



HARVARD Kennedy School

MOSSAVAR-RAHMANI CENTER
for Business and Government **30** YEARS

Improving Information Use By Enhancing Performance Measures

Kendra M. Asmar
Danjell H. Elgebrandt

July 2015

M-RCBG Associate Working Paper Series | No. 47

The views expressed in the M-RCBG Fellows and Graduate Student Research Paper Series are those of the author(s) and do not necessarily reflect those of the Mossavar-Rahmani Center for Business & Government or of Harvard University. The papers in this series have not undergone formal review and approval; they are presented to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). Papers may be downloaded for personal use only.

IMPROVING INFORMATION USE BY ENHANCING PERFORMANCE MEASUREMENT

*-Recommendations for Air Force Materiel
Command*

31 March 2015

Kendra M. Asmar
Danjell H. Elgebrandt
John F. Kennedy School of Government
Harvard University
Master in Public Policy, 2015

Brigadier General (USAF, Ret) Dana Born, Advisor
Phil Hanser, Seminar Leader
John Haigh, Seminar Leader

This PAE reflects the views of the authors and should not be viewed as representing the views of the PAE's external client, nor those of Harvard University or any of its faculty.

The views expressed in this PAE are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

Acknowledgements

We would like to thank our Faculty Advisor, Brigadier General (USAF, Ret) Dana Born for her invaluable help with our PAE. Her feedback and advice has been central both for the general direction of our work as well as with many detailed challenges we have encountered along the way.

We would also like to thank Colonel David Sieve, Gregory Dierker, and Jeffrey Glendenning at Air Force Materiel Command for devoting their time, resources, and expertise in working together with us to complete our PAE.

Lastly, we would like to thank our seminar leader, Phil Hanser, who has been a great support to us from start to finish.

Contents

Acknowledgements 1

Executive Summary 4

Opportunity 6

Problem Statement 8

Methodology 8

Literature Review 8

Case studies 12

Durham Constabulary 12

Implications for AFMC 13

Royal Air Force (RAF) Chaplain Service 13

Implications for AFMC 14

RAF Performance Management 14

Implications for AFMC 15

New York Police Department (NYPD) 15

Implications for AFMC 16

United States Army (USA) 16

Implications for AFMC 17

Store 24 17

Implications for AFMC 18

Lessons Learned 18

Design 18

Implementation 19

Maintenance 19

Recommendations 19

Optimal number of metrics 20

Discussion 20

Making the metrics more usable 22

Discussion 22

Conclusion 24

Appendix A: AFMC Background 25
Appendix B: SMART 28
Appendix C: Metric Evaluation Questions 29
Appendix D: Metric Evaluation 30
Appendix E: Selection Discussion..... 45
Appendix F: Cause/Effect Example 47
Appendix G: Acronyms 48
Appendix H: Glossary..... 49
Bibliography 50

Executive Summary

Sequestration, which took effect in 2013, resulted in over \$1 trillion being cut from the Department of Defense (DOD). Faced with drastic budget cuts and a 9 percent reduction in all financial accounts, Air Force Materiel Command (AFMC) began a drastic reorganization from twelve to five centers to help absorb the effects while retaining combat readiness and stability for the Airmen.¹ Congress requires quarterly performance reports to ensure effectiveness and efficiency. AFMC established 95 performance metrics to determine how the organization is performing and to provide Congress. However, 95 metrics may be too many to consistently track and to actively develop interventions at the command level. Very few studies show — or even evaluate — the effectiveness of performance measurement (PM) systems. Research suggests that PM systems must include the following three aspects to be effective; metrics that:

- inform decision making — if data does not help make decisions, it is unnecessary*
- look towards the future — goals in the strategic plan should be the focus of metrics*
- contain an element of flexibility — once goals are reached, new metrics need to emerge for a continuously developing strategy*

After analyzing AFMC’s 95 metrics, we determined that the following ten performance metrics may be a fruitful focus for the command:

	Metric Name	Aggregated Metrics
1.	<i>Continue to Strengthen AFMC’s Support to the Nuclear Enterprise (AFNWC), Metric 1.1</i>	<i>1.1.1.1 through 1.1.1.5</i>
2.	<i>Advance Today’s & Tomorrow’s Combat Capabilities through Leading-Edge Science & Technology (AFRL), Metric 1.2</i>	<i>1.2.1.1 through 1.2.3.1, and 3.1.1.7</i>
3.	<i>Acquire & Support War-Winning Capabilities ‘Cradle-to-Grave’ (AFLCMC), Metric 1.3</i>	<i>1.3.1.1 through 1.3.2.2, and 3.1.1.1 through 3.1.1.3</i>
4.	<i>Perform World Class Test and Evaluation (AFTC), Metric 1.4</i>	<i>1.4.1.1 through 1.4.2.1</i>
5.	<i>Sustain AF Capabilities through World-Class Depot Maintenance & Supply Chain Management (AFSC), Metric 1.5</i>	<i>1.5.1.1 through 1.5.1.5, and 3.1.1.5 through 3.1.1.6</i>

¹ General Janet C. Wolfenbarger, Harvard Kennedy School Harris Lecture Forum Event, November 6, 2014.

6.	<i>Standardize and Continually Improve Processes</i>	<i>2.1.1.1 through 2.2.1.1</i>
7.	<i>Cost Effectiveness</i>	<i>1.6.1.1 through 1.6.1.3, 3.1.1.4, and 3.1.2.1 through 3.3.2.3</i>
8.	<i>Recruit, Develop, and Retain a Competent Workforce</i>	<i>4.1.1.1 through 4.2.3.2 and 4.4.1.1</i>
9.	<i>Secure and Improve Installations and Infrastructure</i>	<i>4.3.1.1 through 4.3.2.1 and 4.5.1.1 through 4.5.2.1</i>
10.	<i>Assess Health of Each ACS Functional Area & Advocate for Capability Needs, Metric 5.1</i>	<i>5.1.1.1 and 5.1.2.1</i>

These metrics are essential to strategic decision-making. We also suggest “flagging” metrics for review — any metric that is, or projected to be, underperforming should also be discussed at the monthly meetings. We recommend implementing the following best practices (developed from an analysis of literature and case studies) to make their PM system more useful:

- evaluate the set of metrics in use on a regular basis*
- create a uniform system of “traffic light” metric indicators*
- create a standard for metric description and definition*
- use trend analysis on metrics*
- group together related metrics*
- develop a system to store data in order to track and project trends*

A PM system is essential to demonstrate the effectiveness of AFMC’s decisions. Two years after AFMC reorganized, the command has saved over \$6 billion (in direct savings or cost avoidance)², leading the way in budget efficiency for the Air Force. We believe through the implementation of our recommendations, AFMC’s metrics will become even more useful and sustainable.

² General Janet C. Wolfenbarger, Harvard Kennedy School Harris Lecture Forum Event, November 6, 2014.

Opportunity

Air Force Materiel Command (AFMC) currently uses 95 metrics to measure performance with the intent to track progress toward achieving its goals. Due to the restructuring of the command from twelve centers to five centers (Appendix A), Congress requires a quarterly review of the metrics to ensure AFMC is performing efficiently. Thus far, the reorganization has been successful and has saved the Air Force over \$6 billion (in cost avoidance or savings) and Chief of Staff of the Air Force General Mark A. Welsh III has started calling AFMC the “cost-consciousness of the Air Force.”³ Each metric is associated with a priority outlined by the strategic plan. Information on these metrics is stored and managed by a senior AFMC staff member. Every month, he is responsible for collecting information from the “goal champions” — an individual assigned to monitor and input information on their metric — and formatting a presentation for a monthly meeting regarding the status of the metrics. The dashboard is a snapshot of the current health of the organization. The goal champions have the ability to upload additional information on the metric to the dashboard and manually input a trend line. Once a month, the AFMC Commander and senior staff members meet for one to two hours to review 25 to 35 predetermined metrics. The 95 metrics are presented based on the availability of the data and priority of the metric. Each metric is viewed either monthly, quarterly, semi-annually, or annually. The review is meant to “capture active status and trend data,” enabling leadership “to discuss progress, root cause analysis, and mitigation strategies.”⁴ The information discussed at each review is vital; however, 95 metrics is unwieldy at the command level.

AFMC’s strategic plan is intended to communicate a roadmap. The Commander was intimately involved in the development of all 95 metrics; her intention was to use them as an overview of her organization and as guideposts for her command. The importance of performance measurement as a management tool is well-known and can be used to “(1) evaluate; (2) control; (3) budget; (4) motivate; (5) promote; (6) celebrate; (7) learn; and (8) improve.”⁵ The General’s concerns with PM when she took command were similar to the concerns that most other managers see with PM: (1) lack of motivation to set the bar high; (2) tendency to measure what

³ General Janet C. Wolfenbarger, Harvard Kennedy School Harris Lecture Forum Event, November 6, 2014.

⁴ United States Air Force, *Air Force Materiel Command Fiscal Year 2013 Quarterly Metrics*, 2013.

⁵ Robert Behn, “Why Measure Performance,” *Public Administration Review* 63, no. 1 (2003): 588.

external stakeholders care about, though the information may not be valuable internally; and (3) tendency to be large and unwieldy.^{6,7} For example, she did not want to set too low of a goal for her command. She also did not want to measure irrelevant metrics in order to please external stakeholders if the metric provided no use for the command. The baseline used for the metrics is prior performance in the organization and Air Force regulations.

Despite the progress that AFMC has made framing the new organization and establishing their metrics, there are several unresolved concerns that we were tasked with considering. We analyzed each of the 95 metrics to determine if there should be more or less information for each one; evaluate the logic tree of the priority, goal, and metric; and determine if any of the metrics lack value for the command.

When employing a dashboard it is crucial that the organization does not over-rely on the snapshot image of its performance. It is possible for an organization to show, for instance, a highly effective organization despite the long-term prospects looking bleak. Conversely, it is possible for an organization to focus too heavily on a set of metrics that suggest an organization in distress while a more careful analysis suggests that the future is promising. A McKinsey report clarifies this concept, “A patient visiting a doctor may feel fine, for example, but high cholesterol could make it necessary to act now to prevent heart disease. Similarly, a company may show strong growth and returns on capital, but health metrics are needed to determine if that performance is sustainable.”⁸

By refining and limiting the current performance metrics used by AFMC, the organization will be better prepared to accomplish their mission of equipping the Air Force in the changing environment that it faces today. We have based our analysis on the following assumptions to guide our research and recommendations: (1) AFMC has clearly defined strategic goals; (2) the current metrics are tightly aligned with strategic goals or deficits (areas that need

⁶ General Janet C. Wolfenbarger, Harvard Kennedy School Faculty Roundtable Discussion, November 7, 2014.

⁷ Robert Kravchuk and Ronald Schack, “Designing Effective Performance-Measurement Systems under the Government Performance and Results Act of 1993,” 1996.

⁸ Richard Dobbs and Timothy Koller, “Measuring Long-Term Performance,” McKinsey Quarterly, http://www.mckinsey.com/insights/corporate_finance/measuring_long-term_performance.

improvement)⁹; (3) the data gathering process is reliable; and (4) no mandates exist that require the inclusion of any one metric in our recommendations.

Problem Statement

Our analysis focused on:

1. *The optimal number of metrics that should be used at the command level of AFMC.*
 - a. *How often should the organization review these metrics?*
 - b. *How should the metrics be used at the command level, taking into consideration the interdependencies of certain metrics?*
2. *How the metrics can become more usable.*
 - a. *How might the metrics translate data into usable information to aid decision making?*
 - b. *How might the metrics transform into more anticipatory information to prevent a problem before it arises?*

Methodology

We first completed a literature review on performance measurement and cognitive psychology. We then identified organizations that implemented PM systems and balanced scorecards (BSC) to varying degrees of success. We focused on multi-tiered organizations which included for-profit companies, other military branches (both American and foreign), and governmental organizations (such as police departments).

Literature Review

Although there are countless papers written on performance measurement, very few studies illustrate how PM should be used to improve an organization. There are several reasons for this. The business world is fast-moving, with a primary emphasis on profit. Businesses, despite establishing PM systems, rarely follow up to ensure their programs are effective and fulfilling their goals. New management tools continually arise and evolve (for example LEAN, TQM, PM,

⁹ Robert Behn, "Measurement, Management, and Leadership," *Bob Behn's Performance Leadership Report 11*, no. 3 (2013).

BSC, etc.). Managers eagerly adopt the practice, but often the anticipated results do not materialize. There is always pressing competition and other matters requiring managers' attention. Sufficient time and money are often not invested in making sure the new tools work.

Additionally, if a company chose to invest in a study to determine if their PM tool worked, a control group would be hard to select. There are two ways for managers to implement a new system. They could implement the system slowly, in a series of stages, or require the entire company to adopt the tool at the same time. If they implemented the system in stages, the control group may not reflect how the rest of the company will adopt the practice. In other words, one subgroup of an organization does not represent all other subgroups. Scaling up is a difficult process that requires several elements—cost effectiveness, commitment, capacity, community buy-in, and cultural change.¹⁰ If these elements are not met, success of scaling up is unlikely.

The other option is implementing a change throughout the entire company at once. In this case, the control group would be the company. The study would have to compare company performance before and after the tool was implemented. If the study appeared to prove that the tool was effective, other companies would want to implement the tool as well. However, each organization has a unique competitive advantage and culture that makes it difficult to export management tools from one company to another. Given the differing characteristics among companies, a study would only be valid for the company where the tool was tested.

Regardless of how useful performance measurement can be, a PM system will produce few gains if it is poorly designed and implemented. If gains do occur as a result of a PM system, they are often incremental and go unnoticed. When PM is the catalyst for large gains, some other management tool or event usually gets the credit.¹¹ This also leads to a problem with internal validity; it is difficult to identify just one reason why a company starts to perform better.

We also conducted research in cognitive psychology to help inform our recommendations. Working memory is one's ability to keep information readily available in order to process

¹⁰ Mark Fagan, *HKS MLD 601: Operations Management Tool Kit* (2014).

¹¹ David N. Ammons, "Performance Measurement and Managerial Thinking," *Public Performance & Management Review* 25, no. 4 (2002).

information and make decisions.^{12,13} The capacity of working memory is relatively small, with one's limit being seven new "chunks" of information, plus or minus two. (This is contested in academia, with many researchers claiming the average number is closer to three or four.)¹⁴ Grouping together information may help increase the amount of information one can handle. For example, it is easier to remember a phone number in chunks of three or four digits than to remember all ten digits at once. Additionally, the more often one works with the information, the less likely they are to be constrained by the limitations of their working memory.¹⁵ When a person exceeds their capacity to process information, they suffer from information overload. This leads to negative effects such as anxiety, stress, and poor decision making.

There are several best practices that organizations should implement to increase the effectiveness of their PM systems. The metrics should create a "balanced and focused direction, while aligning to your desired end point."¹⁶ This balance should include internal, external, and financial measures.¹⁷ Metrics should "communicate to senior management whether the company is progressing toward stated goals or is stuck in a holding pattern."¹⁸ The metrics chosen should "encourage performance improvement, effectiveness, efficiency, and appropriate levels of internal controls."¹⁹ The goal of metrics is to generate a discussion among leaders, allowing them to make more informed decisions. To enable managers to make decisions, the metrics should be explicitly linked to the goals of the company and should track trends over a shorter period. When a metric stops driving change (because a goal is reached or the environment changes), that metric should be changed or a new one developed to replace it. The SMART

¹² Pascale Michelon, "What is Working Memory? Can it Be Trained?" SmartBrains, <http://sharpbrains.com/blog/2010/11/16/what-is-working-memory-can-it-be-trained/>.

¹³ Connie Malamed, "20 Facts You Must Know About Working Memory," *The eLearnig Coach*, <http://thelearningcoach.com/learning/20-facts-about-working-memory/>.

¹⁴ George Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity For Processing Information," *Psychological Review* 63, (1956).

¹⁵ Connie Malamed, "20 Facts You Must Know About Working Memory," *The eLearnig Coach*, <http://thelearningcoach.com/learning/20-facts-about-working-memory/>.

¹⁶ Bob Champagne, "Too Many KPI's? Tips for Metrics Hoarders," *Performance Perspectives*, <http://epmedge.com/2011/02/24/too-many-kpis-tips-for-metrics-hoarders/>.

¹⁷ R. Johnston, S. Brignall, and L. Fitzgerald, "'Good Enough' Performance Measurement: A Trade-Off between Activity and Action," *The Journal of the Operational Research Society* 53, no. 3 (2002).

¹⁸ "Measuring Success: Making the Most of Performance Metrics," *General Electric Capital Corporation*, 2012, <http://www.americas.gecapital.com/insight-and-ideas/capital-perspectives/measuring-success-making-the-most-of-performance-metrics>.

¹⁹ "University of California [Performance Measurement] Approach," *Oak Ridge Associated Universities*, 2005, <http://www.orau.gov/pbm/documents/overview/uc.html>.

(specific, measurable, accountable, realistic, timely) test can be used to determine the quality of any particular metric (Appendices B and C).²⁰

Enterprise software company Oracle recommends starting with a large number of metrics and gradually reducing the quantity to fit the organization.²¹ Cascading scorecards — scorecards that provide cause-and-effect links among the multiple levels of an organization — are also beneficial for a multi-tiered organization. This allows the data gathering procedure and the use of the metrics to be aligned. Lastly, despite its difficulty, incorporating qualitative insights is important. Qualitative is not equivalent to subjective. The use of “softer” metrics is fine, even recommended, as long as they can be measured objectively. In a 2011 study about why balanced scorecards fail in large organizations, Dr. Nopadol Rompho cites two categories of failures that may jeopardize the implementation and/or effective use of a BSC approach — design failures and process failures.²²

Design failures include too many/few metrics or metrics that do not align with the organization’s strategy. The authors identify the following factors that contribute to process failures: (1) having too few individuals involved in the BSC process; (2) having too long of a development process for the scorecard; (3) treating the scorecard as a one-time measurement project; and (4) lacking senior management commitment. They also warn about the scorecard becoming the center of operations with the organization striving to improve only what is measured in the scorecard rather than using PM as an important management tool among many.

*Mark Graham Brown, in his book *Keeping Score* emphasizes the time aspect of BSCs. A proper scorecard takes into account the past, present and future.²³ He also points out that a proper stakeholder analysis must be performed before metrics can be selected so that no stakeholder is*

²⁰ George Doran, “There’s a S.M.A.R.T. way to write management’s goals and objectives,” *Management Review* 70, no. 11 (1981).

²¹ Frank Buytendijk, “Zen and the Art of the Balanced Scorecard,” Oracle, 2008, <http://www.oracle.com/us/solutions/business-intelligence/063553.pdf>.

²² Nopadol Rompho, “Why the Balanced Scorecard Fails in SMEs: A Case Study, Faculty of Commerce and Accountancy, Thammasat University, 2011, <http://www.ccsenet.org/journal/index.php/ijbm/article/viewFile/10247/8988>.

²³ Mark Graham Brown, *Keeping Score: Using the Right Metrics to drive World-Class Performance* (Boca Raton, FL: CRC Press, 1996).

left without their most important metrics in the scorecard. It is also necessary to consider whether the metrics are leading or lagging indicators. Leading indicators are important as “lagging indicators without leading indicators don’t tell the story behind the outcomes....They don’t provide critical early warnings if you are you off track....On the flip side, leading without lagging drill down too heavily on short-term performance.”²⁴ However, research shows that lagging indicators often “foreshadow movements in leading indicators.”²⁵ Lagging indicators can result in patterns that signal upcoming events. A balanced mix of leading and lagging indicators is necessary.

Case studies

In this section we look at real world cases in which BSCs have been used. We identified generalizable lessons that can be transferred to the challenges AFMC faces.

Below is a brief summary of some of these cases focusing on what the respective organizations tried to accomplish, how they did so and how this is relevant to AFMC. We paid particular attention to the selection of performance metrics but also looked at other relevant aspects when called for.

Durham Constabulary²⁶

While small (1,370 officers and 950 police staff), Durham Constabulary in England faced challenges similar to those of AFMC. As a public organization, performance per cost unit had been unsatisfactorily tracked over the years and the constabulary opted for a BSC approach in order to address the situation.

In its “Plan on a Page,” so named because the goal was to present the strategy in a very condensed format, the constabulary outlined four areas:

²⁴ Kevin Lindsay, “For the Clearest Market Insight, Analyze Both Leading and Lagging Indicators,” *Entrepreneur*, September 11, 2014, <http://www.entrepreneur.com/article/236847>.

²⁵ Alex A. Burkholder, “New Approaches To The Use Of Lagging Indicators,” *Business Economics* 15, no. 3, pg. 20.

²⁶ Bernard Marr and James Creelman, “Performance Management, Analytics, and Business Intelligence: Best Practice Insights from the Police,” *The Advanced Performance Institute* (www.ap-institute.com), 2012, http://www.ap-institute.com/media/465536/120118_durham.pdf.

1. *what we need to be good at (core deliverables)*
2. *what will help us do it (enabling factors)*
3. *how we will align our resources*
4. *how we will deliver value for money*

Using Key Performance Questions (KPQ) and Key Performance Indicators (KPI), the constabulary began mapping measurable metrics to the specific areas which they measure. The constabulary implemented the “Plan on a Page” at the highest organizational level and then proceeded, through a process called cascading, to break it down and apply it on lower levels. This process has the advantage of ensuring that the plans of the sub-areas are aligned with the organization as they are aggregated in a bottom-top approach. For multi-tiered organizations it is crucial to ensure that progress at lower levels contribute to the larger organizational goals. The constabulary also implemented a “business intelligence” approach similar to a dashboard. This enables progress evaluation in real time and helps identify changes due to trends or macro variables from changes that are due to organizational practices.

Implications for AFMC

This case highlights how a changed attitude towards metrics can be a driver of change. It also shows how cascading needs to be properly designed and, lastly, how a snapshot of the organization can be used to facilitate long-term change.

Royal Air Force (RAF) Chaplain Service²⁷

The RAF is of natural interest to our PAE since the RAF is, arguably, the non-US organization that shares the most traits with the USAF. The RAF operates a Chaplaincy service, tending to the spiritual needs of its personnel. The Chaplaincy service developed a strategic plan that utilized a BSC but lacked the proper tools to evaluate it and the BSC fell into disuse.

²⁷ Bernard Marr, Ian Shore, and Paul Major, “‘Believing’ in Performance: Measuring and Managing What Matters in Chaplaincy,” *The Advanced Performance Institute* (www.ap-institute.com), 2012, http://www.ap-institute.com/media/454753/120103_raf_chaplaincy.pdf.

The organization created a “strategic mission plan” which sought to enable Chaplains to prioritize their work against centrally determined priorities. The RAF felt that the output of their Chaplains could not be measured in a traditional measurement sense. That is, merely measuring the number of services held or pastoral visits would not provide useful information in an organization where quality must be prioritized over quantity. This presented the RAF with a challenge: how does one properly measure quality in an organization where two very complex fields — military and religion — overlap?

A distinction was made between metrics that are subjectively measurable and metrics that are objectively measurable. Subjective metrics are of great value but if implemented incorrectly could make the scorecard less accurate. In this case, 11 out of 38 KPQs could be measured objectively. This illustrates an important point: by redefining metrics, they can be properly evaluated and be more useful. It also shows that when metrics are selected, it is important not to rely only on easily accessible ones but instead survey the operation in its entirety and select the metrics that best fit the strategic plan.

Implications for AFMC

The importance of measuring relevant information and not merely information that is easily accessible is stressed in this case. The need to align metrics with the strategic plan shows that selecting metrics is a multi-stage process where one needs to determine if information is relevant, measurable and providing explanatory value to the development of an ongoing transformative process in an organization.

RAF Performance Management²⁸

Another relevant RAF case deals with cascading. Although the RAF engaged in PM, staff felt there was a mismatch between the approach used and the ideal state. A number of comments from station executives highlighted the problems:

²⁸ Bernard Marr and Ian Shore, “Cascading Balanced Scorecards: Using Strategic Maps to make Performance Relevant to RAF Stations,” *The Advanced Performance Institute* (www.ap-institute.com), 2008, http://www.ap-institute.com/media/4411/cascading_balanced_scorecards.pdf.

- *“The current system doesn’t provide us with the relevant information.”*
- *“Measurement is for reporting only — we don’t use any of the data we capture.”*
- *“The current scorecard is not very useful, I would like to use it as a management tool, but it doesn’t capture much data that is relevant to me.”*
- *“We only measure what is easy to measure, not what really matters.”*
- *“Data is not providing us with the necessary insights — we need more subjective assessments.”*

With these problems in mind, the RAF opted for a KPQ/KPI approach. The number of KPQs had grown to an unmanageable number, accompanied by an even greater number of KPIs. Identifying the source of too many metrics to study proved helpful for the RAF as this insight allowed them to narrow the number of metrics used. Also, the stations were informed that since cascading was in focus, the KPQs and KPIs were to be reinvestigated on a regular basis and changed to ensure the best fit between the different levels of the organization.

Implications for AFMC

Since AFMC suspects that they are currently using too many metrics, this case is valuable. The problems cited are problems that need to be kept in mind while surveying the situation at AFMC. There needs to be a well-defined idea regarding what the metrics are ultimately used for.

New York Police Department (NYPD)²⁹

The NYPD embarked on a BSC approach to reduce crime rates in New York City through improved follow-up of its police force’s operations. In 1994, the project started with the unveiling of five new strategies, each with a specific (and measurable) goal, such as reducing the number of guns on the streets of NYC.

Buy-in from the policemen was secured through the use of focus groups designed to gather information necessary to the success of the overall strategy. This was beneficial as it directly involved the people who would carry out the strategies.

²⁹ Eli B. Silverman, *NYPD Battles Crime: Innovative Strategies in Policing*, (Boston: Northeastern University Press, 1999).

The NYPD had a history of failed organizational reforms. The solution came in the form of a program called CompStat, short for “compare statistics.” The program emphasized case ownership by the policemen. Individual officers were given more freedom, unleashing creativity and thinking outside the box. All precincts had to provide data on a weekly basis. The NYPD compiled lists of particularly crime-ridden areas and recorded these trends on area maps.

While CompStat provided the NYPD with valuable information, the flow of information became unmanageable. In order to address this, briefing books were compiled for senior staff highlighting the most important metrics. CompStat also allowed for trend analysis of crimes committed in NYC and illustrated them in a clear, easily understandable format that allowed officers to link cause and effect and better understand root causes. The NYPD used CompStat to clearly define goals and provide an incentive (competition and pride) for police officers.

Implications for AFMC

The need for ownership of the respective metrics is something that aligns with the idea of goal champions at AFMC. The case implies that in an organization with a history of frequent reforms, ensuring that everyone involved believes in the proposed reform is of great importance.

United States Army³⁰ (USA)

The USA was a pioneer in using BSCs and since the early 1990s, scorecards have helped them evaluate combat readiness. The deployment of forces into war zones and potential war zones is preceded by an evaluation using scorecards to assess whether troops are sufficiently prepared.

A Harvard Business School study found that:

The balanced scorecard enabled the US Army to become leaner, more nimble, and technically advanced to achieve its mission of serving the American people, protecting national interests, and fulfilling military responsibilities. Using an aggressive BSC rollout through automation and

³⁰ “Balanced Scorecard Working for the US Military,” Advanced Performance Institute (www.ap-institute.com), <http://www.ap-institute.com/Balanced%20scorecard%20working%20for%20the%20US%20military.html>.

education, the US Army managed to transform its organization of military personnel stationed around the globe.

The BSC used by the Army enabled cascading to lower levels. The US Army Corp of Engineers uses nine metrics, aggregated into four areas, to evaluate the organization's performance.

- Mission: is it supporting the goals of the wider US Army and the nation as a whole?*
- Client/customer: is it working to improve its client, customer and stakeholder relationships?*
- Business practices: how efficient are its processes, and are they improving?*
- Capability and innovation: is it translating innovations, expertise, and learning into knowledge and improved business practices for mission accomplishment?*

Implications for AFMC

The success reported in this case all but proves that the redesign of the AFMC PM system will have significant effects, since a very similar organization has shown that this was the case.

Store 24³¹

Store 24 owns and operates convenience stores in New England. Faced with increasing competition in a mature business environment in the late 1990s, the company decided to implement a BSC to track store performance and assist in their simultaneously implemented “fun and interactive shopping experience” idea, an attempt to redefine the brand. The BSC approach did not deliver as expected and was subsequently abandoned. The business case identifies a number of failures in the strategy:

- the different levels of the organization were not connected*
- the strategy map was not used as a tool to figure out how to improve selected metrics*
- management did not unanimously support the strategy chosen*

³¹ Dennis Campbell, Srikant Datar, Susan Kulp, and V.G. Narayanan, “Testing Strategy with Multiple Performance Measures: Evidence from a Balanced Scorecard at Store 24” (2008).

This case serves as a cautionary tale about what can happen when important success factors such as buy-in, cascading, and perceiving the scorecard as a means to an end are ignored and/or mismanaged.

Implications for AFMC

This case shows that it is not enough to have a sound idea about how to improve a PM system; proper implementation and maintenance is extremely important as well. While the potential of PM in multi-tiered organizations is larger than for other organizational structures, the increased complexity can undermine the efforts to reform the organizations unless managed properly.

Lessons Learned

Based on our studies of relevant literature and cases we have drawn a number of conclusions. Some of these lessons are universal and can be easily adopted; others may need to be carefully adapted to the unique structure and purpose of AFMC.

A number of cases make references to KPIs and KPQs. While AFMC does not use this terminology, we believe that these terms closely resemble the priorities and performance metrics of AFMC. We therefore based our recommendations in part on KPI/KPQ thinking.

As both the literature and the case studies suggest, it is important to not treat a PM system as a one-time change in management. The system must be designed, implemented and maintained with a number of lessons in mind in order for it to maximize the chances of organizational success. We have broken down the insights that the literature and cases have provided us into three categories which will serve as the basis for our recommendations for AFMC.

Design

- ➔ *map strategy against relevant metrics*
- ➔ *identify leading and lagging indicators and balance them against each other*
- ➔ *measure subjective metrics objectively with the proper evaluation criteria*
- ➔ *include softer, less technical strategic goals such as workforce development and retention*

- *do not include metrics just because they are easily available and/or measurable*
- *obtain buy-in on all levels, including for the number of metrics selected*
- *decide on a strategy first and metrics later, assign a metric champion who is not also the person responsible for the operations which the metric measures*
- *ask if the metric directly correlates to success regarding a KPQ (e.g. number of units sold is a better metric than number of units produced)*

Implementation

- *strive for cascading effects in multi-tiered organizations*
- *ensure that organizational and data structure do not generate an unmanageable number of metrics used*
- *be cautious implementing practices that have worked elsewhere without considering the fit with AFMC*

Maintenance

- *develop methods for evaluating what drives changes in metrics*
- *maintain flexibility to drop, add, or redefine metrics over time to ensure their relevancy*
- *ensure that tracking and trend analysis is possible*
- *consider causation; just because the organization does better/worse upon implementing the scorecard does not mean this was because of the scorecard*
- *consider time commitment; focus on a limited set of metrics*
- *develop a process which phases metrics in; too many new metrics at once may distort the result and make it difficult to discern the impact of any one metric*

Recommendations

After an analysis of the 95 metrics currently being used, a literature review and case study analysis, and considering the goals of AFMC, we have created a set of recommendations to address the concerns of AFMC listed above.

Optimal number of metrics

- ➔ *15 to 20 metrics reviewed on a monthly basis*
 - *10 predetermined metrics,*
 - *5 to 10 additional metrics “flagged” for review*
- ➔ *strategic in nature, to aid command-level decision making*

Discussion

Based on the level of expertise, we recommend using no more than 15 to 20 metrics at the command level. Our literature review suggested that people can handle seven pieces of new information before experiencing information overload. However, because of the familiarity of the information, the command staff can handle more chunks of information than seven. While we recommend starting at using this number of metrics, we suggest reevaluating this number (and the recommendations listed below) and adjust accordingly.

We utilized a checklist (see Appendix C) to decide which metrics should be given priority. Our evaluation of the respective metrics can be found in Appendix D. This provided us with a clearer picture regarding the suitability of the metrics. The next step was to balance this quantification of the metrics with a qualitative analysis based in the needs of AFMC and the priorities the Commander defined, ensuring all priorities would be given proper attention in our recommendations. We concluded that, through aggregation, it was possible to select ten metrics of particular importance. This process was vital in determining the best metrics. Our recommendations should be evaluated for effectiveness and fit after three months. If AFMC staff, through their knowledge and expertise, believes the following metrics are not the most effective mix for the Commander to review on a monthly basis, our process can be used to select more suitable metrics.

These metrics recommended below are strategic in nature and lend themselves to strategic decision-making. In addition to these metrics, any metric that is underperforming should be flagged for review (through a trend feature in the technology employed, or by the goal champion; discussion to follow). This system of flagging will increase the number of metrics for

monthly review up to the suggested 15 to 20 metrics. A detailed explanation of why we chose the following ten metrics can be found in Appendix E.

	Metric Name	Aggregated Metrics
1.	<i>Continue to Strengthen AFMC’s Support to the Nuclear Enterprise (AFNWC), Metric 1.1</i>	<i>1.1.1.1 through 1.1.1.5</i>
2.	<i>Advance Today’s & Tomorrow’s Combat Capabilities through Leading-Edge Science & Technology (AFRL), Metric 1.2</i>	<i>1.2.1.1 through 1.2.3.1, and 3.1.1.7</i>
3.	<i>Acquire & Support War-Winning Capabilities ‘Cradle-to-Grave’ (AFLCMC), Metric 1.3</i>	<i>1.3.1.1 through 1.3.2.2, and 3.1.1.1 through 3.1.1.3</i>
4.	<i>Perform World Class Test and Evaluation (AFTC), Metric 1.4</i>	<i>1.4.1.1 through 1.4.2.1</i>
5.	<i>Sustain AF Capabilities through World-Class Depot Maintenance & Supply Chain Management (AFSC), Metric 1.5</i>	<i>1.5.1.1 through 1.5.1.5, and 3.1.1.5 through 3.1.1.6</i>
6.	<i>Standardize and Continually Improve Processes</i>	<i>2.1.1.1 through 2.2.1.1</i>
7.	<i>Cost Effectiveness</i>	<i>1.6.1.1 through 1.6.1.3, 3.1.1.4, and 3.1.2.1 through 3.3.2.3</i>
8.	<i>Recruit, Develop, and Retain a Competent Workforce</i>	<i>4.1.1.1 through 4.2.3.2 and 4.4.1.1</i>
9.	<i>Secure and Improve Installations and Infrastructure</i>	<i>4.3.1.1 through 4.3.2.1 and 4.5.1.1 through 4.5.2.1</i>
10.	<i>Assess Health of Each ACS Functional Area & Advocate for Capability Needs, Metric 5.1</i>	<i>5.1.1.1 and 5.1.2.1</i>

The additional five to ten metrics that are flagged can be based on multiple “triggers:” historical trends, projected trends, goal champions, or the Commander. Goal champions should have the ability to flag metrics for review. They are the individuals with the deepest knowledge about any given metric and may be able to identify an issue before the system can. Although the use of technology is beyond the scope of our analysis, research shows that the presentation of information is vital for decision-making, especially when there is an excessive amount of information.³² Technology and presentation (trend tracking, metric presentation, status availability, and standardized traffic light criteria) are vital in reducing the risk of information overload.

³² Andrea Seaton Kelton and Robin R. Pennington, “Internet Financial Reporting: The Effects of Information Presentation Format and Content Differences on Investor Decision Making,” *Computers in Human Behavior*, 28 (2012).

Metrics should not be used just to flag negative performance, a view supported and articulated by the AFMC Commander.³³ Not only is it important to fix potential problems, it is important to recognize outstanding performance and implement best practices throughout the organization. For this reason, we also recommend that the system or goal champions flag outstanding performance. This can be in the form of consistently performing well above the “green” baseline, or for maintaining a positive trend over time.

Making the metrics more usable

- *standardize “traffic light” system*
- *standardize information available in the “description” section of the metric*
- *make status and trend information readily available*
- *develop a system used to store data in order to track and project trends*
- *group similar metrics together (for example cause and effect metrics)*
- *flag the overarching metric for review if any of the sub-metrics are underperforming; if one sub-metric is red, the main metric should also be red*

Discussion

While going through the metrics, we came up with additional recommendations to help strengthen AFMC’s PM system. Most AFMC metrics use a standard traffic light status system. Green indicates meeting standards; yellow is a metric that, while not yet underperforming, gives cause for concern; red means the metric is not meeting standards. However, the traffic light color coding system is not consistent throughout the organization. For example, some metrics use the standard green, yellow, red while others used green, yellow, orange and red. Some metrics even use different colors (such as blue) to indicate acceptable performance. Even with a deep understanding of the topics being discussed, this can create confusion and increase the chance of information overload. We recommend standardizing the color codes used, with each color communicating the same message for each metric (e.g. meets standards, track progress, and underperforming).

³³ General Janet C. Wolfenbarger, Harvard Kennedy School Harris Lecture Forum Event, November 6, 2014.

Based on the best practices, we also recommend that the description of each metric in the dashboard should include (1) what is being measured; (2) what is the purpose of measuring the metric (why is it being measured and how does it lend to decision making); (3) what is the baseline; and (4) what are the different “traffic light” criterion?

The status and trend of the metric should also be readily available. When talking to AFMC personnel, we realized that determining a trend may be a problem. The current dashboard is a real-time snapshot of the status of the organization. Data must be input into the dashboard, which erases the previous information. The trend is currently tracked with an arrow input by the goal champion. The trend is valuable information, and can be better tracked through an automatic trending function in the dashboard. Information for certain metrics should be collected more often and constantly updated in the system in order to have the most up-to-date information (Appendix D). Historical data should not be erased every time new data is uploaded.

We recognize that it may be difficult to change the programming of the system, but we believe it is necessary to track historical information to better determine trends. This will enable the system to flag a metric that is on a negative trend, even if it does not dip below the “green” criteria. Additionally, we believe that this would help make the metrics more anticipatory. While looking at historical data is informative of past performance, creating a dashboard that can project trends and can flag a metric before it becomes a problem may be even more powerful.

Lastly, in order to reduce the chance of information overload and maximize one’s cognitive ability, we recommend grouping similar metrics together. Each metric is linked to a specific command goal; however, some metrics appear to cut across multiple goals. For example, metric 1.5.1.3 states that manpower is red due to over-manning (more people working than was expected) due to production issues and increased orders. Metric 3.1.2.1, which falls under a different goal (cost effectiveness), points out that AFMC is over-budget on civilian personnel costs (Appendix F). This appears to be a cause and effect relationship, which is not evident in the dashboard. We recommend grouping related metrics such as these together, enabling a more informed decision. All 95 metrics contain important information that the Commander needs to

have access to. However, if similar metrics (for example 1.4.1.1 through 1.4.1.7, which all fall under 1.4.1) are all doing well, then the top-level metric should also be green. If any of the sub-metrics are underperforming, this should immediately trigger the top-level metric to be flagged for further investigation. The ten recommended metrics consist of multiple sub-metrics. This allows the command to have a general idea of performance if there is not a concern, but flags the metric if a concern does exist. At this point, the sub-metrics can be de-aggregated for review.

Conclusion

The potential for improvement through an improved approach to metric review, maintenance, and analysis at AFMC is large. With a \$55-60 billion budget, even minor performance improvements can result in significant cost savings or avoidance. Given the changing world dynamics and looming budget crisis, every cent saved makes a difference. Secretary of Defense Ash Carter emphasized the magnitude of the budget issue on March 18, 2015, stating that the DOD may need to start cutting military members' base pay to meet the budget.³⁴ By adopting the proposed improvements to their PM system, the visibility and transparency of the organization will increase and the process of identifying areas in need of attention will be simplified and sped up. Creating a more manageable dashboard will increase its usefulness.

³⁴ "SecDef Warns of Pay Cuts," Military Officers Association of America, (2015).
http://www.moaa.org/Main_Menu/Take_Action/Top_Issues/Serving_in_Uniform/Compensation/SecDef_Warns_of_Pay_Cuts.html.

Appendix A: AFMC Background

Air Force Materiel Command is the support center for the United States Air Force, responsible for equipping the Air Force with systems to face both present and future threats. It is one of nine Major Commands in the Air Force. Although AFMC only accounts for 13% of Air Force personnel (84,000 people), it is responsible for 40% of the Air Force budget (\$55-60 billion/year). The Air Force relies on AFMC to fulfill five core mission goals:

- develop leading-edge science and technology*
- perform “cradle-to-grave” life cycle weapon systems management*
- accomplish world-class developmental test and evaluation*
- accomplish world-class depot maintenance and supply chain management*
- strengthen the Air Force nuclear enterprise*

AFMC recently underwent an organizational restructure to accomplish its mission more efficiently and effectively. The four missions of AFMC (science and technology, life cycle management, test and evaluation, and sustainment) were aligned to standardize business practices, streamline processes and decision making authority, and provide opportunities for information sharing across the remaining AFMC centers.

This reduction alone saved \$100 million/year in personnel costs.³⁵ Each center is now responsible for only one mission with one primary commander responsible across all geographic locations. Additionally, in the first year, 66 processes were standardized across the organization and 41 instructions rescinded that were no longer necessary.³⁶ The first year after AFMC was reorganized, over \$6 billion was saved (through direct savings or cost avoidance).³⁷ Today, AFMC consists of five centers and 21 HQ AFMC staff directorates who directly support the Commander to execute the mission. Each one of the remaining centers aligns with one of AFMC’s five mission goals:

³⁵ General Janet C. Wolfenbarger, Harvard Kennedy School Harris Lecture Forum Event, November 6, 2014.

³⁶ Ibid.

³⁷ Ibid.

- ➔ *Air Force Research Laboratory (AFRL)*
 - *develop leading-edge science and technology*
- ➔ *Air Force Life Cycle Management Center (AFLCMC)*
 - *acquire and support war-winning capabilities “cradle-to-grave”*
- ➔ *Air Force Test Center (AFTC)*
 - *perform world-class test and evaluation*
- ➔ *Air Force Sustainment Center (AFSC)*
 - *achieve world-class sustainment*
- ➔ *Air Force Nuclear Weapons Center (AFNWC)*
 - *strengthen the nuclear enterprise*

In 2013, General Janet C. Wolfenbarger established a strategic plan to “shape and guide the command’s actions for the next three to five years.”³⁸ Her plan was influenced by the changing environment and fiscal constraints that the DOD faces. In 2012, a twenty-year Air Force Strategic Environment Assessment was conducted, and identified several environmental factors that impact strategic planning:

- ➔ *Potential adversaries (both state and non-state actors) are acquiring and developing the means to challenge the United States military.*
- ➔ *The number, and importance, of non-traditional operations such as asymmetric warfare, humanitarian operations, and urban operations is likely to increase.*
- ➔ *Deterrence is likely to become more challenging for the United States.*
- ➔ *The cost of energy will likely rise.*
- ➔ *New technology will create many opportunities, as well as potential obstacles.*

With the threat of sequestration and knowledge of the above environmental factors, General Wolfenbarger created a strategic plan with the intent to eliminate unnecessary spending and that was sustainable for the duration of her tenure and beyond. She identified five priorities that would align with the priorities of the Chief of Staff of the Air Force (win the fight, shape the

³⁸ *United States Air Force, Air Force Materiel Command Strategic Plan 2013, 2013.*

future, and strengthen the team). Each priority focuses on meeting the needs of the warfighter and is communicated through broad and continuing commitments for AFMC.

- expertly execute warfighter support
- ready and responsive agile combat support
- continuous process improvement
- cost effective mission execution
- workforce management and support

Each priority has at least one supporting goal to meet specific objectives and guide actions. The goals are then broken down further into the 95 metrics used by the AFMC commander. Figure 1 illustrates how the priorities outlined by General Wolfenbarger are aligned with both the Chief of Staff of the Air Force and the core mission goals of AFMC.

Figure 1: Priority Alignment



Appendix B: SMART

- *Specific: clear, as to avoid misinterpretation*
- *Measurable: can be quantified and compared to other data; allow for statistical analysis; avoid yes/no questions*
- *Accountable: the measure must be “owned” by a specific employee base*
- *Realistic: metric must be able to be achieved given the constraints of the organization*
- *Timely: achievable within the given time frame*

Appendix C: Metric Evaluation Questions

1. *Is the metric objectively measurable?*
2. *Does the metric help explain the efficiency or effectiveness of one or more of the stated goals/KPQs?*
3. *Does the metric allow for meaningful trend or statistical analysis?*
4. *Is it known (or at least can it be known) whether the metric is measuring a leading or lagging indicator?*
5. *Is there another metric that measures the same thing but better?*
6. *Does the metric include milestones and/or indicators to express qualitative criteria?*
7. *Is the metric not only measurable but also useful?*
8. *Are assumptions and definitions specified for what constitutes satisfactory performance?*
9. *Is someone assigned to monitor the metric? If so, what periodicity is used?*

Appendix D: Metric Evaluation

Metric	Definition	Measurable?	Standard Stoplight?	Baseline/Benchmark	Recommendations/Notes
1.1	<i>Under development—limited information Aggregate sub-metrics</i>				
1.1.1					
1.1.1.1					
1.1.1.2					
1.1.1.3					
1.1.1.4					
1.1.1.5					
1.2	Advance Today's & Tomorrow's Combat Capabilities through Leading-Edge Science & Technology (AFRL)				
1.2.1	Ensure the Air Force S&T Program addresses the highest priority capability needs of the Air Force (AFRL) <i>To meet goal 1.2, AFRL must deliver solutions to warfighter S&T needs that ensure winning capabilities in the near, mid, and far term.</i>				
1.2.1.1	High Profile Programs: % of High Profile Programs on track to meet cost, schedule, performance, and delivery. Terms of Reference: Green >80%, Yellow: 79%-60%, Red: <60%.	Yes	Yes	Stated in definition	Can you make measures more clear in which of the areas (cost, schedule, performance, delivery) are failing?
1.2.1.2	Rapid Innovation Programs: % of Rapid Innovation Programs on track to meet cost, schedule, performance, and delivery. Terms of Reference: Green >80%, Yellow: 79%-60%, Red: <60%.	Yes	Yes	Stated in definition	Can you make measures more clear in which of the areas (cost, schedule, performance, delivery) are failing?
1.2.2	Execute a balanced, integrated S&T Program that is responsive to Air Force Service Core Functions (AFRL) <i>To meet goal 1.2, AFRL must maintain a balanced S&T investment portfolio that responds to the AF Core Functions in a way that provides maximum value to the warfighter across the near, mid, and far term.</i>				
1.2.2.1	S&T Alignment with Customer Needs: Metric assesses customer satisfaction with the S&T program response to Core Function Support Plan Gaps and also the AF status with closing the Capability Gaps. Terms of Reference: Green - Satisfied/Executing, Yellow - Cautious/Planning, Orange - Apprehensive/Engaged, Red - Dissatisfied/Stalled.	Unclear	No	Stated in definition	Can you clarify what satisfied means? It is good to gather customer data, but must be in a standardized way so the customers are all reporting based on same definitions; also need to include the why they are satisfied or not. This is harder to trend.
1.2.2.2	Scientific Advisory Board – Relevance Assessment: % of SAB assessed Technical Areas that met or exceeded relevance expectations. Terms of Reference: Green- >80%, Yellow-Between 79% - 60%, Orange-Between 59% - 40%, Red-< 40%.	Difficult	No	Stated in definition	If there are clearly defined expectations, then this is measurable. Information on which technical areas are meeting expectations is not available.
1.2.2.3	Technology Transition: Leading and actionable measure of how well AFRL S&T Programs are positioned to transition to their intended customer. Measure is composed of key transition indicators (Documented S&T Need, Cost/Schedule/Performance, Transition Strategy, Customer Transition Funding & Interest). Terms of Reference: Green (>80), Yellow (79-70), Orange (69-60),	Difficult	No	Stated in definition	"Interest" and "strategy" are not clearly defined; may be difficult to measure.

	<i>Red (<60).</i>				
1.2.3	Shape the critical organizational competencies and resources needed to support the S&T program (AFRL) <i>To meet goal 1.2, AFRL must actively evolve its organization (research areas, equipment, facilities) to meet current and future warfighter needs.</i>				
1.2.3.1	Scientific Advisory Board – Quality Assessment: <i>% of SAB assessed Technical Areas that met or exceeded quality expectations for S&T research, equipment, facilities, and expertise. Terms of Reference: Green- >80%, Yellow-Between 79% - 60%, Orange-Between 59% - 40%, Red-< 40%.</i>	<i>Yes</i>	<i>No</i>	<i>Stated in definition</i>	<i>What are "quality expectations?"</i>
S&T Benefits Tracking Report	<i>The S&T Benefits Report provides insight into S&T Advances, Subject Matter Expert Consults, and unique Facility/Tool Uses. The benefits identified serve both current and future warfighters and customers.</i>	<i>No</i>	<i>No</i>	<i>Not stated in definition</i>	<i>Informational but too vague.</i>
1.3	Acquire & Support War-Winning Capabilities ‘Cradle-to-Grave’ (AFLCMC)				
1.3.1	Deliver timely and effective acquisition solutions (AFLCMC): <i>The AFLCMC product is to provide the warfighter's edge by acquiring and supporting war winning aircraft, engines, munitions, electronics, and cyber weapon systems and sub-systems while achieving cost efficiencies.</i>				
1.3.1.1	Acquisition Cost Variance to Baseline: <i>This metric assesses active AF Acquisition Categories (ACATs) I-III Acquisition Program Baseline (APBs) total acquisition cost compared to current estimate total acquisition cost.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Can you include absolute cost levels, rate of change, rate of change of rate of change (change programming to allow quick calculations)? Can you change to leading indicator?</i>
1.3.1.2	Acquisition Schedule Achievement: <i>This metric identifies the next upcoming milestone and determines the difference from objective to current estimate (in months).</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Same issue as above metric.</i>
1.3.1.3	Acquisition Delivery Achievement: <i>This metric assesses the number of programs where planned deliveries are equal to actual deliveries.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Why is there a difference between the schedule metric (1.3.1.2) and this metric?</i>
1.3.1.4	Acquisition Requirements Performance: <i>This metric assesses whether the project/program is meeting, or on track to meet, its technical performance goals, Technology Development Strategy (TDS) exit criteria, or APB performance parameters. This assessment should also identify any significant unplanned performance issue (including integration) that increases risk in any of the assessment areas or in operational suitability and effectiveness (e.g., engine failures, test failures, or software failures).</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Can these be weighted differently based on cost?</i>
1.3.2	Deliver affordable and effective product support (AFLCMC) <i>The AFLCMC product is to provide the warfighter's edge by acquiring and supporting war winning aircraft, engines, munitions, electronics, and cyber weapon systems and sub-systems while achieving cost efficiencies and improve warfighter product outcomes. Two metrics are measured to determine the affordability and effectiveness of product support: Logistics Health Assessment Compliance and System Availability.</i>				
1.3.2.1	Logistics Health Assessment Compliance: <i>This metric assesses a program's status relative to the 12 product support elements at the different life cycle phases. The LHA also provides a platform and</i>	<i>Vague</i>	<i>Yes</i>	<i>Not stated in definition</i>	

	<i>enterprise assessment roll up.</i>				
1.3.2.2	System Availability: <i>This metric measures system availability, number of hours an aircraft/equipment is available to perform its assigned missions.</i>	<i>Yes</i>	<i>No</i>	<i>Not stated in definition</i>	<i>How was the baseline established (why is baseline so low)?</i>
1.4	Perform World Class Test and Evaluation (AFTC)				
1.4.1	Provide credible and timely system performance information to decision makers (AFTC) <i>The primary product of the AFTC is information regarding the performance of a system-under-test (SUT). This information can be delivered to a customer in many different ways from a collection of unanalyzed test data to a fully analyzed and reported system performance. Additionally, to be timely, the final product must be delivered to a timeline that supports the customer's acquisition decision making processes. The objective then is to provide credible (quality) technical deliverables in a cost and schedule effective manner.</i>				
1.4.1.1	Test Project Schedule Effectiveness: <i>This metric measures the ability of the AFTC to meet test project schedule commitments outlined in a customer support agreement for projects that complete in the quarter in question. Projections for ongoing projects are also provided.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>What's the difference between the acquisitions category for schedule/cost/etc. and test project schedule/cost/etc.?</i>
1.4.1.2	Test Project Cost Effectiveness: <i>This metric presents trend information regarding the percentage of test projects that have met, or are over, cost estimates.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Can you show how much they go over by? Some of these look like they are measuring schedule, not cost (problem of double measuring).</i>
1.4.1.3	Technical Deliverables Timeliness: <i>This metric measures the ability of the AFTC to meet commitments outlined in a customer support agreement by delivering technical information on time.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	
1.4.1.4	Test Project Schedule Satisfaction: <i>This metric measures the subjective satisfaction of the customer regarding the schedule control during execution of the project through the use of responses to a standard questionnaire.</i>	<i>Difficult</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Valuable information, but subjective evaluation may present the problem of being told what you want to hear.</i>
1.4.1.5	Test Project Cost Satisfaction: <i>This metric measures the subjective satisfaction of the customer regarding the cost estimate and cost control during execution of the project through the use of responses to a standard questionnaire.</i>	<i>Difficult</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Valuable information, but subjective evaluation may present the problem of being told what you want to hear.</i>
1.4.1.6	Business Satisfaction: <i>This metric assesses responses on AFTC surveys, "Are we easy to do business with?" The calculations will be an annual accumulation of customer survey responses to the question. Site surveys are submitted to AFTC to consolidate the responses.</i>	<i>Difficult</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Valuable information, but subjective evaluation may present the problem of being told what you want to hear.</i>
1.4.1.7	Developmental Test Effectiveness – Deficiencies: <i>This metric provides one indicator of the effectiveness of Developmental Test and Evaluation by ascertaining and presenting the number of deficiencies that were found in Operational Test. Those deficiencies will then be adjudicated to determine if there is a reasonable expectation that the deficiency should have been found during developmental test. Those that should have been found but</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>First one to include a purpose for the metric (to drive DT&E process changes—identify problems earlier, identify roadblocks); all metrics should include a purpose.</i>

	<i>weren't are "escapes" or failures in the developmental test program.</i>				
1.4.2	Align Test & Evaluation infrastructure investment programs with requirements – today & tomorrow (AFTC) <i>This objective is designed to show the health of the AFTC Test and Evaluation (T&E) enterprise.</i>				
1.4.2.1	Capability Readiness: <i>This metric assesses the status of major mission areas with an assessment of the capabilities contained in each mission area.</i>	<i>Difficult</i>	<i>No</i>	<i>Not stated in definition</i>	<i>Can this be based on numbers/percentages?</i>
1.5	Sustain AF Capabilities through World-Class Depot Maintenance & Supply Chain Management (AFSC)				
1.5.1	Be a Reliable, Agile & Responsive Organization Focused on Achieving the Art of the Possible <i>The AFSC needs to consistently meet its customers' requirements today and in the future, as effectively as possible (reliable), while also being able to understand and adapt as these requirements change (agile & responsive). This is underpinned by the need for the AFSC commanders to focus on identifying and implementing the Art of the Possible goals/targets to drive corrective action for those areas where there is a failing to meet existing customer requirements, improving performance and being more effective to their customers. Whilst it is accepted that there are federal mandates the AFSC still needs to be a highly competitive sustainment organization, targeting new business and ensuring that its people, processes and resources are prepared to accept repatriated/new workload.</i>				
1.5.1.1	Weapon System Performance Dashboard: <i>Measures the Center's reliability performance for parts support to operational units and provides an overview of fleet status with an emphasis on AA and TNMCS. Summary of why a WS is not meeting the target TNMCS rate.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	
1.5.1.2	Develop an Integrated Workload Strategy & Plan: <i>Measure of the centers reliability performance to deliver a plan identifying improvements in the development of workload requirements for supply chain, DPEM, and Direct Cite customers that lead to a Depot Maintenance capability through a new Requirements Review and Depot Determination (R2D2) process.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>This is currently a lagging indicator. What is this trying to measure—what is actually being measured to determine status is unclear.</i>
1.5.1.3	Improve Accuracy and Integration of SC & MX Plans: <i>Ensure that the execution of the integrated plan developed from the R2D2 process in Objective 1.5.1.2 improves the accuracy of SC and MX workload plans.</i>	<i>Unclear</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>This is currently an unclear metric--what are you measuring and what is the baseline?</i>
1.5.1.4	First Pass Quality Performance: <i>Measure of the centers reliability performance for providing quality aircraft, quality components and quality engines.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	
1.5.1.5	Stratified Aircraft Production Performance: <i>Measure of the center's DDP and Flow Days reliability and responsiveness performance to meet aircraft production operations.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	
1.6	Execute Mission within AF/DOD/Statutory Limitations (FM)				
1.6.1	Ensure Compliance with Statutory Limitations <i>Captures AFMC's compliance with statutory limitations.</i>				
1.6.1.1	Travel Limitation (FM): <i>Measures travel funding execution which is limited to FY 2015 SAF limitation.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Including quantitative information may make this metric more usable and help identify areas for</i>

					<i>cost saving.</i>
1.6.1.2	Contract Services Limitation (FM): Measures monthly Center obligation execution for contracted services, excluding medical and purchase of goods, per the FY 2015 National Defense Authorization Act requirement to limit contracted services to stated levels.	Yes	No	Not stated in definition	Statutory problems may be better reported in a different venue; may also make sense to keep the red/green status instead of the standard stoplight indicator for this type of metric.
1.6.1.3	Small Business Execution (SB): Measures actual small business obligations compared to AFMC goals.	Yes	No	Not stated in definition	
2.1	Standardize Critical AFMC Processes and Train the Workforce (A8/9)				
2.1.1	Implement Standardized Methodology for Critical AFMC Processes (A8/9): This objective ensures critical processes are standardized and a repeatable methodology is developed to identify and standardize new critical processes. Standardization will be achieved by process owners complying with a checklist that ensures processes are properly documented, trained, and codified.				
2.1.1.1	Progress in Relation to Schedule for Standardize AFMC Mission Processes: Each process owner will develop a schedule to complete the standardization checklist. Progress will be reported in relation to completion per the checklist 30, 60 and 90 days behind schedule thresholds will be established for Yellow, Orange, and Red status, respectively. This metric will be briefed through the AFMC Strategic Plan process. This metric measures the number of standardized processes across AFMC.	Yes	No	Stated in Definition	Can you use the stoplight system? This may be harder to trend, but it is important to know.
2.1.2	Validate AFMC Instruction Portfolio and Align to 5-Center Construct (A8/9): Ensure the AFMC instruction portfolio consists of only value added instructions. Alignment to the 5-center construct (revisions, rewrites, new guidance, etc.) will occur as process owners work through standardizing critical processes.				
2.1.2.1	Progress in Relation to Schedule to Eliminate Non-value Added Instructions: Policy Owners tasked to submit schedules to eliminate non-value added instructions. Thirty, sixty, and ninety day behind schedule thresholds will be established for Yellow, Orange, and Red status, respectively. This metric will be briefed through the AFMC Strategic Plan process.	Yes	No	Stated in definition	Can you track this over a shorter period?
2.2	Continuously Improve Critical AFMC Processes (A8/9)				
2.2.1	Improve AFMC mission execution through CPI and an innovative culture (A8/9) Improve mission execution on critical processes through the application of Continuous Process Improvement efforts to meet established improvement goals/targets and assess the maturity of our innovative culture to achieve the "art of the possible".				
2.2.1.1	Airmen Powered by Innovation (API) Suggestions: NO DEFINITION	N/A	No	Not stated in definition	
2.3	Enforce Standard Processes				
2.3.1	CC's Ensure Strategic Alignment (IG) This objective compares unit performance under Unit Effectiveness Inspection (UEI) Major Graded Area 3; Improving the Unit, Sub-MGA Strategic Alignment as measured by IG and Center CCs. IG-level metric validates internal score and provides cross-center focus. Center-level metric is focused internally on center and subordinate units. It is based on a 2-year rolling average. This metric will assess how well the organization achieves Strategic Alignment based on a 5-tier Adjectival Scale (Outstanding, Highly Effective, Effective, Marginally Effective and Ineffective).				

2.3.1.1	IG's Assessment of Strategic Alignment: <i>This metric depicts the Inspector General's assessment of Strategic Alignment. Strategic Alignment requires Commanders to strive for strategic alignment within their organizations. This includes aligning authorities with mission requirements. Vision and mission statements should lead to strategic plans that include yearly calendars and annual budgets.</i>	Unclear	Yes	Not stated in definition	<i>This may be better reviewed in a different venue; if not, it may be useful to include more quantitative data. Subjective measures are difficult to use.</i>
2.3.1.2	CC's Assessment of Strategic Alignment: <i>Center-level metric focuses internally. CCIP average of Center CC's assessments of Sub-Major Graded Area (MGA) 3.2. (Improving the Unit, Strategic Alignment) based on Commander's Inspection Reports (CCIRs).</i>	Unclear	Yes	Not stated in definition	<i>This may be better addressed in a different venue; this is a Center-level metric and may not need AFMC CC's review unless Center CC suggests a review or a center is under-performing.</i>
2.3.2	CC's Ensure Standardized Process Operations (IG) <i>This objective compares unit performance under Unit Effectiveness Inspection (UEI) Major Graded Area 3; Improving the Unit, Sub-MGA Process Operations as measured by IG and Center CCs. IG-level metric validates internal scores and provides cross-center focus. Center-level metric is focused internally on center and subordinate units. It is based on a 2-year rolling average. This metric will assess how well the organization achieves Process Operations based on a 5-tier Adjectival Scale (Outstanding, Highly Effective, Effective, Marginally Effective and Ineffective).</i>				
2.3.2.1	IG's Assessment of Process Operations: <i>This metric depicts the Inspector General's assessment of Process Operations. Process Operations requires leaders to be aware of critical processes, and to constantly seek to improve and standardize those processes to produce more reliable results. Leaders will seek to remove bottle-necks or limiting factors and ensure risk management principles are applied during daily operations. All risks, including safety and risks to personnel, should be considered when analyzing and improving processes.</i>	Unclear	Yes	Not stated in definition	Same issue as 2.3.1.1 and 2.3.1.2
2.3.2.2	CC's Assessment of Process Operations: <i>Center-level metric focuses internally. CCIP average of Center CC's assessments of Sub-Major Graded Area (MGA) 3.2. (Improving the Unit, Process Operations) based on Commander's Inspection Reports (CCIRs).</i>	Unclear	Yes	Not stated in definition	Same issue as 2.3.1.1 and 2.3.1.2
2.3.3	CCs Ensure an Effective CC's Inspection Program (IG) <i>This objective compares unit performance under Unit Effectiveness Inspection (UEI) Major Graded Area 3; Improving the Unit, Sub-MGA Commander's Inspection Program (CCIP) as measured by IG and Center CCs. IG-level metric validates internal scores and provides cross-center focus. Center-level metric is focused internally on center and subordinate units. It is based on a 2-year rolling average. This metric will assess how well the organization achieves a CCIP based on a 5-tier Adjectival Scale (Outstanding, Highly Effective, Effective, Marginally Effective and Ineffective).</i>				
2.3.3.1	IG's Assessment of CC's Inspection Program: <i>This metric depicts the IG's assessment of the Commander's Inspection Program (CCIP). CCIP requires Commanders to have the legal authority and responsibility to inspect their subordinates and subordinate units. A robust CCIP finds deficiencies and improves mission readiness. Part of this effort must be a Unit Self-Assessment Program where individual Airmen report their compliance with guidance. An independent verification of those</i>	Unclear	Yes	Not stated in definition	Same issue as 2.3.1.1 and 2.3.1.2

	<i>reports provides commanders with additional confidence in their validity.</i>				
2.3.3.2	CC's Assessment of CC's Inspection Program: <i>Center-level metric focuses internally. CCIP average of Center CC's assessments of Sub-Major Graded Area (MGA) 3.3. (Improving the Unit, CCIP) based on Commander's Inspection Reports (CCIRs).</i>	<i>Unclear</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Same issue as 2.3.1.1 and 2.3.1.2</i>
2.3.4	CC's Stan/Eval Prgs Accurately Assess Aircrew Performance and Standardized Air Operations (A3) <i>This objective assesses the Stan/Eval Program Team's ability to accurately assess aircrew performance and standardized air operations. AFMC/A3 provides AFMC/IG validation during Aircrew Performance Evaluation (APE) inspections.</i>				
2.3.4.1	Average Local Standardized/Evaluation (Stan/Eval) Checkride Results (A3): <i>This metric averages local Stan/Eval checkride results on group-equivalent units up to the center level.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Same issue as 2.3.1.1 and 2.3.1.2</i>
2.3.4.2	Average Aircrew Performance Evaluation (APE) Results (A3): <i>This metric averages unit APE Ratings up to the center level and provides comparative analyses of average local Stan/Eval checkride and APE results.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Same issue as 2.3.1.1 and 2.3.1.2</i>
3.1	Demonstrate Cost Effective Mission Execution (FM)				
3.1.1	Deliver cost effective mission execution <i>Ensures the cost-effective application of resources (e.g., budget, manning, capacity, and investments) by the AFMC Centers to execute the AFMC mission.</i>				
3.1.1.1	Program Acquisition Unit Cost (PAUC) (AFLCMC): <i>This metric measures current PAUC cost performance compared to the Acquisition Program Baseline (APB).</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>What is the difference between this and 1.3.1.1? Can they be combined?</i>
3.1.1.2	Acquisition Should Cost (AFLCMC): <i>This metric reflects the realized/projected "should cost" savings which result from the implementation of cost reduction initiatives into all ACAT I/II/III programs throughout program execution, including product support. It compares the realized savings to a realization forecast to measure success.</i>	<i>Vague</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>May be useful to include more quantitative data for this metric.</i>
3.1.1.3	Development Planning Return on Investment (ROI) (AFLCMC): <i>This metric shows linkages and payoffs between S&T investments, core function master plans (CFMPs), and planning for new programs. ROI is calculated by dividing estimated cost avoidance by the cost of the development planning effort.</i>	<i>Yes</i>	<i>No</i>	<i>Not stated in definition</i>	<i>May decrease chance of information overload if the standard stoplight indicator was used. Good goal to base new programs on; what is the baseline and why?</i>
3.1.1.4	Cost Effectiveness Through Competition (PK): <i>This metric measures the level of competition for the command. It reports obligations completed through competition efforts versus total obligations for the fiscal year.</i>	<i>Under Development</i>			
3.1.1.5	CSAG-M Should Cost (AFSC): <i>Measures the Consolidated Sustainment Activity Group - Maintenance (CSAG-M) actual cost of what has been produced to date against should cost, based on</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Why are the baselines set the way they are? Why not flag anything >=0%? Is there a way to make this a leading indicator?</i>

	<i>the earned hours (production) times budgeted rates and fixed overhead.</i>				
3.1.1.6	CSAG-S Expense (AFSC): Measures the Consolidated Sustainment Activity Group - Supply (CSAG-S) actual cost of material and overhead expenses against planned cost.	Yes	Yes	Not stated in definition	
3.1.1.7	S&T Efficiencies (AFRL): Measures AFRL actions to reduce "tail" or support functions and reinvest those resources to "tooth" or scientific research. The goal set for AFRL is to save/reinvest \$148.6M across the FY13-16. This metric is already being reviewed monthly through AFMC/A8 and SAF/AQ as a strategic metric (Objective E4 is the tracking number by SAF).	Yes	Yes	Not stated in definition (but is stated in status details)	Although the threshold is established by SAF, it may be beneficial to set "stretch targets" (see glossary)
3.1.2	Deliver cost effective functional execution (FM) Ensures the cost-effective application of resources (e.g., budget, manning, capacity, and investments) by the MAJCOM functionals to execute the AFMC mission.				
3.1.2.1	Civilian Workyear Execution (O&M and RDT&E) (A1): Measures projected civilian pay and workyear execution to actual execution.	Yes	Yes	Not stated in definition	This may inform strategic decisions.
3.1.2.2	Program Office Support Funding (FM): Measures FY15 funded PMA against the FY15 PMA baseline. Funded PMA per Program may vary based on Center Commander Discretion.	Yes	Yes	Not stated in definition	This may inform strategic decisions.
3.1.2.3	Inspector Cost Averages (IG): Measures the average cost of an inspector per inspection week using planned costs versus actual costs.	Yes	Yes	Not stated in definition	This may reveal options for cost savings. The presentation slides already suggest cost reducers; this is good and an example of the value metrics can add.
3.1.2.4	AFMC Current for Canceled Invoices Outstanding: Metric identifies the loss of current year funding used to pay canceled year invoices.	Yes	Yes	Not stated in definition	This may inform strategic decisions.
3.2	Achieve and Maintain Financial Accountability/Auditability (FM)				
3.2.1	Achieve/sustain financial statement auditability in support of an AF unqualified opinion (FM) Receiving a clean (unqualified) opinion on external audits of the AFMC portion of Air Force financial statements.				
3.2.1.1	Budgetary Resources (FM): Measures the audit readiness activities supported by AFMC, in support of the Air Force audit readiness goal for the Statement of Budgetary Resources (SBR) per National Defense Authorization Act (NDAA) 2013. The NDAA goal is to achieve audit readiness by the end of fiscal year 2014.	Yes	No	Not stated in definition	This may inform strategic decisions.
3.2.1.2	Asset Accountability (A4): Measures the audit readiness activities supported by AFMC, in support of the Air Force audit readiness goal for Mission Critical Assets (MCA) accountability as directed by the Under Secretary of the Air Force (USECAF). The USECAF goal is to achieve audit readiness by the end of calendar year 2015.	Yes	No	Not stated in definition	This may inform strategic decisions.

3.2.1.3	Full Audit (FM): Measures the audit readiness activities supported by AFMC, in support of the Air Force audit readiness goal for the full set of financial statements per the National Defense Authorization Act (NDAA) 2010. The NDAA goal is to achieve audit readiness by the end of fiscal year 2017.	Yes	No	Not stated in definition	This may inform strategic decisions.
3.2.1.4	IT Systems Compliance (A6): Measures the status of achieving audit readiness IAW 10 USC 2222 for AFMC owned or operated financial/financial feeder Defense Business Systems (DBS).	Under Development			
3.3	Achieve Efficiencies in Energy Use (A6/7)				
3.3.1	Achieve Compliance with Federal and Executive Order Mandates (A6/7): This metric measures AFMC energy use compliance, showing past, current and projected consumption, it also discusses measures to the meet the 3% reduction goal, and risks to meeting the compliance standard. This metric measures AFMC energy use compliance, showing past, current and projected consumption, it also discusses measures to the meet the 3% reduction goal, and risks to meeting the compliance standard. This metric measures AFMC water use compliance, showing past, current and projected consumption, it also discusses measures to the meet the reduction goal, and risks to meeting the compliance standard. This metric measures AFMC Renewable Energy Compliance, showing past, current and projected results of renewable energy projects.				
3.3.1.1	AFMC Energy Use Intensity: This metric measures AFMC energy use mandate compliance, showing past, current and projected consumption.	Yes	No	Not stated in definition	If information can be obtained more often, this may benefit the metrics because this information can influence strategic decisions.
3.3.1.2	AFMC Water Use Intensity Compliance: This metric measures AFMC water use mandate compliance, showing past, current and projected consumption.	Yes	Yes	Not stated in definition	If information can be obtained more often, this may benefit the metrics because this information can influence strategic decisions.
3.3.1.3	AFMC Renewable Energy Compliance: This metric measures AFMC renewable energy mandate compliance, showing past, current and projected results of renewable energy projects.	Yes	Yes	Not stated in definition	If information can be obtained more often, this may benefit the metrics because this information can influence strategic decisions. It appears as though old data is being used as 2013 is still an "estimate."
3.3.2	Achieve efficiencies in fuel usage (A3 & A4) A4 general purpose fuel is green and awaiting A3 data on aviation fuel				
3.3.2.1	Achieve efficiencies in fuel usage; AFMC petroleum reduction (A4): This metric captures the reduction of fuel use (vehicle/aviation) by the Command. This goal stems from Executive Order 3423 and the Energy Independence Act.	Yes	No	Not stated in definition	If information can be obtained more often, this may benefit the metrics because this information can influence strategic decisions.
3.3.2.2	Achieve efficiencies in fuel usage; AFMC Alternative Fuel Consumption (A4): This metric assesses AFMC's ability to comply with the Energy Independence and Security Act of 2007 and Executive order 13423, strengthening federal environmental, energy, and transportation management.	Yes	No	Not stated in definition	If information can be obtained more often, this may benefit the metrics because this information can influence strategic decisions.
3.3.2.3	AFMC Aviation Fuel Efficiency (A3): This metric assesses AFMC's ability to improve aviation fuel efficiency.	Yes	No	Not stated in definition	Could benefit from more quantitative information to set the baseline instead of subjective terms such as "likely," and "questionable." If information can be obtained

					<i>more often, this may benefit the metrics because this information can influence strategic decisions.</i>
4.1	Recruit, Develop and Retain a Diverse and Competent Workforce (A1)				
4.1.1	Manage occupations, positions and competencies to meet mission requirements (A1) <i>This objective ensures that the commanders are provided the workforce required to perform their missions. The workforce should be of the requisite size and makeup, and should be competent in the performance of their duties.</i>				
4.1.1.1	APDP Certification Rates (A1): <i>This metric measures the percentage of personnel on Key Leadership Positions (KLPs) who are certified. Individuals filling a KLP must meet the mandatory level III certification for the career field the KLP is assigned within the Grace Period Expiration (GPE), which is 24 months from the time of assignment.</i>	Yes	Yes	Not stated in definition	
4.1.1.2	Fill Rates (Civilian Mission Critical Occupations) (A1): <i>This metric measures if we are hiring the people we need to complete AFMC's mission. There are nine occupations of interest (mission critical occupations), or top series/jobs. These occupations are directly associated with the primary mission of the Command and Center without which mission-critical work cannot be completed. The nine occupations are: aircraft maintenance, contracting, director, engineering, finance/cost, scientist, munitions and maintenance, program manager, and logistics readiness.</i>	Yes	Yes	Not stated in definition	<i>Is it possible to turn this into a leading indicator and project?</i>
4.1.1.3	Military Officer Assignments Equity (A1): <i>Metric measures AFMC 62E (Developmental Engineer), 63A (Acquisition Manager), 11X (Pilot), 12X (Navigator), 21A (Aircraft Maintenance), 21M (Munitions and Missile Maintenance), and 21R (Logistics Readiness) manning rates.</i>	Yes	Yes	Not stated in definition	<i>Is it possible to turn this into a leading indicator and project?</i>
4.1.1.4	Enlisted Assignments (A1): <i>Metric measures enlisted manning rates. Total manning rates should be within 5% of AF average. Skill level manning should be no less than 15% of AF average. Authorized are funded authorizations. Assigned are personnel billeted against those funded authorizations. Data source is MilPDS. Thresholds are the following: green <= 5% difference, yellow 6-15% difference and red >15% difference. Data is collected and reviewed (prior to AFPC matching assignments) at the beginning of each assignment cycle.</i>	Yes	Yes	Stated in definition	<i>Is it possible to turn this into a leading indicator and project?</i>
4.1.1.5	SDE Completions for Civilian Senior Leaders (A1): <i>The metric measures SDE completion status for GS-15s and equivalents at the time of their promotion.</i>	Yes	No	Not stated in definition	<i>May be able to make this more usable with a projection.</i>
4.1.1.6	Officer Development (AAD) (A1): <i>The metric measures Advanced Academic Degree (AAD) completion rates for Second</i>	Yes	Yes	Not stated in definition	<i>You may benefit from using a stretch target here, instead of the rest of the AF as the baseline.</i>

	<i>Lieutenant through Colonel, compared to AF-wide statistics.</i>				
4.1.1.7	Enlisted Development (PME completions) (A1): <i>This metric measures PME completion rates for enlisted personnel eligible for each level of PME. Enlisted PME includes Airman Leadership School (ALS), the Noncommissioned Officer Academy (NCOA), and the Senior Noncommissioned Officer Academy (SNCOA), compared to AF-wide statistics.</i>	Yes	Yes	Not stated in definition	<i>You may benefit from using a stretch target here, instead of the rest of the AF as the baseline.</i>
4.1.1.8	Enlisted Education (CCAF Degrees) (A1): <i>This metric measures the Community College of the Air Force (CCAF) completion rates for personnel within each enlisted tier (E1-E4, E5-E6, and E7-E9) compared to AF-wide statistics.</i>	Yes	Yes	Not stated in definition	<i>You may benefit from using a stretch target here, instead of the rest of the AF as the baseline.</i>
4.1.1.9	Mandatory Supervisory Training Completions (A1): <i>This metric measures the number of civilian first-level supervisors (assigned 180 days or more) who have completed AF Mandatory Supervisory Training (MST) IAW AFI 36-401. MST consists of three courses: USAF Supervisory Course (USAFSC), Civilian Personnel Management Course (CPMC), and Military Personnel Management Course (MPMC).</i>	Yes	No	Not stated in definition	
4.1.2	Advocate & Encourage a Diverse & Inclusive AFMC Workforce (A1) <i>This objective provides a platform for AFMC leadership, at all levels, to promote and strengthen an AFMC culture that values inclusion of all personnel and views diversity as a force multiplier. It monitors the AFMC workplace climate to identify barriers that could prevent personnel from achieving their full potential. Diversity includes, but is not limited to, personal life experiences, geographic and socioeconomic background, cultural knowledge, educational background, work background, language abilities, physical abilities, philosophical/spiritual perspectives, age, race, ethnicity, and gender.</i>				
4.1.2.1	Workforce Diversity (A1): <i>This metric reflects AF and AFMC demographics for gender, race, age, and education levels/types for officers, enlisted, and civilians.</i>	Yes	No	No target	<i>This may be more valuable to look at in another venue; does it lend itself to strategic decisions?</i>
4.2	Enhance the Wellness and Safety of the Workforce & their Families				
4.2.1	Implement & market Comprehensive Airman Fitness (A1) <i>This objective monitors the implementation of the CAF across AFMC. CAF is an overarching philosophy (not a program) for taking care of people. It provides a framework through which the Air Force can deliver relevant programs and services more effectively across the four pillars of fitness (Physical, Social, Mental, Spiritual) ultimately improving well-being, enhancing life balance, and strengthening personal and organizational resilience in Airmen and their Families. CAF begins and ends with leadership at all levels supported by helping agencies across functional communities.</i>				
4.2.1.1	Airmen Fitness Rates (A1): <i>This metric presents status of the military fitness test for AFMC officers and enlisted personnel compared to AF-wide statistics.</i>	Yes	No	Not stated in definition	<i>This is informational, but should this be handled at a lower level? It can be flagged if a problem arises, or is projected to arise.</i>
4.2.1.2	Sexual Assault Reporting (A1): <i>This metric measures the number of sexual assault victim reports at AFMC installations.</i>	Yes	No	No target	<i>This is informational, but should it be handled at a lower level? Does it lend itself to strategic decision making?</i>
4.2.1.3	Active Duty and Civilian Deaths by Suicide (SG): <i>This metric shows the number of deaths by suicide per 100,000 people in</i>	Yes	No	No target	

	<i>AFMC over a calendar year compared to AF-wide statistics.</i>				
4.2.2	Promote principles of healthy living (SG)				
4.2.2.1	Individual Medical Readiness (SG): <i>This metric assesses the Individual Medical Readiness (IMR) compliance rate in five medical areas (immunizations, dental exam, preventive health assessment, medical laboratory, and medical equipment), as well as not having a duty limiting condition. IMR monitoring allows commanders and their medical support providers to monitor the medical readiness status of unit personnel, ensuring a healthy and fit fighting force, medically ready to deploy.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Same issue as above</i>
4.2.2.2	Civilian Health (SG): <i>This metric contains 3 sets of data. (1) Self-reported health risk data from civilians voluntarily submitting Health Risk Assessments and healthy behavior/class attendance as part of AFMC's Civilian Health Promotion Services (2) AF Safety Automated System (AFSAS) data on occupational illnesses reported for AFMC's civilians and (3) Lost Time Case and Lost Duty Days rates per 100 civilians for AFMC civilians.</i>	<i>Yes</i>	<i>No</i>	<i>No target identified</i>	<i>Same issue as above</i>
4.2.3	Reduce Mishaps (SE)				
4.2.3.1	5-yr Average Class C Mishaps (SE): <i>This metric measures the 5-yr rolling average of on and off duty Ground Class C mishaps within AFMC. Class C mishaps are safety mishaps (1) costing less than \$500K and more than \$50K in property damage or (2) any injury, illness or disease that causes one or more loss work days. The goal is to have an ever decreasing 5-yr rolling average. Any measurable increase in the rolling average meets the Red threshold; a zero to two percent decrease in the rolling average meets the Orange threshold; a two to five percent decrease in the rolling average meets the Yellow threshold; and a greater than five percent decrease in the rolling average meets the Green threshold. The metric data is internal to AFMC, is updated quarterly, and reported to the HQ AFMC ESOH Council semiannually. Data used to build the metric is reportable Air Force wide in the Air Force Safety Automated System.</i>	<i>Yes</i>	<i>No</i>	<i>Stated in definition, but buried in words and hard to identify</i>	
4.2.3.2	On-duty Class A Mishaps (SE): <i>This metric measures the number of on duty ground, weapons and flight Class A mishaps. Class A mishaps are safety mishaps (1) costing more than \$2M or (2) fatal or permanent total disability, or (3) destruction of a DOD aircraft. Note a destroyed UAV/UAS is not a Class A mishap unless the criteria in (1) or (2) is meet. The goal is to have no Class A mishaps. Four or more on-duty mishaps meet the "Red" threshold;</i>	<i>Yes</i>	<i>No</i>	<i>Stated in definition, but buried in words and hard to identify</i>	

	<i>three on-duty mishaps meet the "Orange" threshold; two on-duty mishaps meet the "Yellow" threshold; and zero to one on-duty mishaps meet the "Green" threshold. The metric data is internal to AFMC, is updated quarterly, and reported to the HQ AFMC ESOH Council semiannually. Data used to build the metric is reportable Air Force wide in the Air Force Safety Automated System.</i>				
4.3	Protect & Secure AFMC Installations and Sites (A6/7)				
4.3.1	Provide First Responder services and Installation Security IAW Air Force standards (A6/7) <i>This metric uses the National Incident Management system as the structure to measure our ability to provide first responder Installation security services to the Command.</i>				
4.3.1.1	Incident Management and Response: <i>This metric is the strategic roll up of 5 functional areas (command, operations, logistics, planning and administration/finance) which cover Incident Management and Response under the National Incident Management System (NIMS) construct.</i>	<i>Difficult</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Useful information, but may benefit from more quantitative information/approach.</i>
4.3.2	Prevent the compromise, loss, unauthorized access/disclosure of sensitive, controlled unclassified, & classified information (IP) <i>This objective seeks to increase employee awareness of what information needs to be protected, why it needs to be protected, and how to protect it. The strength of an IP-aware corporate culture is measured by tracking security incidents, with a focus on eliminating compromises, losses, and repeat violations, whether at the individual or unit level.</i>				
4.3.2.1	Measures the number of security incidents which occur in AFMC: <i>This metric measures the number of security incidents which occur in AFMC.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Why is this the baseline?</i>
4.4	Deploy Fully Trained & Ready Personnel (A3)				
4.4.1	Meet AEF Deployment Standards (A3)				
4.4.1.1	AEF Execution Focus Areas: <i>Purpose of metric is to assess the commander's performance in getting their deploying Airmen to their final destination on time, with the required equipment, and with all their deployment requirements completed. Grading criteria is the percentage of deployers who have zero mission impact discrepancies.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	
4.5	Champion Infrastructure & Services for our Workforce & Families (A6/7)				
4.5.1	Ensure Installation Support Services Provided IAW AF Standards (A6/7)				
4.5.1.1	Provide Base Support Vehicles and Equipment: <i>This metric is a roll up 11 areas of vehicle mission capable rates. The mission capable rate is the percentage of assigned vehicles in operational service.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>This may benefit from more quantitative information for metric evaluation.</i>
4.5.1.2	Provide Quality Fuels Support: <i>This metric is a strategic roll up of two functional areas (quality and quantity) of both ground and aviation fuel.</i>	<i>Yes</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>Is this related to the previous fuel metrics? Can they be aggregated?</i>
4.5.1.3	Mission Support: <i>This metric is a roll-up of key mission support mission areas (e.g. security, communications, utilities, airfield</i>	<i>Difficult</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>This may benefit from more quantitative information for metric evaluation.</i>

	<i>management, real estate).</i>				
4.5.1.4	Provide Food Services: <i>This metric measures the provision subsistence to essential station messing and supports to personnel receiving basic allowance for subsistence.</i>	<i>Yes</i>	<i>No</i>	<i>Not stated in definition</i>	
4.5.1.5	Provide Child and Youth Services: <i>This metric measures certifications and inspections requirements mandated by the Military Child Care Act of 1989.</i>	<i>Yes</i>	<i>No</i>	<i>Not stated in definition</i>	
4.5.1.6	Provide MWR-Core (Fitness): <i>This metric assesses the availability/accessibility of AFMC's fitness programs against the Air Force standard of being open/accessible 112 hours per week.</i>	<i>Yes</i>	<i>No</i>	<i>Stated in definition</i>	
4.5.2	Ensure Installation Infrastructure Provided IAW AF Standards (A6/7) <i>This objective ensures AFMC's installation infrastructure is provided IAW AF standards.</i>				
4.5.1.1	Infrastructure and Facility Sustainment: <i>This metric measures a roll up of functional areas (Acquisition, Force Support, Contracting, Judge Advocate, Science and Technology, Airfield Operations, Civil Engineering, Safety, Distribution, Health Services, Materiel Management, Chaplain Corps, Logistic Plans, Financial Management, Munitions, Historian, Maintenance, Intelligence, Security Forces, Test and Evaluation, Inspector General, Mission Assurance, Air Force Office of Special Investigations, and Public Affairs) which, collectively, provide comprehensive infrastructure life-cycle management (acquire, sustain, dispose).</i>	<i>Difficult</i>	<i>Yes</i>	<i>Not stated in definition</i>	<i>It is unclear what is being measured.</i>
5.1	Assess the Health of Each ACS Functional Area & Advocate for Capability Needs (A8/9)				
5.1.1	Identify, assess, and report the functional health of ACS through a collaborative process (A8/9) <i>This objective is focused on the health of the Agile Combat Support (ACS) service core function when viewed from a functional community perspective. Functional community health is hidden when ACS is viewed from a capability perspective. ACS is comprised of 24 functional communities: Acquisition (SAF/AQ), Force Support (AF/AI), Contracting (SAF/AQ), Judge Advocate (AF/JA), Science/Tech(SAF/AQ), Airfield Ops (AF/A3/5), CE (AF/A4/7), Safety (AF/SE), Distribution (AF/A4/7), Health Services (AF/SG), Materiel Management (AF/A4/7), Chaplain Corps (AF/HC), Logistics Plans (AF/A4/7), FM (SAF/FM), Munitions (AF/A4/7), Historian (AF/HO), Maintenance (AF/A4/7), Intelligence (AF/A2), Security Forces (AF/A4/7), T&E (AF/TE), IG (SAF/IG), Mission Assurance (SAF/AA), AFOSI (SAF/IG), and Public Affairs (SAF/PA).</i>				
5.1.1.1	Functional Health Assessment Results: <i>This metric reports overall risk for each of the 24 ACS functional communities (see Objective 5.1.1 for a list of the functional communities). The risk is extracted from the most recent functional health assessment, ACS planning data call results, and Core Function Support Plan. In addition to overall risk, mitigation actions taken to reduce overall risk will be tracked and reported semi-annually.</i>			<i>Restricted Information</i>	
5.1.2	Identify, assess & report ACS Core Capability health through a collaborative process (A8/9) <i>This objective is focused on the health of the Agile Combat Support (ACS) service core function when viewed from a capability perspective. ACS is comprised of 5 core capabilities: Field, Base, Protect, Support and Sustain. Capabilities are what ACS deliveries to the warfighter.</i>				

5.1.2.1	Core Capability Risk Assessment Results: <i>This metric reports overall risk for each of the 5 ACS core capabilities (Field, Base, Protect, Support and Sustain). The risk is extracted from the most recent Core Function Support Plan. In addition to overall risk, mitigation actions taken to reduce overall risk will be tracked and reported semi-annually.</i>	<i>Restricted Information</i>
----------------	--	-------------------------------

Appendix E: Selection Discussion

Aggregate 1.1.1.1 through 1.1.1.5 into 1.1.

Aggregate 1.2.1.1 through 1.2.3.1, and 3.1.1.7 into 1.2.

Aggregate 1.3.1.1 through 1.3.2.2, and 3.1.1.1 through 3.1.1.3 into 1.3.

Aggregate 1.4.1.1 through 1.4.2.1 into 1.4.

Aggregate 1.5.1.1 through 1.5.1.5, and 3.1.1.5 through 3.1.1.6 into 1.5.

Each one of the sub-metrics (1.1.1.1 through 1.5.1.5) should be managed at the respective center. If each center is managing their performance well and resolving issues before they become a major issue, there is no need to spend an excessive amount of time on each sub-metric. If any of the sub-metrics are underperforming, or of concern to their goal champion, they can be flagged for review. The additional metrics from cost effectiveness (3.1.1.X) that we recommend aggregating are directly tied to mission execution in the respective centers. If the centers are accomplishing their respective missions, but not doing so cost effectively, the center should be reviewed. This also incorporates our recommendation to aggregate cause-and-effect relationships.

Aggregate 2.1.1.1 through 2.2.1.1 into one metric *that communicates the standardization and improvement of processes. We recommend tracking 2.1.2.1 over a shorter time period if possible. We suggest this as a metric since it ties directly to the strategic plan outlined established in 2013. Process standardization and improvement has saved, and has the potential to save additional, money. The cost savings should be tracked and provided to Congress.*

Aggregate 1.6.1.1 through 1.6.1.3, 3.1.1.4, and 3.1.2.1 through 3.3.2.3 into a cost effectiveness metric. *Many of these metrics relate to compliance with executive orders. Although this is important, it may not be necessary to review these metrics unless there is an issue with compliance. Additionally, many of the metrics are reviewed on a less frequent basis. Essentially, this aggregated metric will be a review of just a few metrics for most of the year. Since the budget and sequestration are major challenges that the DOD faces, cost effectiveness is a very important area. This may result in more of these sub-metrics being flagged for review or further discussion. It may also result in a more focused task group to analyze possible cost reductions. If this is the case, the current metrics may not be informative enough. However, for the purpose of our recommendation and the current metrics being used, we recommend aggregating these sub-metrics.*

Aggregate 4.1.1.1 through 4.2.3.2 and 4.4.1.1 into one metric *that communicates the recruitment, development, and retention of a diverse and competent workforce. When reviewing these metrics, they all communicated the training and status of the workforce itself. Hence, we believe they can be aggregated into one metric that only needs to be de-aggregated when one area is underperforming (or projected to underperform).*

Aggregate 4.3.1.1 through 4.3.2.1 and 4.5.1.1 through 4.5.2.1 into one metric *that communicates the status of the installation, infrastructure, and services provided to the workforce and their families. These metrics all relate to support, services, or infrastructure. Due to their related nature, we again believe that they can be aggregated into one metric and de-aggregated if a problem arises.*

Aggregate 5.1.1.1 and 5.1.2.1 into 5.1. *Due to the nature of these metrics, an in-depth analysis could not be performed. However, given our recommendation of utilizing cascading, we do suggest aggregating these metrics.*

**The Commander should have the ability to flag metrics for review for any reason. All of these metrics incorporate the concept of cascading previously mentioned. There is limited information regarding the Weapon System metrics, so we did not analyze or include them in our recommendation as we did not feel we had sufficient information to make a recommendation regarding these metrics.*

Appendix F: Cause/Effect Example



Objective 1.5.1.3 Improve Accuracy & Integration of SC & MX Plans Scorecard



Measure	YTD Target	YTD Actual	Variance	%	Progress Towards Plan
AFSC Production Direct Earned Hours (DPBH)					
Variance	20,581	21,333	538	98%	
AFSC Depot MX Personnel Variance	22,553	23,689	1,136	95%	
MSTR Funded Orders Variance	140,275	154,973	14,704	93%	
MSTR Financial Performance Variance	\$ 1,800	\$ 2,123	\$ 329	84%	
Overall all Status					

Discussion

- 4 of 4 measures on track
- Manpower red due to over-manning
- Progress coded green due to increase in orders, both OC and OO are manning above their plan, and will continue to climb based on current FY15 forecast.
- WR over-manned in FY14; unchanged workload in FY15, but has a VERA/VSIP in process
 - Offered to 679; 176 accepted to date
- Metric green based on manning needed for FY15 and expected VERA/VSIP results

Causes/Issues

- Order variance and production issues driving manpower red
 - Due to increased orders, both OC and OO are manning above their Plan
 - Due to production issues at WR, while they are currently over-manned, reductions are being deferred until production catches up.

Mitigation

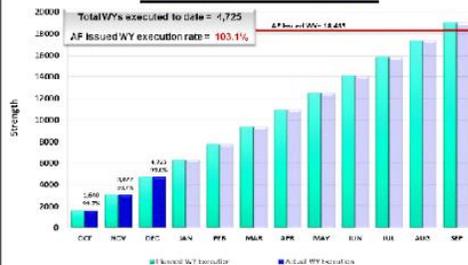
- Will continue to be reviewed monthly by corporate board

● All metrics green
 ● 1 or more metrics yellow
 ● 1 or more metrics red



Metric 3.1.2.1 Civilian Workyear Execution (O&M DBA)

Baseline & Trend



Discussion

- FY15 AF WY Allocation = 18,435
- FY15 AFMC Planned WYs = 19,029
- Centers were directed to create a plan to hire to FY15 Authorized strength of 19,587
- "Achievable" plan projected by the Centers executes 19,029 workyears and moves AFMC closer to FY15 authorized
- AFMC has executed 4,725 WYs to date
 - ~24.1% of 19,029 plan
 - ~25.6% of AF WY Allocation

Causes/Issues

- **FY15 PB ~ \$1.9B**
 - ~\$91M civ pay shortfall to fund AF issued WY of 18,435; resulting from the difference between programming AWC and the actual AWC
- **Additional civ pay shortfall of ~\$64M to fund Center "achievable" plans**
 - Keeps AFMC on track to hire to authorized
 - Prevents a possible 1 for 1 hiring freeze

FY15 Mitigation

- Request additional civ pay to allow Centers to execute towards FY15 authorized levels
- A1 and FM are meeting monthly with Centers to closely monitor execution

Metric POC: HQ AFMC/A1K
Ms. Tammy Lyons. DSN 767-7203

● 100% - 98%
 ● 97.9% - 96%
 ● <96%

1

Appendix G: Acronyms

AFMC — Air Force Materiel Command

BSC — balanced scorecard

CC — Commander

DOD — Department of Defense

HQ — Headquarters

IAW — in accordance with

KPQ — key performance questions

KPI — key success indicators

PM — performance measurement

Stan/Eval — standardization and evaluation

TQM — total quality management

USAF — United States Air Force

Appendix H: Glossary

Competitive advantage — *unique characteristic of a business that gives them an advantage against other competition; could be personnel, technology, or culture, among others*

External validity — *a study is externally valid if the results of the experiment are generalizable to populations other than the study*

Internal validity — *a study is internally valid if the effect is clearly attributed to the independent variable*

Key performance questions — *questions that capture exactly what you need to know, to track and monitor strategy execution and implementation*

Key success indicators — *help an organization define and measure progress toward organizational goals*

Lagging indicators — *indicators that show past performance*

Leading indicators — *indicators that signal future performance*

Stretch targets — *targets that require improvement and innovation to be met; a target that cannot be easily achieved*

Warfighter — *the operational branch of the Air Force; mainly, pilots*

Bibliography

- Adams, Chris. "What Is a Human's Cognitive Capability." <http://ergonomics.about.com/od/ergonomicbasics/f/What-Is-a-Human-Cognitive-Capability.htm>.
- Air Force Materiel Command Fiscal Year 2013 Quarterly Metrics. United States Air Force. 2013.
- Ammons, David N. "Performance Measurement and Managerial Thinking." *Public Performance & Management Review* 25, no. 4 (2002): 344-47.
- "Avoiding Working Memory Overload in Your Surveys." *Net Promoter Score (NPS) Monitoring and Analysis*. July 16, 2013. <http://www.npsmonitoring.com/blog/2013/07/avoiding-working-memory-overload-surveys/>.
- "Balanced Scorecard Working for the US Military." *Advanced Performance Institute* (www.ap-institute.com). <http://www.ap-institute.com/Balanced%20scorecard%20working%20for%20the%20US%20military.html>.
- Behn, Robert. "Avoid Being Caught in the KPI Quagmire." *Bob Behn's Performance Leadership Report*, 12, no. 3 (2014).
- Behn, Robert. "Measurement, Management, and Leadership." *Bob Behn's Performance Leadership Report*, 11, no. 3 (2013).
- Behn, Robert. "What Performance Management Is and Is Not." *Bob Behn's Performance Leadership Report*, 12, no. 1 (2014).
- Behn, Robert D. "Why Measure Performance? Different Purposes Require Different Measures." *Public Administration Review* 63, no. 1 (2003): 586-606.
- Behn, Robert. "Performance Leadership: 11 Better Practices That Can Ratchet Up Performance." *IBM Center for the Business of Government*. May 2004.
- Bouckaert, Geert, Wouter Van Dooren, Harry P. Hatry, Elaine Morley, Scott P. Bryant, Harry P. Hatry, Blaine Liner, Elisa Vinson, Ryan Allen, Pat Dusenbury, Scott Bryant, and Ron Snell. "Performance Measurement: Getting Results." *Public Performance & Management Review* 25, no. 3 (2002): 329-35.
- Brown, Mark Graham. *Keeping Score: Using the Right Metrics to Drive World-Class Performance*. (Boca Raton, FL: CRC Press, 1996).

- Burkholder, Alex A. "New Approaches To The Use Of Lagging Indicators." *Business Economics* 15, no. 3 (1980): 20.
- Buytendijk, Frank. "Zen and the Art of the Balanced Scorecard." Oracle. 2008. <http://www.oracle.com/us/solutions/business-intelligence/063553.pdf>.
- Camoës, Jorge. "Fibonacci, Working Memory and Information Overload." *The Excel Charts Blog*. March 10, 2008. <http://www.excelcharts.com/blog/fibonacci-working-memory-and-information-overload/>.
- Campbell, Dennis, Srikant M. Datar, Susan L. Kulp, and V.G Narayanan. "Testing Strategy with Multiple Performance Measures Evidence from a Balanced Scorecard at Store24." 2008.
- Champagne, Bob. "Too Many KPI's? - Tips for Metrics Hoarders..." *Performance Perspectives*. February 24, 2011. <http://epmedge.com/2011/02/24/too-many-kpis-tips-for-metrics-hoarders/>.
- Chan, Siu Y. "The Use of Graphs as Decision Aids in Relation to Information Overload and Managerial Decision Quality." *Journal of Information Science* 27, no. 6 (2001): 417-25.
- Demski, Joel. "Optimal Performance Measurement." *Journal of Accounting Research* 10, no. 2 (1972): 243-58.
- "Developing Performance Metrics." *Developing Performance Metrics*. <http://www.orau.gov/pbm/documents/overview/uc.html>.
- Dobbs, Richard and Koller, Timothy. "Measuring Long-Term Performance." *McKinsey Quarterly*. http://www.mckinsey.com/insights/corporate_finance/measuring_long-term_performance.
- Doran, George T. "There's a S.M.A.R.T. way to write management's goals and objectives". *Management Review*, 70, no. 11 (1981): 35–36.
- Doran, George. "Why SMART Objectives Don't Work." *Rapid Business Improvement*. September 19, 2011. <https://rapidbusinessimprovement.wordpress.com/2011/09/19/why-smart-objectives-dont-work/>.
- "Establishing Effective Metrics for New Product Development Success." *Siemens PLM Software*. 2009.
- Gerrish, Ed. "The Impact of Performance Management on Performance in Public Organizations: A Meta-Analysis." 2014.

- Glaser, Mark. "Tailoring Performance Measurement to Fit the Organization: From Generic to Germane." *Public Productivity & Management Review* 14, no. 3 (1991): 303-19.
- Halper, Fern. "How To Build A Metrics-Driven Company: Getting from Data to Metrics." *Dashboard Insight*. March 29, 2011. <http://www.dashboardinsight.com/articles/business-performance-management/how-to-build-a-metrics-driven-company.aspx>.
- Johnston, R., S. Brignall, and L. Fitzgerald. "'Good Enough' Performance Measurement: A Trade-off between Activity and Action." *Journal of the Operational Research Society* 53, no. 3 (2002): 256-62.
- Kelton, Andrea Seaton and Pennington, Robin R. "Internet Financial Reporting: The Effects of Information Presentation Format and Content Differences on Investor Decision Making." *Computers in Human Behavior*, 28 (2012).
- Kimberg, Daniel Y., Mark D'esposito, and Martha J. Farah. "Cognitive Functions in the Prefrontal Cortex--Working Memory and Executive Control." *Current Directions in Psychological Science* 6, no. 6 (1997): 185-92.
- Klemm, William R. "Training Working Memory: Why and How." *Psychology Today*. March 26, 2012. <https://www.psychologytoday.com/blog/memory-medic/201203/training-working-memory-why-and-how>.
- "KPOs & KPIs." *Strategy Pal*. <http://xtra.strategypal.com/audit/strategy-questions/kpqs-a-kpis/>.
- Kravchuk, Robert S., and Ronald W. Schack. "Designing Effective Performance-Measurement Systems under the Government Performance and Results Act of 1993." *Public Administration Review* 56, no. 4 (1996): 348-358.
- Kua, Patrick. "An Appropriate Use of Metrics." *Martinfowler.com*. February 19, 2013. <http://martinfowler.com/articles/useOfMetrics.html>.
- Lindsay, Kevin. "For the Clearest Market Insight, Analyze Both Leading and Lagging Indicators." *Entrepreneur*. September 11, 2014. <http://www.entrepreneur.com/article/236847>.
- Marr, Bernard. "20 Years of Measuring and Managing Business Performance: From KPIs and Dashboards to Performance Analytics and Big Data." *The Advanced Performance Institute* (www.ap-institute.com) and *Acute: The BIRT Company*. 2012.
- Marr, Bernard and Creelman, James. "Performance Management, Analytics, and Business Intelligence: Best Practice Insights from the Police." *The Advanced Performance*

- Institute (www.ap-institute.com). 2012, http://www.ap-institute.com/media/465536/120118_durham.pdf.*
- Marr, Bernard and Shore, Ian. "Cascading Balanced Scorecards: Using Strategic Maps to make Performance Relevant to RAF Stations." The Advanced Performance Institute (www.ap-institute.com). 2008. http://www.ap-institute.com/media/4411/cascading_balanced_scorecards.pdf.*
- Marr, Bernard, Shore, Ian, and Major, Paul. "'Believing' in Performance: Measuring and Managing What Matters in Chaplaincy," The Advanced Performance Institute (www.ap-institute.com). 2012. http://www.ap-institute.com/media/454753/120103_raf_chaplaincy.pdf.*
- McLeod, Saul. "Information Processing." Simply Psychology. January 1, 2008. <http://www.simplypsychology.org/information-processing.html>.*
- McLeod, Saul. "Working Memory." Simply Psychology. January 1, 2012. http://www.simplypsychology.org/working_memory.html.*
- Malamed, Connie. "20 Facts You Must Know About Working Memory." The ELearning Coach. Accessed March 22, 2015. <http://thelearningcoach.com/learning/20-facts-about-working-memory/>.*
- Marguis, Hank. "7 Dirty Little Truths About Metrics." ITSM Solutions Newsletter. July 5, 2006. <http://www.itsmsolutions.com/newsletters/DITYvol2iss26.htm>.*
- "Measuring Success: Making the Most of Performance Metrics." GE Capital. <http://www.americas.gecapital.com/insight-and-ideas/capital-perspectives/measuring-success-making-the-most-of-performance-metrics>.*
- Michelon, Pascale. "What is Working Memory? Can it Be Trained?" SmartBrains. <http://sharpbrains.com/blog/2010/11/16/what-is-working-memory-can-it-be-trained/>.*
- Miller, George A. "The Magical Number Seven, Plus Or Minus Two: Some Limits On Our Capacity For Processing Information." Psychological Review 63 (1956): 81-97.*
- Miller, George A. "Information and Memory." Scientific American, 1956, 42-46.*
- Morin, Amanda. "5 Ways Kids Use Working Memory to Learn." Understood.org. December 16, 2013. <https://www.understood.org/en/learning-attention-issues/child-learning-disabilities/executive-functioning-issues/5-ways-kids-use-working-memory-to-learn>.*
- "Public Procurement Practice: Performance Metrics." Principles and Practices of Public Procurement. 2012.*

- Reh, F. John. "Key Performance Indicators or KPI." *About Money*.
<http://management.about.com/cs/generalmanagement/a/keyperfundic.htm>.
- Rompho, Nopadol. "Why the Balanced Scorecard Fails in SMEs: A Case Study." *Faculty of Commerce and Accountancy*. 2011.
<http://www.ccsenet.org/journal/index.php/ijbm/article/viewFile/10247/8988>.
- "SecDef Warns of Pay Cuts." *Military Officers Association of America*. March 20, 2015.
http://www.moaa.org/Main_Menu/Take_Action/Top_Issues/Serving_in_Uniform/Compensation/SecDef_Warns_of_Pay_Cuts.html.
- Silverman, Eli B. *NYPD Battles Crime: Innovative Strategies in Policing*. Boston: Northeastern University Press, 1999.
- "Understanding Information Overload." *Infogineering RSS*.
<http://www.infogineering.net/understanding-information-overload.htm>.
- Van Winkle, William. "Information Overload." *Information Overload*. <http://www.gdrc.org/icts/i-overload/infoload.html>.
- "What Is Information Overload?" *WiseGEEK*. <http://www.wisegeek.org/what-is-information-overload.htm>.
- Wipawayangkool, Kamphol. "Alleviating Knowledge Overload In IT-HR Collaboration In Information Security Management Via Transactive Memory Systems." *Journal of Knowledge Management Practice* 14, no. 3 (2013).