Massachusetts Institute of Technology
Engineering Systems Division

Working Paper Series

ESD-WP-2003-03

METRICS PILOT PROJECT FOR MILITARY AVIONICS SUSTAINMENT: EXPERIMENTAL DESIGN AND IMPLEMENTATION PLAN

Kirkor Bozdogan
MIT Co-Lead of the Enterprise Integration Team (EIT) of the
LEAN SUSTAINMENT INITIATIVE (LSI)

Benjamin M. Brandt
Capt. Brandt (USAF) is a graduate student Research Assistant and
Candidate for the MS Degree in Technology and Policy at MIT.

Joseph M. Sussman
MIT Co-Lead of the Enterprise Integration Team (EIT) of the
LEAN SUSTAINMENT INITIATIVE (LSI)

JUNE, 2003

WORKING PAPER
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ENGINEERING SYSTEMS DIVISION

METRICS PILOT PROJECT FOR MILITARY AVIONICS SUSTAINMENT: EXPERIMENTAL DESIGN AND IMPLEMENTATION PLAN

Kirkor Bozdogan[#]
Benjamin M. Brandt*
Joseph M. Sussman[#]

Center for Technology, Policy, and Industrial Development
Massachusetts Institute of Technology
Cambridge, Massachusetts

October 10, 2002

[#]Dr. Bozdogan and Prof. Sussman serve as MIT Co-Leads of the Enterprise Integration Team (EIT) of the LEAN SUSTAINMENT INITIATIVE (LSI).

*Capt. Brandt (USAF) is a graduate student Research Assistant and Candidate for the MS Degree in Technology and Policy at MIT.

Acknowledgements: This working paper has benefited from discussions the authors have had with a number of individuals at the US Air Force Ogden Air Logistics Center during a site-visit on 8-10 July 2002, whose cooperation and assistance is gratefully acknowledged: Jim Lengyel (OO-ALC/MAWW, Nancy Long (OO-ALC/LGPB), Dave Jensen (OO-ALC/MALA), Sam Thomas (OO-ALC/MALA), Capt. Dominic Clementz (OO-ALC/MALA), Mike Jackson (OO-ALC/LGFBR), Chuck Viganski (Lockheed Martin; support to OO-ALC/LGFBR), and Fred Smullin and Tony Pombo (of Total Quality Systems, Inc., contractor for the Depot Repair Information Local Server – DRILS). The authors also acknowledge helpful discussions with Prof. John R. Hauser (MIT Sloan School of Management).

Further Inquiries: Please direct further inquiries or comments to Kirk Bozdogan (MIT), bozdogan@mit.edu; TEL: 617 253-8540; FAX: 617 258-7845

EXECUTIVE SUMMARY

This working paper outlines the design of an experiment, employing a pilot project, for identifying and validating new metrics for managing the US Air Force military avionics sustainment system. The paper also presents a plan for implementing the pilot project. The experimental design allows for the quantitifation of the effects of the new metrics, while controlling for the effects of other factors impacting the observed outcomes.

Underlying the pilot project, and the proposed experimental design, are three main hypotheses derived from earlier research: (a) currently used metrics foster *local optimization* rather than *system-wide optimization*; (b) they do not allow measures of progress towards the achievement of system-wide goals and objectives, and, hence, do not allow visibility into the impact of depot maintenance on the warfighter; and (c) they are driving the "wrong behavior," causing suboptimal decisions governing maintenance and repair priorities and practices and, as a result, undermining the efficiency and effectiveness of the sustainment system, despite the fact that the Air Force sustainment system has a dedicated and highly skilled workforce supporting the warfighter.

For the purposes of this pilot project, a *nonequivalent comparison group* experimental design is proposed. Such a design makes use of *before* and *after,* as well as *with* (treatment) and *without* (treatment) comparisons for two independent groups that are identical or highly similar to each other in terms of their essential characteristics. Here, one group is exposed to the *treatment,* while the other serves as the *control* group. *Treatment* refers to the introduction of new metrics, including "new operating rules"; *control* means an absence of new metrics. The Ogden Air Logistics Center (ALC) avionics sustainment site is proposed to serve as the *treatment* case and the Warner Robins ALC avionics sustainment site is proposed to serve as the *control* case. The two comparison groups are not equivalent in the sense that they differ in terms of the *new metrics* that would have a measurable impact on the outcomes.

In both cases, sustainment is defined broadly to encompass both the maintenance, repair and overhaul (MRO) operations and the attendant procurement, materiel, financial and supply chain management functions and organizations. A proposed back-up plan is to use a simpler *one-group-pretest-posttest* experimental design framework focusing only on the Ogden ALC avionics sustainment site. Under this back-up design *before* (i.e., a defined period prior to the introduction of new metrics) would represent the *control* case and *after* (i.e., a defined period following the introduction of new metrics) would represent the *treatment* case.

The pilot project is proposed to be executed during a one year period. At the conclusion of the project, the results will be evaluated and plans for further pilot projects will be developed and implemented, as appropriate. Further pilot projects may entail, for instance, an evaluation of the effects of an expanded treatment regime, including both new metrics and introduction of lean practices, into depot repair operations and the supporting supplier base.

The pilot project will focus on MRO operations relating to a specific set of pre-selected components (end-items or Line Replaceable Units – LRUs) that will effectively serve as the *units of analysis* in the experiment. These end-items consist of two samples: five high MICAP (mission capability) end-items, where the lack of serviceable items results in not-fully-mission-capable aircraft, and a sample of five "supportable" end-items that are normally provided to the operating bases in response to backorder requisitions. While supporting the MICAP items is taken to represent fulfilling urgent customer needs, meeting the backorder requisitions is defined as fulfilling normal customer needs. "Customer" is defined as the combat units; for the purposes of this pilot project, "customer" encompasses the operating bases.

The following list of the MICAP and normally supportable end-items, to serve as the units of analysis at the Ogden ALC, are currently being considered:

Sample of MICAP End-Items

Shop; National Stock Number Designation	Description	Abbreviation
Radio Frequency Shop		
1270-01-233-0011WF	Modular Low-Power Radio Frequency	MLPRF
5985-01-212-2950WF	Antenna	ANTENNA
1270-01-102-2962WF	Low-Power Radio Frequency	LPRF
Display and Indicators Shop		
6625-01-193-8861WF	Multi-Function Display	MFD
Microwave Shop		
1270-01-238-3662WF	Dual Mode Transmitter	DMT
1270-01-132-6867WF	Low Noise Assembly*	LNA

NOTE: *Shop Replaceable Unit (SRU), which is an important part of MLPRF.

Sample of Supportable End-Items

Shop; National Stock Number Designation	Description	Abbreviation
Displays and Indicators Shop		
1270-01-468-8658WF	Heads-Up Display Electronic Unit	HUD EU
5826-01-052-1945NT	Mode Select Coupler	MSC
Computer and Inertial Shop		
6615-01-448-6152WF	Digital Flight Control Computer	DFLCC
6615-01-042-7834WF	Rate Gyro Assembly	RGA
Processor and Pneumatics Shop		
5998-01-080-3978WF	Jettison Remote Interface Unit	JRIU
1290-01-109-1499WF	Missile Release Interface Unit	MRIU

These end-items will be matched with the same or analogous end-items maintained and repaired at the Warner Robins ALC. In particular, the high MICAP items will be examined closely to make sure that those selected share technological commonality and are highly comparable in terms of the underlying causes of their MICAP status. An option being considered is to focus directly on supportable end-items reflecting a stable maintenance environment and defer, for now, an analysis of the MICAP items since they may often involve a "chaos" environment and may introduce an added layer of complexity into the pilot project.

Focusing on these specific end-items, the pilot project will test a new set of proposed *outcome metrics* as well as *enabling metrics*, and conduct an evaluation of their impact on the efficiency and effectiveness of the depot sustainment system. *Outcome metrics* represent measures of value the depot maintenance system, taken as a total enterprise, delivers to the customer. They are customer satisfaction metrics. They gauge how effectively the customer's urgent and normal needs are met, how well they are met (e.g., in terms of product quality, customer wait-time), and how cost-effectively they are met (e.g., average cost of repair). *Enabling metrics* gauge how well the sustainment system performs various processes, functions and practices – by making sure that the organization as a whole is doing the "right job" as well as doing the "job right" – to deliver the defined outcome metrics benefiting the customer.

The following set of new metrics are proposed for the pilot project:

New Metrics		Brief Description
OUTCOME METRICS		
	Urgent customer requirements satisfaction rate (UCRSR)	Ratio of the total number of serviceable end-items (in MICAP status) provided during a given period *to* total number of high MICAP items requisitioned during that period.
	Normal customer requirements satisfaction rate (NCRSR)	Ratio of the total number of serviceable end-items provided during a given period *to* the total number of backorders issued by bases.
	Weighted customer requirements satisfaction rate (WCRSR)	Weighted average of UCRSR and NCRSR, where the weights are quantified on the basis of the previously negotiated quantities using the "Quasi-EXPRESS" experiment. At WR-ALC, the weights may be derived from EXPRESS-driven inductions of MICAP and backorder items.
	Unit cost of maintenance	Actual incurred unit cost of maintenance, reflecting full costs of materials, labor and cost of utilization of capital equipment, using the prevailing direct labor, overhead, and general and administrative (G&A) rates. Actual labor hours include accumulated labor hours for repairing end-items across all shops, including SRU repair.
	Product quality	Total number of serviceable end-items produced by depot maintenance during a given period that are found to be defective (end-items with Quality Deficiency Reports (QDRs) sent back to the depot and, upon re-testing, are found to be defective).
	Customer wait time	Total elapsed time (hours) from issuance of a requisition until receipt of a serviceable end-item at base supply that is available to base maintenance upon request.
ENABLING METRICS		
	End-item shopfloor flow time (variance)	The statistical variance in shopfloor flow time, measuring variability (a Six Sigma concept to reduce process variation to drive continuous improvement).
	Cost of maintenance and repair (variance)	The statistical variance in unit cost of maintenance and repair, based on actually incurred costs and using prevailing direct labor, overhead and G&A rates (intended to eliminate waste and drive down unit costs through standard work processes and other lean methods).
	Productivity	Total number of serviceable end-items produced by depot repair per unit of labor-hours (e.g., 1000 labor-hours); not adjusted for unscheduled work-stoppages due to equipment failure in order to motivate depot-repair to put greater emphasis on preventive maintenance, a lean concept.

	Responsiveness	The ratio of *required Takt time* to the *observed Takt time,* to gauge how well depot maintenance as a system is responding to the pace of real customer demand for serviceable end-items. Detailed explanation is given in the text. *Takt time* is a basic lean concept in designing manufacturing operations to evolve a "pull-based" production system.
	End-item repair parts combined fill rate	Combined fill rate for a specific set of pre-defined repair parts and materials, based on previous repair history of end-items showing most frequently failing parts. This is in contrast with currently used issue or stockage effectiveness metrics at the individual part level. The metric is intended to motivate supply organizations to collaborate more closely in supporting the "fixer" and to motivate a much closer working relationship between the repair, supply and financial organizations.
	Supplier delivery performance	The "order-to-delivery" time for repair parts and materials obtained from suppliers, including the Defense Logistics Agency (DLA), measured in terms of total elapsed time from order-to-delivery for 50%, 75% and 95%, respectively, for all repair parts and materials requisitioned.

These proposed metrics are grounded in previous research which reveals that the causal structure of metrics used by the Air Force is a lot more complex than the conventional top-down hierarchical view of metrics [See references given at the end of this paper]. This suggests the adoption of an adaptive control feedback mechanism approach to metrics, not a top-down command-and-control approach. The main aim of these new metrics is to help coordinate the actions of numerous organizational entities, teams and processes through the establishment and clear communication of common goals, help foster a new culture, and develop metrics that motivate and reward teams striving to optimize system-level goals.

The implementation of these new *outcome metrics* and *enabling metrics* includes the concurrent adoption of a number of *new operating rules*, to pave the way for the adoption of these new metrics by removing some obvious roadblocks. These new operating rules include both incentive systems and rewards. It is expected that improvements in these *outcome metrics* (customer satisfaction measures) will impact the warfighter directly: (a) by leading to a *reduction* in TNMCS (Total Not Mission Capable for Supply), through improvements in base-supply of serviceable items, that should directly help to increase the FMC (Fully-Mission-Capable) rates; and (b) by promoting greater efficiency in base-repair operations – through reduced cannibalization in light of greater availability of serviceable items, such that the available resources can be put to more productive pursuits – which would tend to reduce the TNMCM (Total Not Mission Capable for Maintenance) rates and thereby increase the FMC rates.

A basic tenet of this proposed plan is that the pilot project should be "owned" by the principal participating organizations that will be responsible for its execution based on agreed-upon ground rules and operating procedures. The role of the MIT LSI researchers will be to provide technical support and data analysis.

The execution of the pilot project will require the development and tracking of a detailed set of data, for the previous year as well as during the time the pilot project is being executed, in order to conduct the necessary analytical tasks required to test the validity and benefits of the new metrics. Weighed against the expected future benefits of the new metrics, the extra cost of such a data collection effort is an investment well-worth making -- representing a potentially large benefit-cost ratio.

The expected benefits of the pilot project include significant improvements in the efficiency and effectiveness of the avionics sustainment system that can be migrated to many other component repair environments within the Air Force "organic" sustainment system. One of the key benefits would be showing the impact of depot maintenance on the warfighter. Also, the pilot project is expected to provide a framework for evaluating tradeoffs, leading to better decision-making. In addition, the pilot project is expected to prototype and codify a rigorous approach for introducing new metrics into the Air Force's metrics structure and for continually improving existing metrics. Finally, the pilot project will provide the commercial providers of contract maintenance services a new process for improving their performance metrics in a way that is synchronized with expected improvements within the "organic" sustainment system, so that the entire Air Force sustainment system can be optimized to provide the best support to the warfighter at the least cost.

To ensure successful execution of the pilot project, the participating organizations are invited to review and approve the proposed experimental design and implementation plan on a fast-track basis, come to a common understanding of the new working relationships required to implement the project through concerted action, help develop the necessary data and put in place the needed data monitoring steps, and work with the MIT LSI research team on a collaborative basis to help launch and continue to actively support the pilot project.

I. INTRODUCTION

This working paper is designed to provide a proposed conceptual and operational plan for the implementation of the **Metrics Pilot Project on the Sustainment of Avionics Systems.** The pilot project was recommended by the Enterprise Integration Team (EIT) of the Lean Sustainment Initiative (LSI) at a meeting of the LSI Steering Group on 19 December 2001 and was approved by the Steering Group. The Steering Group consists of the senior leadership of the LSI stakeholders community encompassing both the Air Force sustainment system and commercial providers of contract maintenance and repair services supporting the Air Force. This proposed planning document is presented for review and approval by the participating stakeholder organizations.

The rest of the working paper is organized into the following parts:

- Motivation
- Objectives
- Hypotheses
- Experimental design
- New metrics and related data requirements
- Project structure

- Resource requirements
- Expected deliverables and benefits
- Major tasks and project schedule; and,
- Next steps.

A glossary of the acronyms used in the paper is given in Attachment A.

Throughout the paper, the terms "depot-repair" and "depot maintenance" may be used interchangeably. The intent here is to define the avionics sustainment system broadly to encompass both the MRO operations and the attendant procurement, materiel, financial and supply chain management functions and organizations. This is consistent with the broader view adopted in the paper that the depot maintenance system, however it may be currently structured organizationally, is, in effect, in the business of creating and delivering best value to the *real* customer, which is the combat units. For reasons detailed in the text, definition of the "customer" is extended to encompass the operating bases, where base supply effectively serves as the actual interface between the "customer" and the depot maintenance system.

II. MOTIVATION

A major motivation for performing this pilot project is to explore, identify, define and test metrics and analytical methods for linking improvements in performing avionics-related component maintenance and repair services to the achievement of system-wide performance improvements (e.g., increased rates of fully-mission capable (FMC) aircraft). This would enable tradeoff decisions *between* the achievement of system-level objectives (e.g., increasing the rates of FMC-aircraft) *and* actions designed to improve component-repair operations (e.g., at the depot level). Hence, there appears to be a real need not only for a disciplined process for identifying more effective metrics, and formal ways of validating their effectiveness, but also for an analytical link between *local* improvement actions and *system-level* performance outcomes.

Generally speaking, it is instructive to test out the feasibility, workability and usefulness of new business processes and practices in a controlled "pilot project" setting before their wholesale implementation throughout the Air Force sustainment enterprise. It would similarly be beneficial to test the adoption of new performance metrics in a controlled experimental setting. Thus, a well-designed pilot project concentrating on avionics MRO operations is expected to validate the usefulness and expected benefits of a new set of "intervention" metrics driving avionics-related sustainment operations. The results of such a pilot project are expected to have considerable spillover benefits for other component-level repair operations as well in both government and industry.

III. OBJECTIVE

The objective of this pilot project is to evaluate the overall effectiveness, operational feasibility, workability and potential benefits of a new or modified set of metrics designed to foster system-wide optimization of the sustainment system to benefit the warfighter, concentrating on "organic" avionics MRO operations and the related procurement, materiel, finance and supply organizations, in a controlled experimental setting.

IV. HYPOTHESES

The main hypotheses motivating the pilot project, which can be tested by conducting the proposed experiment, are presented below. Following this discussion, interactions between the bases and depot-repair are explored and the implications of these interactions for experimental design are outlined. A more detail discussion of the underlying research leading to the main hypotheses outlined below can be found in references given at the end of this paper.

A. Main Hypotheses

Three main hypotheses driving the pilot project are stated and discussed below.

Hypothesis 1: Currently used metrics generally foster *local optimization* rather than *system-wide optimization.* Consequently, the performance of the sustainment system, in terms of its efficiency and effectiveness, is compromised and the sustainment system does not provide the best support to the warfighter. *If currently used metrics were realigned or largely replaced by new metrics that drive behavior towards improving the overall performance of the sustainment system rather than showing better financial or other performance measures locally, then the efficiency and effectiveness of the "organic" avionics sustainment system will improve, resulting in better support of the warfighter.*

Discussion: Current levels of performance of the "organic" avionics sustainment system are the outcome of many root causes reflecting a complex set of interactions among numerous factors, including currently used metrics, driving production priorities and practices. These factors include the quality of the capital stock (e.g., testing equipment), capability of the workforce, funding availability, policies governing investment in new spares, availability of repair parts and materials, contracting practices, organizational structure and the division of responsibility among various organizational units as well as the degree of cooperation among them, and a host of broader government policies, regulations and procedures governing many aspects of the sustainment system. The root causes underlying current performance levels require a wholesale attack to eliminate them. These root causes, or the factors contributing to them, may not disappear in the course of the pilot project. That is, introduction of new metrics cannot be expected to completely eradicate the negative effects of some or all of these factors. However, new metrics may well positively influence some of them to make a significant difference in terms of inducing local behavior that is more conducive to system-wide optimization, resulting in enhanced system-wide performance outcomes.

Hypothesis 2: Currently used metrics generally do not allow measures of progress towards the achievement of system-wide goals and objectives; the metrics do not allow visibility into the impact of depot maintenance on the warfighter. For example, they do not allow for quantifying the *incremental impact*, on F-16 mission capability rates, of an increase in metrics gauging the availability of repair parts and materials for the maintenance of the F-16 avionics system through an *incremental investment* in more parts and materials. Consequently, informed tradeoff decisions cannot be made at the system level to provide more cost-effective support to the warfighter. *If existing metrics were realigned or largely replaced by a new set of metrics that are linked together at multiple levels in the form of a cascading, mutually-supporting, chain of metrics, then the incremental contribution of sustainment initiatives at various levels to overall system-wide* *performance outcomes can be more readily and transparently gauged, yielding informed tradeoff decisions to deliver better value to the warfighter.*

Discussion: The metrics currently used by the Air Force sustainment system generally do not allow measurement of progress towards the achievement of system-wide goals for several reasons. A major reason is that depot maintenance metrics are fundamentally in conflict with customer support metrics. Also, currently used metrics are not consistent *vertically* from the highest to the lowest echelons (i.e., from the "corporate" Department of Defense to the Air Staff to AFMC to the shop-floor). Further, they are not consistent *horizontally*, from the flightline to the shop-floor and beyond, encompassing the supplier network supporting depot-repair. The supplier network covers numerous suppliers providing new spares and repair parts, the Defense Logistics Agency (DLA), and commercial providers of contract maintenance, repair and overhaul (MRO) services supporting the depots. In addition, currently used metrics often reflect financial and other measures driving internal priorities of organizational silos. Moreover, they lack accountability and can often be "gamed" by managers at virtually all levels. Finally, there is considerable confusion between metrics, performance measures, process indicators and status reports.

These shortcomings of the currently used metrics can be overcome by realigning or largely replacing existing metrics with a new set of metrics that are linked together and are mutually-supportive, both and *vertically* (e.g., from the Air Staff to AFMC to the shop-floor) and *horizontally* (e.g., from the flightline to the shop-floor to the supplier network. The "House of Metrics" approach mentioned in a companion LSI Working Paper[1] and outlined in an MIT Master's Thesis performing an exploratory evaluation of the current metrics used by the Air Force sustainment system[2], can be helpful in constructing such a metrics structure, informed by the quantitative empirical results of the "Metrics Thermostat" research at MIT focusing on the F-16 sustainment system[3]. Substantively, these results can be aided by the introduction of lean performance metrics derived from lean principles for managing complex modern enterprises, tailored to the sustainment system. Lean thinking is defined as the "dynamic, knowledge-driven and customer-focused process by which all people in a defined enterprise continuously eliminate waste with the goal of creating value."[4]

Hypothesis 3: Currently used metrics by the Air Force sustainment system are driving the "wrong" behavior, causing suboptimal decisions governing maintenance and repair priorities and practices and, as a result, undermining the efficiency and effectiveness of the sustainment system. *If existing metrics were realigned or largely replaced by a new set of metrics that provide greater incentives for both "doing the right job" and "doing the job right", the efficiency and*

---

[1] Kirkor Bozdogan, "Summary of Findings, Current Projects and Planned Activities," Working Paper, Enterprise Integration Team of the Lean Sustainment Initiative, Massachusetts Institute of Technology, 30 August 2002.
[2] Stuart McGillivray, "Air Force Sustainment Performance Metrics: An Exploratory Evaluation," Master's Thesis, Engineering in Logistics, Massachusetts Institute of Technology, June 2002.
[3] See John R. Hauser, "Metrics Thermostat," *The Journal of Product Innovation Management*, Vol. 18, No. 3 (May 2001), 134-153. For a description of LSI research applying the "Metrics Thermostat" approach to the Air Force sustainment system, see Bozdogan, *op cit*. and refer to Keith A. Russell (Lt. Comdr., USCG), "Reengineering Metrics Systems for Aircraft Sustainment Teams: A Metrics Thermostat for Use in Strategic Priority Management," MS Thesis, Aeronautics and Astronautics *and* Technology and Policy, Massachusetts Institute of Technology, December 2000.
[4] See Earll Murman, et al., *Lean Enterprise Value: Insights from MIT's Lean Aerospace Initiative* (Great Britain [Houndmills, Basingstoke, Hampshire RG1 6XS] and New York: Palgrave, 2002), p. 90.

effectiveness of the sustainment system would improve, resulting in better support for the warfighter.

Discussion: Examples of "wrong" behavior that could be induced by the currently used metrics include improving the sales performance of a shop by placing higher priority on the maintenance of certain assets to offset losses in other areas, inducting into repair those end-items that are easier to fix and that bring quick sales benefits, or "gaming" the system in various ways (e.g., budgeting for high overtime, negotiating for higher standard rates). Other forms of "wrong" behavior that might be caused, at least in part, by currently used metrics may include cannibalization[5], lack of availability of repair parts and materials, extensive downtime for testing equipment for lack of the required repair parts, insufficient cross-training of the workforce, excessive delays in contracting, and lack of coordination among the various organizational entities responsible for depot maintenance and supply functions. It is difficult to know, *a priori*, how new metrics might realign existing incentives such that various classes of "wrong" behavior can be reduced or eliminated. However, it is possible to introduce certain "gatekeeper" metrics forcing greater cooperation, for example among various organizational entities responsible for the supply of both wholesale and retail items necessary for depot repair. Similarly, well-defined outcome metrics could induce greater cooperation between depot repair and supply operations.

It must be emphatically stated that the Air Force sustainment system has a dedicated and highly skilled workforce supporting the warfighter. Any references in this document to possible "wrong" behavior that might be induced by currently used metrics pertains only to possible organizational behavior that might be shaped by at least some of these metrics; such references in no way reflect a negative comment on the capability and commitment of the sustainment workforce, individually or taken as whole.

B. Base-Depot Interactions

The main hypotheses presented above involve certain base-depot interactions that need to be considered in the experimental design for the pilot project. To give a simple example, base repair may engage in extensive cannibalization to make up for lack of serviceable end-items or availability of repair parts and materials, perhaps because depot maintenance may have placed the wrong priority on what items to repair. Alternatively, to conserve base budgets, the bases may engage in extensive cannibalization on items before they are sent to the depot for repair, causing the depot to undertake much more expensive repair on those end-items in part because they may have suffered collateral damage in the course of cannibalization actions. Further, depot-repair may be forced to tie-up a considerable number of labor hours and testing-equipment time to the testing of end-items forwarded from bases because they were identified as "Cannot Duplicate" (CND) at the bases and that are later classified by depot repair as "No Fault Found" (NFF). Still another

---

[5] Cannibalization refers to the removal of a serviceable component from one end-item which is already awaiting parts (AWP-G) for use in making another end-item being repaired serviceable. Usually a distinction is made between robbacks and cannibalization. Robbacks refer to cases of removing a serviceable component from an end-item in AWP status, where the component is not physically attached to that end-item, for use in making another end-item in the repair process serviceable. In the case of cannibalization, the serviceable component that is removed from an end-item in AWP status is physically attached to that end-item. This may be a distinction without a difference. See Air Force Materiel Command (AFMC), Hill Air Force Base, Industrial and Logistics Training Division, *AWP Process Overview Training Program, Student Handout*, Command Course MOODMM000300SU, Revision January 2001, p. 5-1.

example may include cases where, competing for repair parts and materials, bases often trump depot-repair, thus depriving depot-repair of needed parts and materials and adversely affecting depot repair. These examples can be extended to include others.

The main issue here is how best to control for these strong two-way interactions between the bases and depot-repair, essentially to make sure that the base-to-depot influences (i.e., negative or positive effects flowing from the bases to depot-repair) do not end up confounding the effects of the new metrics applied to depot-maintenance. The attention here, then, is focused on how best to take into account the key base-to-depot influences; the proposed experimental design outlined below takes precautions regarding such base-to-depot influences.

Meanwhile, it is argued that the depot-to-base influences – that are expected to be mostly positive under the pilot project -- can be safely ignored. The reasoning for this assumption can be laid out as follows. It can be argued that any improvements in depot-maintenance efficiency and effectiveness due to the new metrics will result in improvements in base supply availability. This, in turn, will help reduce the numerical value of the Total Not Mission Capable Supply (TNMCS) metric. As a result, cannibalization at base-repair will most likely be reduced. Consequently, workforce and testing equipment resources will be freed-up for more productive uses. This will then most likely lead to a reduction in the numerical value of the Total Not Mission Capable Maintenance" metric. Ultimately, such a positively-reinforcing chain of developments at the base-level – all due to improvements in depot-maintenance performance – can be expected to result in higher TMC rates for the warfighter.

All else remaining equal, because of these expected positive effects flowing from depot-maintenance to base-level supply and repair operations, over time any negative feedback effects flowing from depot-maintenance to the bases will be reduced in importance. Thus, for the purposes of this pilot project, it should prove sufficient to control for possible base-to-depot influences (negative or positive), while ignoring any possible depot-to-base negative influences (this assumes that the effects of new metrics will be positive),[6] keeping in mind that that is argued that assumed that the boundaries of the pilot project can be defined as depot repair plus the supporting supply organizations and networks. Bases are thus taken as the "customer" and outcome metrics are defined as measures of customer satisfaction, as noted below in more detail.

V. EXPERIMENTAL DESIGN

This section of the paper outlines the experimental design options, defines the two comparison groups being considered for the pilot project (i.e., the Ogden ALC and Warner Robins ALC avionics sustainment systems), describes the units of analysis, and specifies the planned time period for the pilot project. A discussion of the new metrics, as well as other factors of interest in designing and executing the pilot project, is presented in the next section.

A. Experimental Design Options

[6] For the sake of completeness, it is assumed that depot-repair, under the influence of new metrics, will not experience a deterioration in its performance in terms of the repair of SRUs (Shop Replaceable Units), thus possibly falling short in its supply of serviceable SRUs, to the extent that depot-repair in fact supplies not only serviceable LRUs (Line Replaceable Units) but also serviceable SRUs to Consolidated Serviceable Inventory (CSI) for shipment to the bases.

There are a number of experimental design options that can be considered for the pilot project. Among these, two basic designs, in particular, are given primary attention. The first is the ***nonequivalent comparison group design***, which makes use of before and after comparisons for two independent groups that are identical or highly similar to each other in terms of their essential characteristics, where one group is exposed to the treatment (i.e., new metrics) while the other serves as the control group, receiving no treatment. That is, the two groups are not equivalent in terms of the *new metrics* that would have a bearing on the outcomes. The second is the ***one-group pretest-posttest design***, where the same group serves effectively both as the *control* case (before the treatment starts) and as the *treatment* case (after the treatment starts). The latter design option must guard against the counterfactual argument that the two may differ for reasons other than the treatment itself.

Typically, in scientific investigations each member of a population being studied has an equal chance of being chosen. This means, for instance, that each end-item being repaired at a given depot has an equal chance of being included in the study. This means random selection of the individual units of analysis. Moreover, each of these individual end-items has an equal chance of being assigned to various experimental conditions (e.g., different levels of treatment; treatment versus control). Randomization is essential to experimental design in the physical sciences. There are, however, a large class of situations, particularly in the social sciences, where randomization is not possible since real life situations cannot be controlled as in a laboratory. In such situations, it is necessary to make use of "quasi-experimental designs," where individual units are not assigned to various experimental conditions through a random process. While "quasi-experimental designs" lack random assignment of individual units to experimental conditions, they still must meet the same requirements expected of experimental designs involving randomization in terms of drawing causal inferences.

In the pilot project, a "quasi-experimental" design is used, since the individual units of analysis in the experiment (i.e., specific avionics end-items being examined) are not selected at random but rather on the basis of deliberate prior choice reflecting certain criteria (e.g., focusing on a pre-selected sample high MICAP[7] end-items *and* a pre-selected sample of supportable end-items) and, further, they are not randomly assigned one or the other of the two comparison groups. Moreover, the two comparison groups are not abstract constructs (e.g., different levels of treatment; treatment versus control) but rather two separate physical sites. By definition, individual units of analysis at each site cannot be randomly assigned to one or the other site. That is, at each site, randomization can occur only in the selection of the individual items for analysis, not in terms of how they are then assigned to one or the other site (i.e., to treatment versus control cases).

Regardless of what type of experimental design is chosen for the pilot project, the analytical objective remains the same. A first-order analytical objective in the pilot project is to test whether th*e treatment* (i.e., new metrics) have statistically significant effects on the measured outcomes, while accounting for the possible effects of both the confounding and control variables on the observed outcomes. If the effects are statistically significant, a second-order objective is to develop reliable measures of the differences *between* the actual observed outcomes due to the

[7] MICAP (mission capability) is a backorder priority designation to denote a condition where an aircraft is not mission capable for lack of a component; the requisition for that component by base-supply is called a MICAP requisition.

treatment (i.e., new metrics) *and* what would have happened to the measured outcomes in the absence of the treatment. The formal process pursued is grounded in established scientific principles of experimental design and data analysis.

1. Nonequivalent Comparison Group Design

The first and preferred option is a design which uses two comparison groups, one serving as the treatment group and the other serving as an untreated control group, with pretest and posttest data for both. This is known as the "Quasi-Experimental Design with both Control Groups and Pretests." A detailed technical description can be found in Shadish, Cook and Campbell.[8] This design is often called the nonequivalent comparison group design, in light of the fact that the two comparison groups, while they are considered alike or highly similar to each other prior to the onset of the treatment, become nonequivalent once the treatment starts. *Treatment* here involves the introduction of new metrics. Lack of treatment, or *control*, means an absence of new metrics. That is, the same metrics as before continue to be implemented at the control site, which is not exposed to the new metrics.

In this design, the Ogden ALC avionics sustainment system would serve as the *treatment group*, while the Warner Robins ALC avionics system would serve as the *control group.* The avionics sustainment system is defined as the avionics-related "organic" depot maintenance, repair and overhaul operations, as well as the supporting supplier networks encompassing the Defense Logistics Agency (DLA) and commercial providers of new spares, repair parts and components. While the Ogden ALC concentrates on the sustainment of the F-16 avionics system, Warner Robins ALC focuses, among other things, on the F-15 avionics system. The two avionics systems are generally comparable in terms of their state-of-technology and technical system architecture (i.e., federated system architecture, where overall control, and a certain amount of functionality of each avionics subsystem, is delegated to a central mission computer).

For both comparison groups, both *pretest* and *posttest* data (i.e., data to be collected both before and after the start of the treatment) would be developed on all metrics, as well as on other factors of interest. These "other factors of interest" refer to other variables, in addition to the new metrics, that may influence the observed outcomes. These include variables that change concurrently with the changes in the observed outcomes and are known as *confounding* factors, whose respective effects on the outcomes need to be separated from the effects of the new metrics. They also include those factors, known as *control* variables, which refer to key characteristics of the two test sites and of individual units of analysis consisting of pre-selected avionics end-items.

2. One-Group Pretest-Posttest Design

Under this design, the same site (i.e., Ogden ALC) is taken to serve as both the control and the treatment in the experiment. Pretest observations, involving no new metrics, effectively serve as the control. Posttest observations, following the introduction of the new metrics, serve as the treatment. Both pretest and posttest observations refer to all factors of interest, including the new metrics, for the end-items being analyzed which represent the individual "respondents" in the

[8] See William R. Shadish, Thomas D. Cook, and Donald T. Campbell, *Experimental Quasi-Experimental Designs for Generalized Causal Inference* (Boston and New York: Houghton Mifflin Company, 2002), pp. 135-170.

experiment. Data on other end-items, as well as on SRUs, will also be collected to test for any interaction effects.

As pointed out earlier, such a "quasi-experimental" design offers greater simplicity and ease of implementation compared with the two comparison groups design. However, it must have adequate safeguards built into it to make sure that reliable causal inferences can be drawn from it. One possible weakness of such a design is that the pretest and posttest numerical values of the outcome variables may differ not necessarily because of the effects of the treatment but because of other factors not related to the treatment, after properly accounting for any significant changes in the essential characteristics of the Ogden ALC sustainment system.

To provide protection against any threats to the validity of causal inferences from such a design, two additional safeguards can be introduced: adding a second pretest prior to the first, and using a nonequivalent dependent variable. In effect, having two pretests help detect and control for any biases that might exist in estimating the effects of the treatment if only the pretest and the posttest observations were compared. The nonequivalent dependent variable refers to a measure that is not expected to change because of the treatment. For example, introduction of new metrics is expected to affect a particular set of high-MICAP end-items that are difficult to sustain and a set of supportable end-items that are normally quite supportable, in terms of being able to provide the necessary maintenance, repair and overhaul services. However, the numerical values of the metrics for other fairly difficult-to-support end-items are not expected to show any effect due to the introduction of new metrics. One or more of these end-items could serve as the nonequivalent dependent variable, to show that the observed outcomes are not the result of a more general event or development affecting all types of end-items.

B. Treatment and Control Groups

If the two comparison groups design is adopted in the pilot project, the treatment and control groups (sites), as well as reporting requirements for each, would be defined as follows, while noting, again, that in the event of choosing the one-group pretest-posttest design, the Ogden ALC sustainment system would serve as both the control and the treatment site, as noted above.

1. Treatment Group

The treatment group at OO-ALC consists of four organizations:

- OO-ALC/MALA: The F-16 Avionics Branch, which provides maintenance and repair services on F-16 avionics end-items, consisting of both Line Replaceable Units (LRUs) and Shop Replaceable Units (SRUs).
- OO-ALC/LGF: The organizational entity responsible for all F-16 logistics operations.
- OO-ALC/LGS: The organizational entity responsible for the supply of all "wholesale" repair parts and materials.
- Defense Logistics Agency (DLA): The organizational entity primarily responsible for managing all consumable supplies, including hardware items, used by the military services. Traditionally, DLA buys supply items in large quantities to benefit from economies of scale, stores them in distribution depots until they are requested by the service depots, and then ships them to the appropriate depot facilities. Over the years, DLA's responsibility has

expanded from the supply of consumable items and piece parts to include progressively more and more complex aircraft parts and components.

OO-ALC/LGF funds the activities of OO-ALC/MALA (The F-16 Avionics Branch) and, from an organizational standpoint, serves as the "customer" for the repair services provided. It provides provisioning, cataloging, requirements determination, acquisition, distribution, repair and disposal of parts and components, and engineering and technical support services. OO-ALC/LGF also serves as a "retail" supplier to OO-ALC/MALA, providing roughly 10% of all repair parts it needs. OO-ALC/LGS, which is set up as a parallel organization to OO-ALC/LGF, is the main supplier of repair parts and materials, including piece parts. This organization serves as the primary liaison with DLA and has personnel stationed at the Shop Service Center (SSC), which houses the most frequently needed repair parts and manages the shop's workload. The definition of the participating organizations can be expanded to include others, such as contracting and finance, as appropriate.

Currently, the three organizations at OO-ALC are conducting a "work-around" experiment to meet warfighter demand without directly using the EXPRESS system. EXPRESS is the main information and decision-support system the Air Force uses for managing the depot-based component repair operations. More specifically, it prioritizes component repair requirements (i.e., specifying the priority order or sequence in which specific components should be inducted into the repair process) and also guides prioritized distribution of serviceable components to the operating combat units. The three organizational entities at OO-ALC involved in the provision of avionics-related maintenance and repair services believe that EXPRESS is not meeting warfighter needs correctly. Consequently, they have been together pursuing a "quasi-EXPRESS" process in order to better meet warfighter demand.

The "Quasi-EXPRESS" approach has driven OO-ALC/LGF and OO-ALC/MALA to coordinate their respective activities with each other to determine the quantity and mix of end-items to be produced by the repair process. OO-ALC/LGS is also involved in the "Quasi-EXPRESS" process, as item managers need to be informed of negotiated quantities in order to plan for, acquire and have in place the needed repair-parts and materials. While communications between and across these organizational entities have clearly improved recently through weekly meetings, it may be too early to gauge the degree to which this has benefited the effectiveness of the depot repair process. The new metrics outlined below are designed to further encourage closer communication and cooperation among these organizational entities.

There are problems with current levels of sustainment performance; OO-ALC/LGF and OO-ALC/MALA have contrasting views on the causal factors underlying current levels of performance. While one group points to lack of availability of repair parts as the main source of the problem the other disclaims the parts availability issue and instead cites insufficient manpower cross-training and shop capacity issues the main sources of the problem. The pilot project is being designed to account for these disparate views in analyzing the effects of the new metrics.

As the primary treatment group, the F-16 Avionics Production Branch (OO-ALC/MALA), will be key in evaluating the effects of the new metrics. Currently, its performance is measured on the basis of such metrics as production *versus* "negotiated requirements," productivity measures (yield, indirect labor ratio), and additional metrics required by LGFBR (e.g., depot shop flow time, logistics response time, awaiting parts-parts required (AWP-G), awaiting parts-parts supportable (AWP-F). The LGFBR-required metrics will continue to be observed during the pilot project, even though they will not necessarily be the primary measures of performance in the pilot project itself. However, they are important to analyze to determine the effect of the new metrics on the shop's performance based on existing metrics.

Contacts:

OO-ALC/LG Col. Audrey L. Wolff, Deputy Director, Logistics Management (To be contacted)
MALA: Mr. Dave Jensen, Branch Chief
Capt. Dominic Clementz, Deputy Branch Chief

LGFBR: Mr. Mike Jackson, LGFBR Section Chief
Mr. Chuck Vigansky, LGFBR Contractor

LGS: Ms. Marlene Wright, Division Chief

LGP: Mr. Jim Lengyel, Acting Division Chief, Government Co-Lead (LSI Enterprise Integration Team)

2. Control Group

The control group in the pilot project is designated as WR-ALC/LYP (Avionics) and related supply organizations. The nucleus of the control group is expected to be the F-15 avionics production shop (OO-ALC/LYPF), WR-ALC/LGS (Depot Supply Division), and the Defense Logistics Agency (DLA). While there are some differences in the avionics end-items repaired at OO-ALC and WR-ALC for the F-16 and F-15, respectively, the workload and workflow processes are quite similar. The F-15 avionics production shop relies on outside organizations to develop accurate predictions of demand. It also relies on the Defense Logistics Agency and other supply organizations for the repair parts and materials it needs. Further discussions with participating organizations will identify and seek the concurrence of other organizations associated with F-15 avionics repair operations. The definition of the participating organizations at the WR-ALC, as well, can be expanded to include others, such as contracting and finance, as appropriate.

The control group organizations are expected to continue providing weekly and monthly data on the existing metrics, as well as data on both confounding factors and control variables, as defined below. In addition, they are expected to provide any narrative statements identifying key issues affecting the metrics for each time period.

Contacts:

WR-ALC/LYP: Col. Wallace ("Skip") A. Collins, Avionics Production Division (To be contacted)
LYPO: Maj. Timothy Nesley, Chief, Avionics Production Division Chief
LYPF: Mr. James Roeder, Chief, F-15 Branch, Avionics Production Division
LYPM: Mr. Jimmy Beeland (To be contacted)
LGSH: Ms. Mary Anne Schubert, Depot Supply Home Office Chief (To be contacted)

C. Units of Analysis

Depot-maintenance provides two types of component repair services: the repair of all unserviceable end-items generated internally within the depot (or in other depots) through the programmed depot maintenance of aircraft (e.g., F-15, C-5, KC-135, etc.) *and* the repair of unserviceable end-items flowing to depot-maintenance from the operating bases. This pilot project focuses on the latter set of components or end-items, which account for much of the repair work done at the depots. The internal demand for component repair services is typically fairly small, predictable, and constant. Hence, internal demand tied to programmed aircraft depot maintenance is assumed to remain invariant during the pilot project and is not explicitly addressed in calculating the effects of the new metrics on the outcome variables, although they are included in analyses of enabling metrics.

Also, depot maintenance, through its various supply organizations, also acquires new spares, such that the flow of serviceable items to the bases from the Consolidated Serviceable Inventory (CSI) at the depots could consist of as mix of repaired items and new spares. Since the key issue is the supply of serviceable end-items to the operating bases, no clear distinction is drawn here between repaired items and new spares, where the latter play a basically small role in the provision of serviceable items to the bases. More to the point, the intent here is to hold those responsible for the provision of new spares and those responsible for repair services to be held jointly accountable for the outcome of their efforts, which are, by definition, linked together. This treatment should preempt the refrain often heard that satisfying the customer's needs is not solely the responsibility of depot-repair and that any failure in providing the needed serviceable items should be levied upon the supply chain manager.

The units of analysis in the pilot project refers to specific pre-selected components (end-items; line replaceable units – LRUs). Two categories of end-items are defined for analysis: MICAP avionics end-items (those components that are in particularly critical demand, based on past data on both MICAP incidents and MICAP hours) *and* supportable avionics end-items (those components that can be repaired to meet normal customer demand for serviceable items).

The end-items in these two categories reflect the proposition that MICAPs and backorders are the items most required by the warfighter on the flightline. The first five end-items that are selected represent the top five MICAP end-items during the period of July 2001 to June 2002. In addition to these specific end-items, OO-ALC/MALA wants to monitor a particular Shop Replaceable Unit (SRU), the Low Noise Amplifier (LNA), which is a major part of the MLPRF end-item and which represents particularly serious supportability problems. The last six end-items are deemed "supportable" end-items recommended for analysis by OO-ALC/MALA. "Supportable" means the shop has available carcasses, the required production capacity, and needed parts and funding for repair.

<u>Group of MICAP End-Items (End Items Repaired at Ogden ALC)</u>

1. *Radio Frequency Shop*

- MLPRF –1270-01-233-0011WF – Modular Low-Power Radio Frequency

- Antenna – 5985-01-212-2950WF
- LPRF – 1270-01-102-2962WF – Low-Power Radio Frequency

2. *Displays and Indicators Shop*

- MFD – 6625-01-193-8861WF – Multi-Function Display

3. *Microwave Shop*

- DMT – 1270-01-238-3662WF – Dual Mode Transmitter
- LNA – 1270-01-132-6867WF – Low Noise Assembly (SRU)

<u>Supportable End-Items</u>

1. *Displays and Indicators Shop*

- HUD EU – 1270-01-468-8658WF – Heads-Up Display Electronic Unit
- MSC – 5826-01-052-1945NT – Mode Select Coupler

2. *Computer and Inertial Shop*

- DFLCC – 6615-01-448-6152WF – Digital Flight Control Computer
- RGA – 6615-01-042-7834WF – Rate Gyro Assembly

3. *Processor and Pneumatics Shop*

- JRIU – 5998-01-080-3978WF – Jettison Remote Interface Unit
- MRIU – 1290-01-109-1499WF – Missile Release Interface Unit

These predetermined end-items in both categories will be matched with similar end-items repaired at the WR-ALC avionics production shop.

Although the pilot project will focus directly on these pre-selected individual end-items to measure the effects of the new metrics, data on all other MICAP and backorder items (as two broad groups) will also be developed to test for any positive or negative effects of the new metrics on these two groups of end-items. The specific question of interest here is to see whether depot-repair, by concentrating on the repair tasks related to the pre-selected items -- because they are being directly monitored in the pilot project -- may end up placing less emphasis on other end-items.

D. Time Period

The time period for the first phase of the pilot project will be one calendar year. It is important to give the participating organizational entities sufficient lead time prior to the initiation of the pilot project so that they can review and approve the pilot project implementation plan and put in place preparatory steps for launching the pilot project. Also, the pilot project should be reviewed at the conclusion of the first six-month period in order to make any mid-course corrections in executing the pilot project, should that be considered desirable.

At the conclusion of the pilot project, the results will be evaluated and plans for further pilot projects will be developed and implemented, as appropriate. One option might be to evaluate the effects of an expanded treatment regime to include both new metrics and the introduction of key lean practices into depot repair operations and the supporting supplier base.

VI. NEW METRICS AND RELATED DATA REQUIREMENTS

This section focuses on the new metrics and related data requirements. The latter include data on confounding factors, control variables and other measures of interest, as well as data on existing metrics. Both pretest and posttest data will be collected for the new metrics, confounding factors and control variables on a weekly (if possible) and monthly basis. Similarly, data will continue to be collected and made available on the existing metrics currently being used by the Air Force in managing the sustainment system. Such data are required for conducting a rigorous evaluation of the validity and benefits of the new metrics. Although the collection of the required data may entail additional investment in the near term, the anticipated results should provide significant benefits far outweighing any incremental near-term data collection costs.

A. New Metrics

New metrics are categorized into two main groups: *outcome metrics* and *enabling metrics*. *Outcome metrics* represent, for the purposes of the pilot project, delivery of value to the customer. They are customer satisfaction metrics. They gauge how effectively the customer's urgent and normal needs are met, how well they are met (e.g., in terms of product quality, customer-wait-time), and how cost-efficiently they are met (e.g., average cost of repair). *Enabling metrics* gauge how well the sustainment system performs various processes, functions and practices – by making sure that the organization is doing the "right job" as well as doing the "job right" -- to deliver the defined outcome metrics benefiting the customer,

1. Outcome Metrics

The concept of outcome metrics, derived from the idea of delivering value to the customer, can be further generalized to encompass the creation and delivery of value to all stakeholders. [9] In this broader conceptualization, stakeholders would include, for instance, the Department of Defense and the Air Force at the "corporate" level, in the sense of delivering efficient, reliable, and responsive maintenance and repair services. Stakeholders would also include the warfighter, the workforce, and the supplier network, all linked together through the construction of a robust value proposition serving as the basis for creating and delivering value.

For the purposes of this pilot project, a narrower definition of the "customer" is adopted, focusing directly at the "bases." It is assumed that improving customer satisfaction at the base-level will simultaneously improve the delivery of value to other stakeholders as well (e.g., to the warfighter,

[9] For a discussion of a framework for value creation and delivery at the enterprise level (e.g., program enterprises, multi-program enterprises, the US aerospace enterprise) refer to Earll Murman, *et al*., *op cit*.

to the Air Force at the "corporate" level). Nevertheless, to demonstrate how the new metrics can lead to improvements in the delivery of value to these stakeholders, the pilot will provide a link between *incremental* improvements in depot-repair performance levels and resulting *incremental* improvements in flightline performance metrics (e.g., fully-mission-capability rates).

The following customer need satisfaction metrics, which also represent the new *outcome metrics*, are identified (for each targeted end-item to be analyzed in the pilot project):

- ***Urgent customer requirements satisfaction rate (UCRSR):*** This is defined as the ratio of total number of serviceable end-items provided[10] during a given period *to* total number of high MICAP items as defined at the base level.

- ***Normal customer requirements satisfaction rate (NCRSR):*** This is defined as the ratio of the total number of serviceable end-items provided during a given period *to* total number of backorders issued by bases.

- ***Weighted customer requirements satisfaction rate (WCRSR):*** This is defined as a weighted average of UCRSR and NCRSR. The assigned weights signify the relative importance of each and are numerically estimated by quantifying the relative importance that is implicitly attached to each, based on previously negotiated quantities through the "Quasi-EXPRESS" experiment. These weights are taken as a first-approximation of the true weights that could be estimated, if true customer (warfighter) needs were known. Neither EXPRESS nor the "Quasi-EXPRESS" approach is likely to serve as the most reliable (true) measures of actual customer demand. In the absence of such "true" estimators of actual customer demand, the "Quasi-EXPRESS" results are taken as second-best "true" estimators. In the case of the WR-ALC, to the extent that induction decisions are driven by EXPRESS, previous inductions of MICAP items and backorder items can be used to derive the "revealed" weights attached to each category of customer demand.

- ***Cost:*** This metric is designed to measure the actual (incurred) cost of the maintenance and repair services provided for a given end-item by serial number. This requires the determination of the actual full costs of both materials and labor, as well as the cost of utilization of capital (plant and equipment) which is typically built into the prevailing overhead rates. For the purposes of this metric, it would suffice to use the prevailing (already negotiated) direct labor, overhead and general and administrative (G&A) rates. However, in computing labor hours, it is important to capture the total number of actual labor hours incurred rather than standard hours. Moreover, actual labor hours should include accumulated labor hours for repairing a particular end-item across all shops, including labor hours required to repair the SRUs embedded in that end-item.

  During the pilot project, it is assumed that for each serviceable end-item the customer will continue to be charged at prices reflecting previously-negotiated rates, even though the "actual

[10] The word "provided" is used here deliberately, to convey the idea that the supply of serviceable end-items from the Consolidated Serviceable Inventory (CSI) at the depot to the operating bases consists of both repaired items and new spares. The upshot of this is that the metric is chosen to foster closer cooperation between the "fixer" and the supply organizations, to help integrate investment decisions concerning the procurement of new spares more closely into the depot-repair process and, hence, into the resupply pipeline to meet customer needs on a timely basis.

cost" figures computed as just noted might be somewhat lower, reflecting possible efficiency gains due to the introduction of new metrics. Comparing the "business-as-usual" prices charged to the customer reflecting previously-negotiated rates *with* the actual incurred costs during the pilot project, where the latter might entail lower prices due to possible expected efficiency gains, would help determine answers to the following two questions: (a) would the pilot project result in any cost savings to the customer; and (b) what would be the magnitude of such cost savings?

- ***Product quality**:* This metric measures the percent of the total number of serviceable end-items produced by depot-repair (and shipped to CSI for distribution to the operating bases) during a given period that are found to be defective. There are two ways of computing this metric for a particular end-item. The preferred way would be to define the *denominator* as the total number of that end-item "produced" by depot-repair during a given period and shipped as a serviceable item to Consolidated Serviceable Inventory (CSI). The numerator would be slightly complicated to compute. It would involve tracking each one of those serviceable items by serial number and count how many of them had Quality Deficiency Reports (QDRs) filed against them, upon testing at the base-level, and then shipped back to depot-repair. Next, it would be necessary to count how many of those QDR items re-tested by depot-repair were actually found to have quality defects. This final number of items found to have defects – whatever type or number of defects within each end-item tested – would then represent the numerator.

  A less cumbersome but also less accurate way would be to simply take the total number of a given end-item produced during a given period, as the denominator, and take, as the numerator, the number of that same end-item (with a QDR attached to it) which is re-tested at depot-repair during the same period and found to be defective. Another way of computing the numerator, for a given end-item, is to count the total number with a QDR designation re-tested at depot-repair *minus* the total number that is designated as No Fault Found (NFF).
  Of course, in using this second method, the serial numbers of the end-item in question that make up the numerator will not match the serial numbers of those that make up the denominator. This is another way of saying that the time-profiles of the serial numbers in the numerator and in the denominator are quite different. This may distort the quality metric that is computed, particularly when depot-repair production during a particular accounting period has dropped precipitously (e.g., in July, 2002) compared with production levels during previous periods generating those very same quality defects.

- ***Customer Wait Time:*** This metric measures the total elapsed time in hours (including pick, pack and ship time, total transportation time including any in-transit wait time, unpacking time) for each pre-selected end-item from the moment a requisition for it is issued by the bases until the time a serviceable end-item is received at base supply and is, in fact, available to base maintenance upon request. Total elapsed time is used to encourage better coordination along the pipeline to minimize the time it takes to provide the needed serviceable end-item to base supply so that it is available to base maintenance or for delivery straight to the flightline. In this definition, "customer" is defined, for all practical purposes, as "base supply." It is important to note that this definition differs from that just adopted by the Air Force sustainment system which measures "customer wait time" as the total elapsed time between when base maintenance requests a serviceable end-item and when base supply makes that item available for use. It can be argued that improving the "customer wait time" as proposed for the pilot project would, in fact, nullify the need for the second definition.

The implementation of these new *outcome metrics* (customer satisfaction measures) includes the concurrent adoption of a number of new operating rules. These new operating rules will include, but not be limited to, the following:

- The F-16 Avionics Production Branch (OO-ALC/MALA), working through the relevant Shop Service Centers (SSCs), will be able to place orders for repair parts and materials ahead of induction, for the pre-selected end-items that are being directly studied in the pilot project.

- OO-ALC/MALA, LGF, LGS, and DLA will work jointly through meetings on a weekly basis in planning repair requirements, procurement of the necessary repair parts and materials, and any pre-kitting activities in support of the MRO operations related to the pre-selected end-items.

- LGF and LGS, working in concert with the relevant SSCs, will be authorized to acquire the necessary repair parts and materials directly from commercial suppliers when and if DLA is considered unresponsive to the data availability needs at depot-repair on a timely basis.

- LGF, LGS and DLA, working together in close cooperation, will be authorized to streamline contracting processes for the repair parts and materials required for the pre-selected end-items, in order to bring under contract outside commercial suppliers in the shortest time possible; they will further be authorized to enter into long-term partnerships and strategic alliances with selected suppliers.

- DLA will be authorized to acquire repair parts and materials for any of the pre-selected end-items without having to apply the Economic Order Quantity (EOQ) method, to ensure that the required repair parts and materials are made available to depot-repair when needed on a timely basis. Further, DLA will refrain from the practice of eliminating stocks of parts and materials not used during the previous two years for the specific pre-selected end-items.

- The F-16 Avionics Production Branch will be waived from having to show "sales" benefits for a number of the pre-selected end-items, chosen at random, to allow all shops to concentrate on improving their performance on these particular end-items rather than on covering its costs. This is expected to motivate depot-repair to strive to achieve system-wide optimization rather than local (sales) optimization. The implementation of this operating rule will be important in testing two of the main hypotheses outlined earlier.

It is expected that improvements in these *outcome metrics* (customer satisfaction measures) will impact the warfighter directly: (a) by leading to a *reduction* in TNMCS (Total Not mission Capable for Supply), through improvements in base-supply of serviceable items, that should directly help to increase the FMC (Fully-Mission-Capable) rates; and (b) by promoting greater efficiency in base-repair operations – through reduced cannibalization in light of greater availability of serviceable items, such that the available resources can be put to more productive pursuits – which would tend to reduce the TNMCM (Total Not Mission Capable for Maintenance) rates and thereby increase the FMC rates.

2. Enabling Metrics

Enabling metrics refer to those few pivotal metrics that drive an enterprise's overall effort at various levels towards the achievement of its overarching goals and objectives. These metrics are operative at multiple levels and across organizational processes and functions as well as across organizational boundaries. They are designed to motivate people to strive to achieve global optimization across the enterprise. Well-designed enabling metrics foster a culture that both motivate and reward efforts for doing the "right job" as well as for doing the "job right." This means not only making the right choices but also allocating the right types and levels of effort to performing the selected projects or tasks. Most metrics only concentrate on "doing the job right" (e.g., at minimum cost, eliminating waste), losing sight of the organization's central mission to provide the right types of products and services necessary for delivering value. Well-designed enabling metrics also overcome short-term orientation, risk aversion, parochialism, "not-invented-here" syndrome, and related behavioral traits that undermine the enterprise's efficiency and effectiveness. Among the enabling metrics, some can be identified as "gatekeeper" metrics in the sense that they exert significant leveraging influence on the outcome metrics.

For this pilot project, the following enabling metrics are proposed:

- ***End-item shopfloor flow time**:* This metric is already being tracked by the Air Force sustainment system. Of greater interest here, however, is tracking the *variance* in shopfloor flow time for the pre-selected end-items. Variance, a statistical concept, measures variability. The idea is to encourage the depot repair process to minimize variability. This is a central idea from Six-Sigma thinking and practice; it is also historically an integral part of lean thinking and is an important driver continuous quality improvement[11]. The depot repair process inherently exhibits considerable variability that is arguably greater than that observed in normal manufacturing operations. Still, such a metric motivating continuous reduction in variability would foster more standardized workflow processes, improved worker training, self-inspection, and other desirable lean practices. The end result of reducing variability is

---

[11]Six Sigma, which has its roots in the application of probability theory to statistical quality control, has widened its scope in recent years to encompass an integrative management tool for achieving continuous improvement across the entire enterprise. The relationship between Six Sigma and lean thinking can be brought into sharp relief in the context of production operations: while Six Sigma stresses *quality improvement* through elimination of all sources of variation, lean thinking concentrates on *speed* through continuous *defect-free flow* across the entire enterprise value stream. Viewed at the enterprise level, Six Sigma is an important enabler of lean thinking. For a comparative discussion of lean thinking and Six Sigma, see Kirkor Bozdogan, "Lean and Six Sigma: An Overview," Draft Working Paper, Massachusetts Institute of Technology, Lean Aerospace Initiative (July 24, 2002), 10 pp.

For further information on lean thinking, see, for example, Murman, et al., as well as James Womack, Daniel Jones and Daniel Roos, *The Machine that Changed the World: The Story of Lean Production* (New York: Rawson Associates, 1990; and James Womack and Daniel Jones, *Lean Thinking* (New York: Simon & Schuster, 1996).

For further information on Six Sigma, see, for example, Mikel J. Harry and J. Ronald Lawson, *Six Sigma Producibility Analysis and Process Characterization* (Reading, MA: Addison-Wesley Publishing Company, 1992); Peter S. Pande, Robert P. Neuman and Roland R. Cavanagh, *The Six Sigma Way* (New York: McGraw- Hill, 2000); and George Eckes, *The Six Sigma Revolution* (New York: John Wiley & Sons, Inc., 2001).

expected to be shorter mean shopfloor flow time, which would translate into increased capacity, enhanced capability, higher productivity and greater throughput.

- ***Cost of maintenance and repair****:* This metric, as well, is currently being tracked, most likely in the form of "standard" cost, reflecting previously-negotiated rates. The main interest in the pilot project is in "actual" incurred cost, as noted earlier in connection with the discussion concerning outcome metrics. Of greater interest here is the variability (variance) in the actual cost of performing maintenance and repair services for each end-item. Reducing variation in the cost of repair services would result in lower average costs, as many Six Sigma enterprises have already found out. Also, such a metric placing emphasis on reducing cost variation would foster the adoption of basic lean practices for identifying and eliminating all sources of waste (e.g., through value stream mapping and analysis; standard work; preventive maintenance; 6S techniques; mistake proofing; cross-training; *kaizen* events).

- ***Productivity:*** This metric measures the total number of a given Line Replaceable Unit (LRU) or end-item (in physical units, not sales) produced by depot-repair per 100 labor-hours (or per 1000 labor-hours) utilized by all workers to repair that end-item, from that end-item's initial induction into the shop to its shipment as a serviceable item to Consolidated Serviceable Inventory (CSI). In computing this metric, it is important to accumulate the total number of labor hours actually incurred across all shops in connection with the maintenance and repair of a particular end-item. This would include labor hours used in repairing all shop-replaceable units (SRUs) embedded in a given end-item. An alternative, acceptable, way of capturing this metric would be to express it as the total (cumulative) number of labor hours utilized to produce one unit of a particular end-item.

  This metric is not intended to gauge the overall quality of the effort being exerted by depot-repair but rather to encourage more efficient allocation of the available resources (both labor and capital) to the performance of the tasks at hand. There is a rich history in economics addressing the issue of measuring productivity of firms, industries and countries as the ratio of output *to* all types of factor inputs employed (labor, capital, materials) – by using "total factor productivity" measures. The focus here is on a fairly simple "single factor" productivity measure. Thus, for the purposes of this pilot project, a basic measure of "labor productivity" is considered sufficient, while controlling for the "up-time" availability of the testing equipment.[12] The number of labor hours actually employed is not adjusted for unscheduled work-stoppages due to equipment failure in order to motivate depot-repair to put greater emphasis on preventive maintenance.

- ***Responsiveness:*** This metric measures the degree to which depot-repair is responsive to objectively-determined, resource-unconstrained, customer demand that it is supposed to meet.[13]The metric is measured as the ratio of the *required Takt time* to the *observed Takt time*. Takt time is a concept from lean production indicating the desired tempo at which the production system should operate to meet the pace of customer demand. For example, if depot-repair must produce a serviceable end-item of a particular type every 15 hours in order

[12] This is done by introducing a control variable which gauges, for the pre-selected end-items, the depot-repair wait-time due to unscheduled downtime in testing equipment as a percent of total shopfloor flow time.
[13] In this discussion, contract repair is assumed to be an extension of depot-repair. Thus, the responsiveness metric is intended to capture the combined responsiveness of both "organic" depot-repair and contract repair.

to meet the customer's rate of demand for this end-item, this is the required Takt time. If the actual or observed Takt time is instead 30 hours, then depot-repair responsiveness is 0.50 (i.e., 15 hours ÷ 30 hours).[14] A plausible interpretation of such a responsiveness rate is that the production facility is simply not being sufficiently responsive to the tempo of customer demand for the particular product in question. This could be the result of a combination of factors, such as lack of visibility into the volume and time-profile of customer demand, sheer inefficiency in production operations, and presence of barriers to greater efficiency. The production facility can improve its responsiveness only by tackling these factors head-on.

The responsiveness metric as defined is important because it gauges how closely depot-repair is operating with reference to the required (or target) Takt time. The gap between the target Takt time and the currently observed Takt time can be explained by a combination of factors, including resource constraints, the various policy and regulatory barriers under which depot-repair operates, and funding constraints that it faces.

The computation of the *required Takt time* for depot-repair calls for a definition of the true customer demand for depot-level component repair. This can be estimated by summing up total demand from the various sources. One source of demand, for instance, is the demand associated with aircraft brought into the depot for either programmed depot maintenance or upgrade. Another source of demand is foreign countries which have fleets obtained under foreign military sales (FMS). The largest source of demand is base-level demand, which encompasses the operating combat units (e.g., Air Combat Command- ACC; Pacific Air Forces –PACAF; United States Air Forces Europe – USAFE; Air National Guard – ANG; Air Force Reserve --AFR), as well as a number of special units (e.g., Air Force Materiel Command – AFMC; Air Education and Training Command– USAFE). Total demand from the bases, for instance, can be estimated by considering the total number of end-items that are designated as "Not Repairable This Station" (NRTS) at the base-level.[15] This represents the total volume of unserviceable (or reparable) end-items depot-repair is expected to turn into serviceable end-items for shipment back to the bases, assuming that base maintenance will be fully able to repair all "Repairable This Station" (RTS) items, for which the necessary repair parts and materials would be made available by depot supply.

---

[14] How such a measure can be computed for a given product can be quickly illustrated as follows. Suppose the total scheduled production time for a facility during a given month is 280 hours (i.e., (2 shifts/day) x (7 effective working hours/shift) x (5 days/week) x (4 weeks/month)). If customer demand (rate of consumption) is 14 units per month, then the *required Takt time* is 20 hours (i.e., (280 hours/month) ÷ (14 units/month)). Suppose, however, that the production facility is actually producing only 7 units/month, rather than 14 units/month. Then, the *observed Takt time* is 40 hours/unit/month (i.e., (280 hours period) ÷ (7 units/month)). Thus, the ratio of the *required Takt time* to the *observed Takt time* is 0.50 (i.e., (20 hours/unit/month) ÷ (40 hours/unit/month).

[15] An alternative way of gauging true demand is to count the total number of MICAP items and backorder requisitions received by the depot from the operating bases. The resulting total should approximate the total number of NRTS items. This follows from the fact that at the base-level each unserviceable asset is either designated as "Not Repairable This Station" (NRTS) and sent to depot-repair, or kept for repair by base maintenance, or condemned by base supply. Normally, each unserviceable asset is turned into base supply in exchange for an available serviceable asset and base supply originates a requisition for each such unserviceable asset. Thus, the total number of requisitions (for both MICAP items and backorder items) should approximate the total number of NRTS items.

If depot-maintenance -- including contract repair support by commercial providers-- keeps turning out the requisite number of serviceable end-items, base-supply will have a steady-state resupply of these serviceable items, ensuring sustained FMC rates reflecting prevailing norms of readiness. In reality, of course, not all of the MICAP and backorder items get repaired, primarily because of funding constraints. Further, of those that are inducted into depot-repair, some may have higher priority rating than others and are given higher priority attention, and others may have to wait there for long periods of time due to lack of availability of the needed repair parts and materials (i.e., these items become designated as AWP-G, assets waiting for repair parts for which backorders have been placed).

Under such a scenario, the depot-maintenance system would simply not be able to supply the required number of serviceable items, especially while it is underfunded to invest in new spares to replace condemnations. Then, there would result, by definition, a deficit of serviceable end-items. As long as such a deficit continues, the combat forces would experience a degradation in terms of their FMC rates. Such a scenario would seem to approximate pretty closely the current state of affairs. Thus, in computing the observed current Takt time, total demand can be taken as the demand resulting form the "Quasi-EXPRESS" negotiation process currently being used at the Ogden ALC (i.e., "negotiated demand"). It is important to note that this "negotiated demand" would, in all likelihood, differ, perhaps significantly, from the actual or real demand. The reason is that the "negotiated demand" reflects not the real needs but what is feasible to provide in light of the existing funding and other resource constraints.

- ***End-item repair parts combined fill rate:*** This metric measures the repair parts and materials fill rate for a given end-item, taken together, rather than for individual parts needed for repairing that end-item. The importance of such a metric can be demonstrated by using a simple example. Suppose, for instance, that a given end-item requires parts A, B, C and D in order to repair it. Suppose, further, that the issue effectiveness rate (or stockage effectiveness rate) for the individual parts are, respectively, 0.80, 0.90, 0.85, and 0.95, such that the supply organization can make the claim that its performance is very good to excellent, under the circumstances. A more objective view, however, would cast doubt on such a claim. The reason is that the joint probability of the availability of these parts, taken together, to depot-repair at a given point is, roughly speaking, more like 0.60, which is a far cry from anything that depot-repair should find acceptable. It can be seen, in somewhat more technical terms, that if the respective probabilities of the the availability of the required parts are P_A= 0.80, P_B= 0.90, P_C= 0.85, and P_D= 0.95, then the joint probability of the combined availability of all four parts taken together at a given point is the intersection (multiplication) of the respective probabilities (i.e., (0.80)x(0.90)x(0.85)x(0.95)= 0.58).

  In the above example, if parts A, B, C, D are in fact frequently failing parts, then having all four available to depot-repair would lead to a significant improvement in the efficiency and effectiveness of the depot-repair process. To the extent that these parts are currently supplied by separate supply organizations, such a combined fill-rate metric would provide a strong incentive for them to cooperate. A particularly desirable outcome would be if they synchronize their activities and provide pre-kitting of the required repair parts and materials.

This metric is similar in concept to the "Equipment Repair Order" (ERO) fill rate metric introduced by Fricker and Robbins for employment by the US Marine Corps.[16] The ERO fill rate metric, which is limited to critical repairs, measures the percentage of all critical repairs that receive all of their high-priority parts from local supply. Fricker and Goodhart, using actual Marine Corps data, have found through simulations that employing the ERO fill rate metric, along with two new techniques, can lead to significantly improved supply system performance.[17] The metric proposed here for the pilot project differs from the ERO fill rate in the sense that it is focused on all or most frequently failing parts for each individual end-item, rather than focusing on specific repair tasks that are considered critical repairs.

- ***Supplier delivery performance:*** Currently the Air Force sustainment system lacks adequate measures of supply chain performance. This metric is proposed as a first step to gauge the performance of the supplier network supporting the sustainment system. The metric quantifies the "order-to-delivery" time for repair parts and materials obtained from suppliers, by using three related performance measures: elapsed time required from order-to-delivery for 50% of all repair parts and materials requisitioned; elapsed time required from order-to-delivery for 75 % of all repair parts and materials requisitioned; and elapsed time required for order-to-delivery of 95 % of all parts and materials requisitioned.

B. Confounding Factors

Confounding factors are those variables that may be correlated with the outcome variables -- or that may covary with the outcome variables -- such that there is the risk of attributing the observed outcomes to the treatment itself when in fact they may have been caused by these other factors that change in the same direction as the treatment factors. A list of the confounding factors that are of particular interest in this case study is given below.

- ***Burn rate:*** This metric measures the total amount of dollars actually spent on depot-repair during a given period. It is conjectured that higher levels of spending can lead to increased numerical values for the outcome variables, where such effects may be confounded with the treatment effects. Even though the burn rate, as defined, relates to all end-items going through depot-repair, it is reasonable to expect that higher levels of aggregate spending may also well affect the performance of depot-repair on the specific groups of end-items being studied.

- ***Labor hours:*** All else remaining equal, allocation of a greater number of labor hours, while keeping constant the capital stock (e.g., testing equipment capacity) may well influence the outcome variables. This metric is defined as the total number of labor hours (both normal shifts and overtime) actually allocated during a given period, computed separately for the LRU shops, the SRU shops, and the combined total for the entire facility. It is conjectured that the allocation of a greater level of labor effort in the aggregate could influence the levels of effort more specifically directed to the end-items being studied, thus affecting the outcome metrics for these end-items.

---

[16] Ronald D. Fricker, Jr., and Marc Robbins, *Retooling for the Logistics Revolution: Designing Marine Corps Inventories to Support the Warfighter* (Santa Monica, CA: RAND, MR-1096-USMC, 2000).
[17] Ronald D. Fricker, Jr., and Capt. Christopher A. Goodhart, "Applying a Bootstrap Approach for Setting Reorder Points in Inventory Systems," *Naval Research Logistics*, Vol. 47, No. 6 (September 2000).

- ***Testing rework:*** One of the ways in which the operating bases can influence the efficiency and effectiveness of depot-repair is the extra workload they may cause depot-repair to undertake, thus resulting in misallocation of scarce depot-repair resources. An example is the practice of testing "serviceable" end-items shipped to base supply from CSI, after they are produced by depot-repair, which are found to have discrepancies. These items are then sent back to depot-repair, where they are re-tested. Typically, in a majority of the cases, depot-repair classifies these re-tested items as "No Fault Found," which means that the depot is unable to replicate the discrepancies found at the base-level and that these items are indeed serviceable. To the extent that depot-repair, compared with the bases, has older testing equipment or test equipment with different capabilities, such discrepancies may be unavoidable. Nevertheless, a considerable amount of scarce depot-repair resources are claimed by the end-items that are re-tested. Regardless of the testing results (i.e., No Fault Found or quality defects found), the amount of this workload can have a serious impact on the overall efficiency and effectiveness of depot-repair. An increase or decrease in the total number of end-items that are re-tested at depot-repair during a given period can impact the overall performance of depot-repair in terms of the outcome variables. Therefore, a testing rework factor is proposed as a confounding variable. For the purposes at hand, this factor measures the total number of re-tested end-items as a percent of all end-items inducted into depot-repair.

C. Control Variables

Control variables refer to specific characteristics, situational factors, policies and procedures and other attributes that differentially affect the groups, or individual units of analysis, that are being studied in the experiment. For example, two comparison groups may differ sharply in terms of the age composition, educational attainment, skill levels, or unionization of the workforce. One group may have much more modern testing equipment than the other, or may be comparatively more advanced in using information technologies. Such differences might have differential impacts on the performance of the two comparison groups. The objective is to equalize the two groups in terms of these observed differential characteristics by holding the most important of these factors constant or as invariant as possible. A list of the control variables that are of particular interest in this pilot project is given below.

- ***Information access:*** The Depot Repair Information Local Server (DRILS) system now operational at the Ogden ALC avionics repair facility provides direct visibility into the repair history of specific end-items inducted into repair as well as repair parts usage for each end-item. For example, the DRILS system can be used in troubleshooting particular serial numbers on all end-items. By creating a serial number history, "red-flags" are put in place for certain serial numbers that are inducted, on average, more often than others. This helps to identify continuing problems with certain repair parts and manufacturers, for instance

  DRILS is not yet available for use at the Warner Robins avionics facility. Thus, this control variable is proposed to account for the apparent differences between the two repair sites, in terms of information availability and access, that would have a differential impact on their respective performance. To control for these differences, the following "information access rating system" is offered:

Level 1: No ready access to prior repair history and parts usage for any given end-item for any backshop.
Level 11: Only hardcopy/logbook access to any given end-item for any backshop.
Level III: Real-time, technician-entered, electronic access for a limited number of end-items (< 25%) for any backshop.
Level IV: Real-time, technician-entered, electronic access for a majority (25% to 85%) of end-items for all backshops.
Level V: Real-time, technician-entered, electronic access for all end-items for all backshops.

- ***Test equipment availability:*** This variable controls for any differences between the two repair facilities in terms of the availability of test equipment for performing the required testing tasks. It is defined as the total number of hours a given end-item has to wait of wait time for testing services due to unscheduled equipment downtime as percent of the total shopfloor flow time for that end-item.

- ***Workforce experience:*** This measure is intended to control for any differences between the comparison groups in terms of the level of experience of the technical workforce. The metric proposed here is the average number of years of on-the-job work experience of the technical workforce in the LRU shops, in the SRU shops, and for the facility as a whole.

- ***Worker training:*** This measure is intended to control for any differences between the comparison groups due to different levels of exposure to basic lean principles and practices. It is measured as the cumulative percentage of all workers who have had at least two days of formal classroom instruction and training in basic lean practices. These practices encompass, for example, kanban-based "pull" systems for just-in-time (JIT) production, value stream mapping and analysis, 6S (Establishing Visual Order), Kaizen events, standard work, Takt time, load leveling, mistake proofing, root cause analysis, single-piece flow, cellular design, total preventive maintenance, process control (including Six Sigma principles) and related lean methods.

- ***Cellular production:*** This measure represents another variable for controlling for any differences between the comparison groups in terms of their use of lean practices. This particular measure is chosen as an illustrative case of lean implementation. It is defined as the percentage of all SRUs that are repaired by using cellular manufacturing principles and methods. It is conjectured that any improvements in SRU repair will have a wider leveraging effect on the LRU repair tasks, since a considerable part of the challenge in LRU-related repair involves the timely availability of serviceable SRUs from the SRU shops. It is hence thought that, all else remaining equal, lean deployment in connection with SRU-related repair tasks can be taken as a fairly reliable gauge of the extent to which lean practices have been adopted in avionics MRO operations.

- ***Resource constraint:*** This metric is intended to account for any significant differences between the two comparison groups in terms of the level or severity of the resource constraints they may be facing in the course of the pilot project. For example, if one site is much more resource-constrained than the other, than their respective performance levels need to be adjusted for such differences. One measure of the degree of resource constraint faced by each site is the percent of all reparable end-items at Consolidated Repairable Inventory (CRI) that is not inducted into depot-repair due to any number of resource constraints (e.g., funding limitation, lack of available parts, shop capacity).[18] For example, if there are 100 NRTS items at CRI of which 5 had to be condemned, 95 end-items are available for induction into depot-repair. Of these, let us assume that 20 are not inducted, due to a combination of resource constraints. Then, the numerical value of the resource constraint measure proposed here would be 0.21 (i.e., 1.0 – 75/95).[19] It is conjectured that a higher constraint factor would have a larger adverse impact on performance, all else remaining equal. Hence, this control variable is intended to equalize the performance of the two comparison groups for any observed differences in terms of the severity of constraints under which they may be operating.

- ***Induction policies:*** Differences in induction policies may impact the performance of the two sites in different ways. The Ogden ALC is following a "Quasi-EXPRESS" process reflecting a negotiated production level and an induction policy supporting the production of the negotiated output levels. This means substantially reduced utilization of the prioritization scheme determined by EXPRESS on a daily basis. Meanwhile, the Warner Robins ALC may still be operating in accordance with the prioritization process provided by EXPRESS. This control variable is intended to account for such differences in induction policy. A proposed way of quantifying this factor is the total number of end-items inducted into depot-repair bypassing EXPRESS as a percent of all end-items inducted into depot-repair.

D. Other Data Requirements

The successful execution of the pilot project will require the following additional data:

- ***Current metrics:*** Data on all currently used metrics, for all end-items, will continue to be developed and will be made available for use in the pilot project.

- ***Cannibalization and robback data:*** Data on both cannibalization and robback practices will be required for each of the selected end-items being studied in the pilot project to gauge the effects of new metrics on the extent to which cannibalization and robback practices. Data will be required on both cannibalization and robback *incidents* and *hours*.

As noted earlier, such additional data will be needed for analytical purposes only during the pilot project to validate the new metrics. This should help save excessive data collection costs in the future.

[18] The Supportability Module of EXPRESS, which is the central information and decision-support system used by the Air Force sustainment system for the prioritization of depot-repair tasks as well as for the distribution of serviceable items to the operating bases, contains a Supportability Module which operates on all EXPRESS-determined priority repair requirements. The Supportability Module applies effectively a screening process testing for the availability of four types of resources in the following order: carcasses, repair dollars, component parts, and shop capacity. If an end-item on the EXPRESS priority list fails to meet any of these resource availability tests, it is not inducted into repair. For a more detailed description, see Maurice Carter and Ronald W. Clarke, "EXPRESS Planning Module," *Air Force Journal of Logistics*, Vol. XXIII, No. 4 (Winter 1999), 23-26.

[19] The total number of reparable end-items that can be inducted into depot-repair, as defined here, differs somewhat from that calculated in EXPRESS. The total number of carcasses available for induction is estimated in EXPRESS as the sum of all carcasses on-hand (i.e., NRTS items already at CRI *less* those condemned) *plus* those in the in-transit pipeline to the depot *plus* those that are expected to be sent from the operating bases to the depot over the planning period. See Carter and Clarke, *Ibid*., p. 25.

VII. PROJECT STRUCTURE

The pilot project operational team consists of those involved with avionics production at both OO-ALC and WR-ALC and, more broadly, encompasses all of the participating organizations. In fact, the pilot project is expected to be "owned" by the principal participating organizations, which will be responsible for its execution based on agreed-upon ground rules and operating procedures. The government support staff at OO-ALC and WR-ALC are expected to help collect the metrics and related data required for the pilot project. The role of the MIT LSI researchers will be to provide technical support and data analysis. It is expected that the participating organizations will be kept abreast of on-going pilot project outcomes based on analyses of the data being generated.

The participating organizations are encouraged to examine this proposed planning document closely to make sure that the various metrics and related data requirements are clearly and correctly specified. They are further encouraged to offer any improvements that will enhance the quality and usefulness of the pilot project.

Finally, the participating organizations are encouraged to organize themselves in any way that would enhance the success of the pilot project. This might include, for instance, inclusion of DLA personnel, HQ AFMC/LG personnel, representatives of MAJCOMS and other potential stakeholders in the execution and "ownership" of the pilot project.

VIII. RESOURCE REQUIREMENTS

Resource requirements for the implementation of this pilot project (e.g., people, travel, related expenses) would include costs associated with the participation of MIT personnel in this pilot project. Additional resource requirements from the participating organizations would include the cost of monitoring the pilot project, additional data collection costs, and meetings. Further costs may entail investments by the participating supply organizations to ensure the timely availability of the required repair parts and materials for the end-items being studied in the pilot project.

IX. EXPECTED DELIVERABLES AND BENEFITS

Expected deliverables include:

- Briefings, as required, to Air Force leadership;
- Results of site-visits and meetings related to the pilot project;
- Documentation of the pilot project, including experimental design, data sources, principal findings and recommendations for more effective metrics;
- New analytical process for identifying and validating new metrics that the Air Force can use more widely to drive transformational change;
- High-level communication of principal findings to a wide audience of decision-makers in the Air Force, as well as to other policymakers, as appropriate.

The pilot project is expected to generate the following types of benefits:

- Improvements in the efficiency and effectiveness of the US Air Force's military avionics sustainment system through the use of more effective metrics (better resource allocation tied to a proper set of metrics; fostering better prioritization of repair and supply actions);
- Showing the impact of depot maintenance on the readiness of the combat units;
- Improved analytical framework for evaluating tradeoffs and better decision-making (providing an analytical link between *incremental* improvements at the local level and the consequent *incremental* improvements at the system-wide level);

Principal beneficiaries will include warfighters, as well as depot repair and supporting supply organizations. The design and execution of the pilot project is also expected to benefit the LSI industry partners by demonstrating a scientifically rigorous, relatively inexpensive and effective way for improving their performance metrics.

X. MAJOR TASKS AND PROJECT SCHEDULE

Major tasks to be undertaken in implementing the pilot project and the related project timetable is presented below.

Major Tasks/Activities	**Location**	**Timetable (Notional for tasks not yet completed)**
INITIAL ASSESSMENT PHASE (Includes previously completed tasks)		
TASK 1: Conduct Initial (Precursor) Metrics Research ***(Completed)***	Cambridge, MA	Dec 01-May 02
• "House of Metrics" Research	Cambridge, MA	Dec 01-May 01
• Metrics data collection and analysis for avionics end-items	Cambridge, MA	Jan–Jul 02
TASK 2: Develop Preliminary Design of the Pilot Project -- "New Metrics", including "New Operating Rules" ***(Completed)***	OO-ALC	Jun-Sept 02
• Preliminary design of pilot project – kick-off discussions	OO-ALC Cambridge, MA	July 8-10, 02
• Preliminary experimental design	Cambridge, MA	Jul-Aug 02
• Develop operational plan	OO-ALC; WR-ALC; AFMC/LG	Aug 02
• Review operational plan and provide feedback	OO-ALC; WR-ALC; AFMC/LG; Cambridge, MA	Sep - Oct 02
TASK 3: Finalize Pilot Project Operational Plan and Technical Experimental Design	OO-ALC; WR-ALC; AFMC/LG; Cambridge, MA	Nov –Dec 02
TASK 4: Execute Pilot Project	OO-ALC; WR-ALC; AFMC/LG	Jan – Dec 03

TASK 4: Analyze Interim Results and Prepare Report	Cambridge, MA	Jan -March 04

XI. NEXT STEPS

The following action steps are recommended to ensure the successful initiation of the pilot project:

- Review and approval of the proposed implementation plan by the participating stakeholder organizations, as well as any feedback from them regarding any improvements in the proposed implementation plan.

- Development of a general understanding among the participating organizational entities for closer working relationships among them and for concerted action by them in support of the pilot project and the required data collection and monitoring efforts.

- Implementation of the required data collection and monitoring methods and protocols to make sure that the data requirements for the proposed outcome metrics, enabling metrics, confounding factors and control variables, as well as other data needs, can be met on a timely and on-going basis.

- Provision of the required data on all metrics as detailed above for the period preceding the initiation of the pilot project for the development of all pretest data necessary for analytical purposes to estimate the effects of the proposed new metrics.

RESEARCH REFERENCES (LEAN SUSTAINMENT INITIATIVE AT MIT)

Kirkor Bozdogan, "Summary of Findings, Current Projects and Planned Activities," Working Paper, Enterprise Integration Team of the Lean Sustainment Initiative, Massachusetts Institute of Technology, 30 August 2002.

Kirkor Bozdogan and Joseph M. Sussman (in collaboration with Thomas Allen and Wesley L. Harris), "Initial Top-Level Characterization of the Air Force Sustainment System," White Paper, Lean Sustainment Initiative, Massachusetts Institute of Technology, 3 October 1999, 21 pp.

Kathleen N. Chase, "Aircraft Sustainment Operations: Alternatives for USAF Inventory Management," MS Thesis in Management, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, May 2002 (Performed under the auspices of the Lean Sustainment Initiative; Thesis Supervisor: Dr. Donald B. Rosenfield).

Stephen P. Czerwonka (Lt. Comdr., USCG), "Avionics Life-Cycle Forecasting Model," MS Thesis in Aeronautics and Astronautics, May 2000 (Performed under the auspices of the Lean Sustainment Initiative; Thesis Supervisors: Prof. John J. Deyst and Dr. Donald B. Rosenfield).

Lean Sustainment Research, “Major Findings and Future Directions,” Draft Summary Paper, Massachusetts Institute of Technology, 14 July 1998, 16 pp.

Lean Sustainment Initiative (LSI), “Goals, Objectives and Metrics,” Executive Summary, Massachusetts Institute of Technology, 5 August 1998, 8 pp.

Lean Sustainment Initiative (LSI), “Depot Repair Efficiency and Responsiveness,” Executive Summary, Massachusetts Institute of Technology, 14 August 1998, 19 pp.

Lean Sustainment Initiative (LSI), “Execution and Prioritization of Repair Support System (EXPRESS),” Executive Summary, 20 August 1998, 16 pp.

Lean Sustainment Initiative (LSI), “Goals, Objectives and Metrics Task,” Progress Report, Massachusetts Institute of Technology, 31 August 2000, 5 pp.

Dennis F. X. Mathaisel, “On the Relationship between EXPRESS and Air Force Objectives and Metrics: Acase Study,” Draft Working Paper, Lean Sustainment Initiative, Massachusetts Institute of Technology, April 1999, 15 pp.

Stuart J. McGillivray, “Air Force Sustainment Performance Metrics: An Exploratory Evaluation”, Master’s Thesis in Engineering in Logistics, Massachusetts Institute of Technology, May 2002 (Performed under the auspices of the Lean Sustainment Initiative; Thesis Supervisors: Dr. Kirkor Bozdogan and Prof. Joseph M. Sussman).

Keith A. Russell (Lt. Comdr., USCG), “Reengineering Metrics Systems for Aircraft Sustainment Teams: A Metrics Thermostat for Use in Strategic Priority Management,” MS Thesis in Aeronautics and Astronautics *and* Technology and Policy, Massachusetts institute of Technology, December 2000 (Performed under the auspices of the Lean Sustainment Initiative; Thesis Supervisors: Prof. John R. Hauser and Prof. Wesley L. Harris).

Luis C. Tsuji, “Tradeoffs in Air Force Maintenance: Squadron Size, Inventory Policy and Cannibalization,” MS Thesis in Technology and Policy, Massachusetts Institute of Technology, September 1999 (Performed under the auspices of the Lean Sustainment Initiative; Thesis Supervisor: Prof. Joseph M. Sussman).

Jason M. Van Wey, “Alternative Sustainment Options for Military Aircraft: Technology and Regulatory Policy Issues,” MS Thesis in Technology and Policy, Massachusetts Institute of Technology, September 1999 (Performed under the auspices of the Lean Sustainment Initiative; Thesis Supervisor: Dr. Kirkor Bozdogan).

James Wolters, “Top-Level Characterization of the Air Force Sustainment System Using a System Dynamics Model,” White Paper, Lean Sustainment Initiative, Massachusetts Institute of Technology, 31 May 2000 (Research Advisors: Prof. Joseph M. Sussman and Prof. Thomas J. Allen)

D. Yoo, "Database on USAF Sustainment Metrics," Draft Working Paper, Lean Sustainment Initiative, Massachusetts Institute of Technology, February 2002, 21 pp.

RELATED MIT RESEARCH ON METRICS

David H. Campos, "Impact of Performance Measurement and Goal Setting on Supply Chain Responsiveness: An Experiment," MS Thesis in Management *and* Civil and Environmental Engineering in conjunction with the Leaders in Manufacturing Program, Massachusetts Institute of Technology, June 2001 (Thesis Supervisors: Prof. Duncan I. Simester, Mr. James B. Rice, Jr.; Thesis Reader: Prof. Joseph M. Sussman).

Carl Frank, "Metrics Thermostat for Strategic Priorities in Military System Acquisition Projects," MS Thesis, Management of Technology, Massachusetts Institute of Technology, June 2000 (Thesis Supervisor: Prof. John R. Hauser).

John R. Hauser, "Research, Development and Engineering Metrics," *Management Science,* Vol. 44, No. 12, Part 1 of 2 (December 1998), 1670-1689.

John R. Hauser, "Metrics Thermostat," *The Journal of Product Innovation Management*, 18 (2001), 134-153.

John R. Hauser and Gerald Katz, "Metrics: You are What You Measure!," *European Management Journal,* Vol. 16, No. 5 (1998), 516-528.

Burt D. LaFountain, "An Empirical Exploration of Metrics for Product Development Teams," MS Thesis in Mechanical Engineering, Massachusetts Institute of Technology, January 1999 (Thesis Supervisor: Prof. John R. Hauser).

Arpita Majumder, "Strategic Metrics for Product Development at Ford Motor Company," MS Thesis, Engineering and Management, Massachusetts Institute of Technology, June 2000 (Thesis Supervisor: Prof. John R. Hauser).

Jeffry D. Moffitt, "Applying the Metrics Thermostat to Naval Acquisitions for Improving the Total Ownership Cost-Effectiveness of new Systems," MS Thesis, Operations Research, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, June 2001 (Thesis Supervisor: Prof. John R. Hauser).

Duncan I. Simester, John R. Hauser, Birger Wernerfelt, and Roland T. Rust, "Implementing Quality Improvement Programs Designed to Enhance Customer Satisfaction: Quasi-Experiments in the United States and Spain," *Journal of Marketing Research*, Vol. XXXVII (February 2000), 102-112.

SELECTED PRESENTATIONS

T. J. Allen, J.M. Sussman and J. Wolters, "System Dynamics Modeling of the USAF Sustainment System," 13 June 2000.

K. Bozdogan and J. Sussman, “Enterprise Integration Focus Team Research Status,” LSI Co-Leads Meeting, Massachusetts Institute of Technology, Cambridge, MA, 12 December 2000.

K. Bozdogan, J.Lengyel, J. Sussman and L. Williams, “Goals, Objectives and Metrics: Research Overview,” LSI Co-Leads Meeting, Massachusetts Institute of Technology, Cambridge, MA, 1 August 2001.

K. Bozdogan, J. Lengyel, J. Sussman and T. Peterson, “Goals, Objectives and Metrics: Overview,” LSI Steering Group Meeting, Warner Robins, GA, 19 December 2001.

K. Bozdogan, B. Brandt, J. Lenyel, J. Sussman, T, Peterson, “Goals, Objective and Metrics,” Metrics Pilot Project Meeting, Ogden Air Logistics Center (OO-ALC), Ogden, UT, 9 July 2002.

J. Hauser, “Applications of Metrics Thermostat,” LSI Co-Leads Meeting, Massachusetts Institute of Technology, Cambridge, MA, 12 December 2000.

J. Lengyel, J. Shultz, K. Bozdogan and J. Sussman, “Enterprise Integration Focus Team,” LSI Steering Group Meeting, Dayton, OH, 20 March 2001.

K. Russell (Lt. Comdr., USCG), “Goals, Objectives and Metrics – Characterization of the Sustainment of the F-16,” LSI Co-Leads Meeting, Massachusetts Institute of Technology, Cambridge, MA, 12 December 2000.

ATTACHMENT A
GLOSSARY

ACC	Air Combat Command
AETC	Air Education and Training Command
AFB	Air Force Base
AFMC	Air Force Materiel Command
AFR	Air Force Reserve
ALC	Air Logistics Center
ANG	Air National Command
AWP	Awaiting Parts
AWP-F	Awaiting Parts, Parts Supportable
AWP-G	Awaiting Parts, Parts Required
BIE	Base Issue Effectiveness
BOM	Bill of Materials
BSE	Base Stockage Effectiveness
CANN	Cannibalization
CND	Cannot Duplicate
CONUS	Continental United States
CRI	Consolidated Repairable Inventory
CSI	Consolidated Serviceable Inventory
DFLCC	Digital Flight Control Computer
DLA	Defense Logistics Agency
DMAG	Depot Maintenance Activity Group
DMT	Dual Mode Transmitter
DoD	Department of Defense
DRILS	Depot Repair Information Local Server
EIT	Enterprise Integration Team (of Lean Sustainment Initiative)
EOQ	Economic Order Quantity
ERO	Equipment Repair Order
EXPRESS	Execution and Prioritization of Repair Support System
FMC	Fully Mission Capable
FMS	Foreign Military Sales
GAO	General Accounting Office
HQ AFMC	Headquarters, Air Force Materiel Command
HUD EU	Heads-Up Display Electronic Unit
IE	Issue Effectiveness
JRIU	Jettison Remote Interface Unit
LG	Logistics
LGF	F-16 Logistics Requirements Division
LGFBR	F-16 Logistics Analysis Section (within LGF)
LGS	Logistics Supply Division
LNA	Low Noise Amplifier

LPRF	Low-Power Radio Frequency
LRT	Logistics Response Time
LRU	Line Replaceable Unit
LSI	Lean Sustainment Initiative
LYPF	F-15 Logistics Analysis Branch
LYPO	F-15 Avionics Production Branch
MAJCOMS	Major Commands
MALA	F-16 Avionics Production Branch
MC	Mission Capable
MFD	Multi-Function Display
MICAP	Mission Capability
MIT	Massachusetts Institute of Technology
MLPRF	Modular Low-Power Radio Frequency
MRIU	Missile Release Interface Unit
MRO	Maintenance, Repair and Overhaul
MSC	Mode Select Coupler
MT	Maintenance Technicians
MTC	Maintenance Technicians Cross-Trained
NCRSR	Normal Customer Requirement Satisfaction Rate
NM	Materiel Manager
NMC	Not Mission Capable
NRTS	Not Repairable This Station
NSN	National Stock Number
OCONUS	Outside the Continental United States
OO-ALC	Ogden Air Logistics Center
O&ST	Order and Ship Time
OWO	On Work Order (When an end-item is moved from AWP-G to OWO status)
PACAF	Pacific Air Forces
POS	Primary Operating Stock
QDR	Quality Deficiency Report
RGA	Rate Gyro Assembly
RO	Requisition Objective
RSP	Readiness Spares Packages
RTOK	Retest Okay
RTS	Repairable This Station
SBSS	Standard Base Supply System
SE	Stockage Effectiveness
SMAG	Supply Management Activity Group
SPO	System Program Office
SRU	Shop Replaceable Unit
SSC	Shop Service Center
TNMCM	Total Not Mission Capable Maintenance
TNMCS	Total Not Mission Capable Supply
USAF	United States Air Force
USAFE	United States Air Forces Europe
USGS	United States Coast Guard
UCRSR	Urgent Customer Requirements Satisfaction Rate

WCRSR	Weighted Customer Requirements Satisfaction Rate
WR-ALC	Warner Robins Air Logistics Center