation and/or causality analysis	
Effectiveness?	
Cost to gather data?	
Time to implement?	
Breadth or Scope?	
Using and Refining	
Collect data on leading and lagging indicators	
Dashboard data	
Commentary data	
Metric review data	
User data	
Test and trend suite of lagging and leading indicators	
Trend leading indicator performance in relation to lagging indicator performance	
Re-evaluate and review (Change in working environment, Change in organizational goals, Annually)	

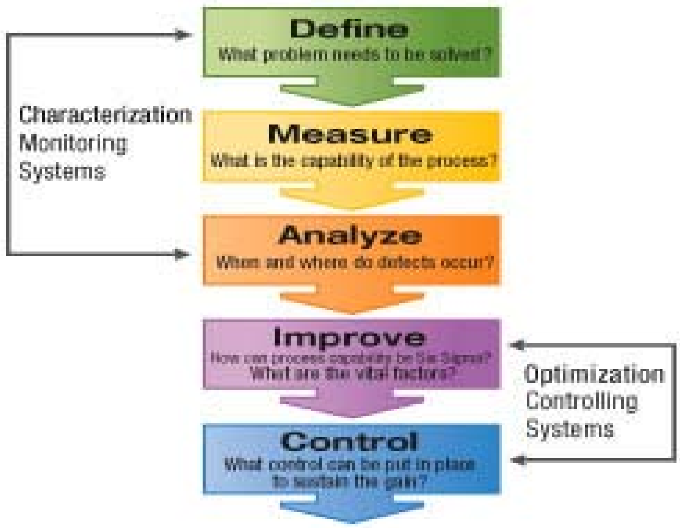
Appendix 2: Leading Indicator Library Template

<table>
<tr><td colspan="2">Area or Category: (e.g. Specific process, Funding, General Operations, Safety, Security, Energy, Environment, Business)</td></tr>
<tr><td colspan="2">Contact information: (e.g. name and e-mail)</td></tr>
<tr><td>Lagging Indicator name:</td><td rowspan="2">Lagging Indicator description:</td></tr>
<tr><td>Driver: (“what keeps you up at night”)</td></tr>
<tr><td colspan="2">#1 Leading Indicator</td></tr>
<tr><td>Name:</td><td rowspan="2">Description: (e.g. explain what it measures, purpose, periodicity, type, units)</td></tr>
<tr><td>Target/Goal and Thresholds:</td></tr>
<tr><td colspan="2">Anticipated or Actual Results: (e.g. actions taken, effectiveness, unintended consequences, external factors influencing performance, cost and time to implement, information display (e.g. graphs, charts, tables))</td></tr>
<tr><td colspan="2">#2 Leading indicator</td></tr>
<tr><td>Name:</td><td rowspan="2">Description: (e.g. explain what it measures, purpose, periodicity, type, units)</td></tr>
<tr><td>Target/Goal and Thresholds:</td></tr>
<tr><td colspan="2">Anticipated or Actual Results: (e.g. actions taken, effectiveness, unintended consequences, external factors influencing performance, cost and time to implement, information display (e.g. graphs, charts, tables))</td></tr>
<tr><td colspan="2">#3 Leading indicator</td></tr>
<tr><td>Name:</td><td rowspan="2">Description: (e.g. explain what it measures, purpose, periodicity, type, units)</td></tr>
<tr><td>Target/Goal and Thresholds:</td></tr>
<tr><td colspan="2">Anticipated or Actual Results: (e.g. actions taken, effectiveness, unintended consequences, external factors influencing performance, cost and time to implement, information display (e.g. graphs, charts, tables))</td></tr>
</table>

Appendix 3: Generic DMAIC process

The DMAIC process is a core component of the Six Sigma methodology. It is used when making improvements to an existing process. (For creation of new processes and products, the DMADV framework is followed.)

DMAIC is an acronym for the 5 key phases in a process improvement project: Define, Measure, Analyze, Improve, and Control



Define

In the ***Define*** phase, the project team clarifies the purpose and scope of the project and confirms that a DMAIC project is in fact appropriate. A key tool for this phase is the project charter. The project leader, typically a Black Belt or Green Belt, coordinates with team members and other stakeholders to compile the charter document. The charter should spell out the objective of the project, the expected timeline and budget, the scope, and the major players.

Also created during the Define phase is the SIPOC – a diagram which identifies the process being examined, the inputs to and outputs of the process, and the relevant suppliers and customers. This ensures that all team members understand the process itself at a high level.

Another important aspect of the Define phase is the gathering of VOC (Voice Of the Customer) data. Six Sigma is focused on finding out directly from customers what their idea of quality is, and how well the current process meets that standard.

Appendix 3: Generic DMAIC process (cont.)

Measure

During ***Measure***, the focus is on gathering data to describe the current situation. It is critical to identify the appropriate process measures and gather sufficient baseline data, so that once improvements are made the impact can be verified empirically.

A detailed process map is created, including documentation of variations in how the process is carried out. With this information the project team can begin to see some of the factors that may be affecting process performance.

Analyze

The purpose of the ***Analyze*** step is to determine the root causes of the process problems and inefficiencies. A variety of methods are used to identify potential root causes, narrow down the possibilities, and confirm the relationship between the suspected causes and the performance of the process.

Statistical analysis is a key component of this step, and is used to demonstrate these relationships.

Improve

The next step, ***Improve,*** involves establishing a means of countering the root causes. Techniques involve brainstorming, FMEA (Failure Modes and Effects Analysis), and piloting the improvement plan before rolling it out in full.

The same data that was obtained during Measure to establish the baseline is again gathered after improvements are in place. Data analysis and charting techniques are used to confirm that performance has in fact improved sufficiently to meet the project's goal.

Control

Finally in the ***Control*** phase, steps are taken to ensure that the gains obtained during Improve are maintained. Common tasks include setting up ongoing data tracking and a plan for identifying when the process performance starts to slip and taking appropriate action. At the end of this phase, the project manager transfers ownership back to the process owner, and the team communicates the project results to all stakeholders.

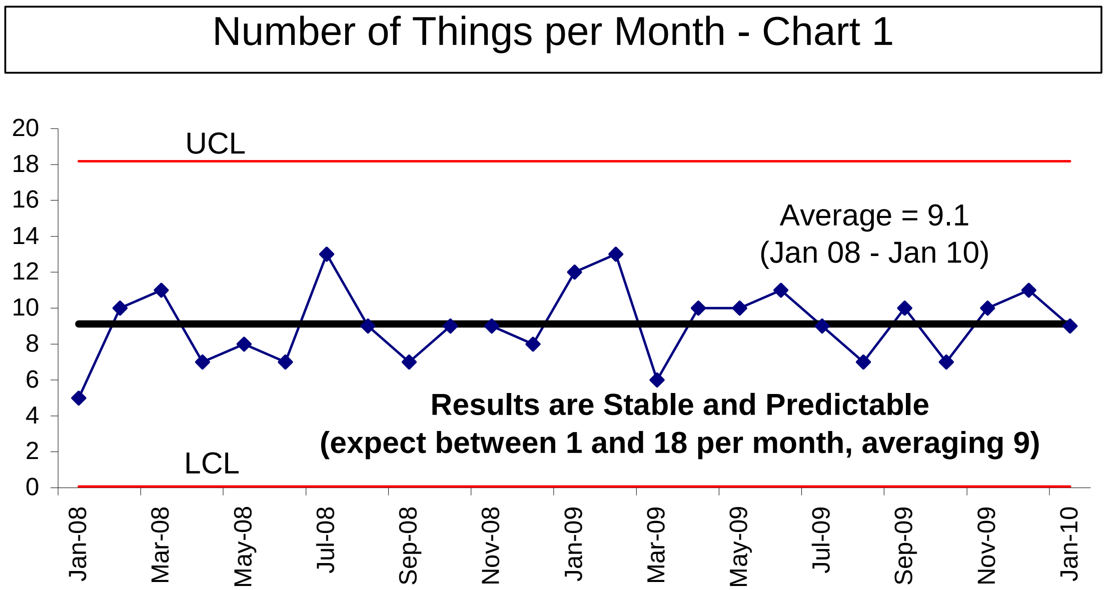
Conclusion

By following the Six Sigma DMAIC methodology, process improvement can be accomplished in a way that is systematic, sustainable, confirmed with data, and in alignment with customer and stakeholder quality expectations.

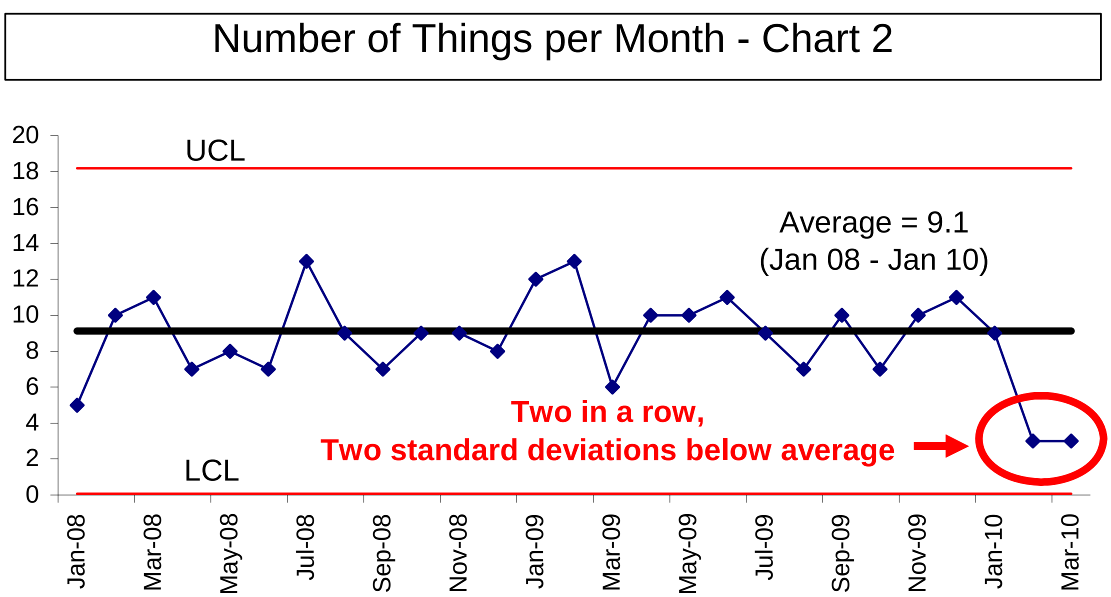
From **Introduction to DMAIC** by Heidi Wiesenfelder as published in Bright Hub
http://www.brighthub.com/office/project-management/articles/23736.aspx

Appendix 4: Example of Trend Analysis Using Statistical Process Control

Let us assume we have chosen an indicator (leading or lagging) and are plotting it on a monthly basis. The example is titled “Things Per Month”, and the first 25 months have been stable and predictable.



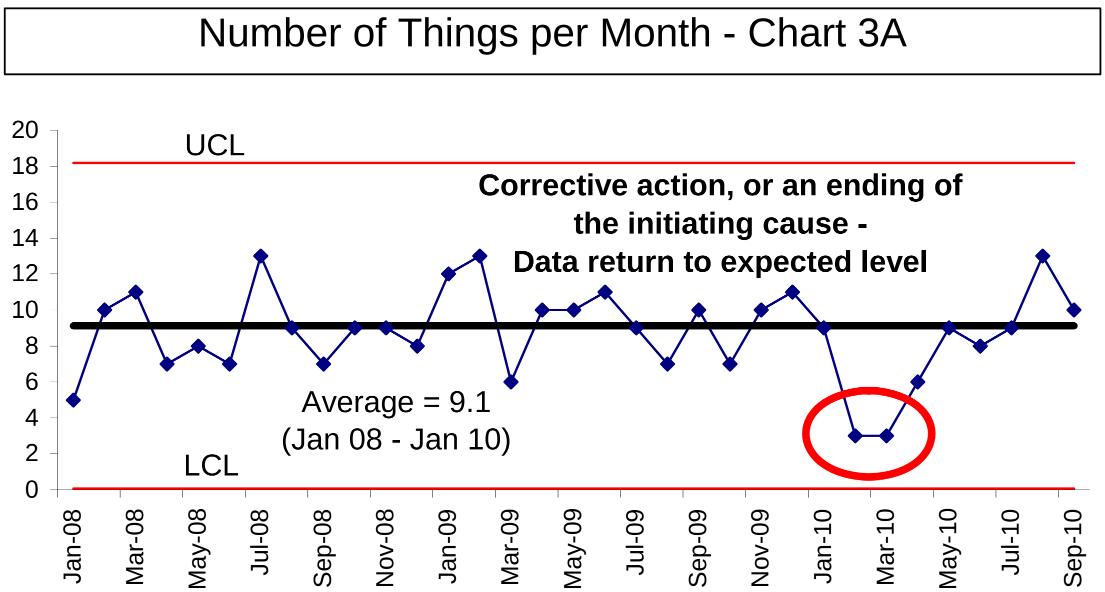
Next, let us assume something has caused the number of things per month to drop. Initially, we don’t know the reason, but we do detect that there are two months in a row more than two standard deviations below average.



Appendix 4: Example of Trend Analysis Using Statistical Process Control (cont.)

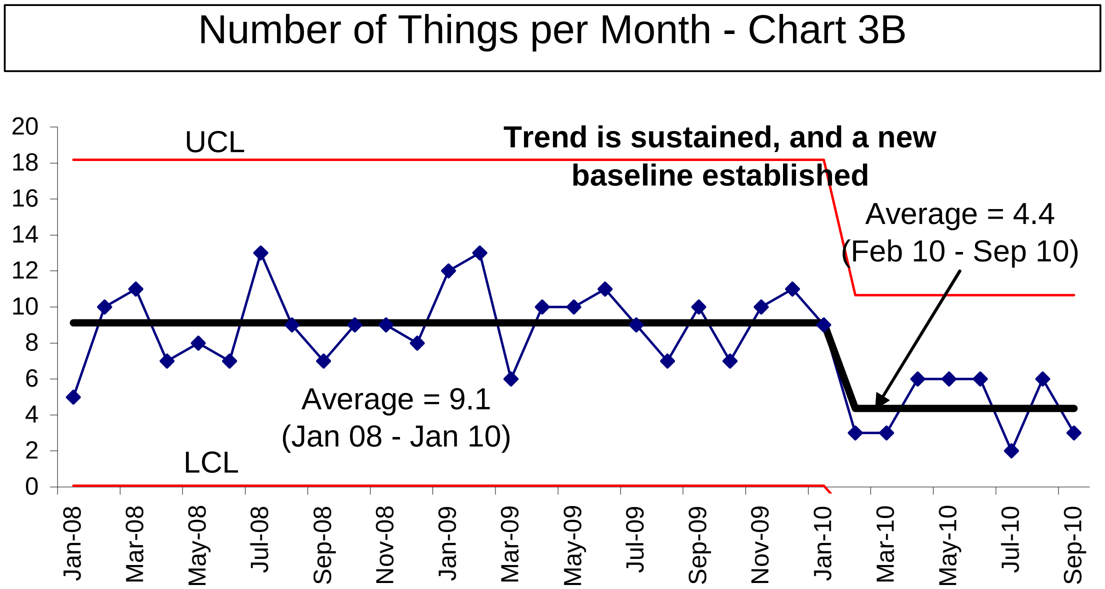
Following this detection of a trend, there will be one of two results. Either whatever caused the trend will end and the data return to normal, or whatever caused the trend will remain in effect, and the data will stabilize at a new baseline level with a new prediction.

The case below is represents either a corrective action was taken (because "Things per Month" dropping was a non-desirable trend) or the source of the trend only was in effect for two months (a scheduler of "Things to Do" was on vacation). In this case, we leave the trend circled for historical purposes, and the average and control limits are not shifted.



The case below represents a permanent shift in the rate of "Things per Month". Either the reduction was a desirable trend (such as events) and we ensured that what caused the trend remained in effect, or there was some permanent shift in the process outside of our control that we have not been able to correct for. In this case, a new baseline average and control limits were established, and we note that the prediction of future performance is between 0 and 10 per month, averaging 4 per month.

Appendix 4: Example of Trend Analysis Using Statistical Process Control (cont.)



When to (and whether to) develop a new baseline is best left to the analyst, though the decision should be a result of

- A significant trend has occurred (following the trend rules)
- Knowledge of what caused the trend (and is it expected to be permanent)
- Have the data made "three changes of direction" since the trend started.

Empirical evidence has shown the three-changes-of-direction criteria to be reasonable criteria to determine if enough data are present for a baseline calculation.

Appendix 5: Table of Some Common Causal Analysis Methods

Technique	Description	Use	Advantages	Disadvantages
Events and Causal Factor Charting	A chronology of key events or conditions, represented graphically, that lead to an unwanted outcome or condition.	Good to use for all event timelines.	Provides a visual display of the analyst or team's understanding of the contributing events or conditions.	May draw conclusions regarding unrelated conditions as being related and causal.
Why-Why Analysis	A technique of asking a series of why questions for a given condition or fault. Successive answers can begin to reveal underlying root causes for the condition.	Good for condition-related problem.	Provides a structured evaluation process.	Without care to answer each "why" with the immediately influencing events or conditions, will make leaps in logic.
Barrier Analysis	Provides for systemic identification of barriers (i.e., admin controls, engineering barriers, etc.) that failed to prevent or mitigate a hazardous condition or event.	Good tool for examining existing defenses that proved vulnerable and led to an identified failure or problem.	Provides an understanding of defense failure modes, defense in depth, and multiple defense vulnerabilities necessary for the given pathway to failure.	Requires training to use effectively.
Change Analysis	A chart that pairs a listing of events or conditions in an event or error-likely situation with those in an error-free situation.	Helps to focus on changes in equipment, facility, process, personnel, or routine that may have contributed to problem.	Provides systematic approach to understanding changes in unfolding event that influenced the resulting outcome.	Requires training to use effectively.
Ishikawa, Fishbone, or Cause-Effect Diagram	A diagram that displays the major and minor causes that lead to particular result.	Good tool for examining condition-related problems, such as process failures or audit findings. Can be paired with a Why-Why analysis.	• Easy to apply with minimal training • Provides analyst with relevant topical areas that may affect outcome, leading to a "why" analysis in each area.	Without care to answer each "why" with the immediately influencing events or conditions, will make leaps in logic.
Failure Modes and Effects Analysis (FMEA)	A bottom-up method of identifying and prioritizing failure modes according to consequences, frequency, ease of detection, and risk.	Good tool for identifying potential failures and causes; can be used in design process.	Good for exhaustively cataloging all possible faults.	May be limited by team's experience with previous failures.
Fault Tree Analysis	A top-down method of analyzing the effects of initiating faults and events on a complex system.	Fault tree constructed for showing the impact of an undesired fault and finding ways for system improvement.	Good at showing how resistant a system is to faults.	Not good at examining multiple failures.
Pareto Analysis	A technique built on observation that 80% of problems are produced by 20% of the causes.	Uses a distribution chart to describe the key contributors to a problem.	Provides a compelling graphical display of key contributors.	May need to be combined with other techniques to arrive at causes.
Human Performance Improvement	A technique for identifying human performance gaps. It includes a review of significant error precursors, mode of operations, and equipment or facility conditions.	Not actually a separate causal analysis technique, but a method of informing other causal analysis investigations.	Good at identifying the human and work environment context for error conditions.	Requires training to use effectively.