Effects of Lab Demo on Employee Engagement

A Dissertation submitted
to the Graduate School
Valdosta State University

in partial fulfillment of requirements
for the degree of

DOCTOR OF PUBLIC ADMINISTRATION

in Public Administration

in the Department of Political Science
of the College of Arts and Sciences

May, 2017

John J. MacGaffick

M.P.A., Troy University, 2010
B.S., Park University, 2007

This dissertation, "The Effects of Lab Demo on Employee Engagement," by John J. MacGaffick, is approved by:

Dissertation Committee Chair

James W. Peterson, Ph.D.
Professor of Political Science

Committee Members

Robert P. Yehl, Ph.D.
Assistant Professor of Political Science

Leigh R. Swicord, D.P.A.
Adjunct Professor of Public Administration

Dean of the Graduate School

James T. LaPlant, Ph.D.
Professor of Political Science

FAIR USE

DUPLICATION

I authorize the Head of Interlibrary Loan or the Head of Archives at the Odum Library at Valdosta State University to arrange for duplication of this dissertation for educational or scholarly purposes when so requested by a library user. The duplication shall be at the user's expense.

Signature 

I refuse permission for this dissertation to be duplicated in whole or in part.

Signature _____

ABSTRACT

Lab Demo is coming! Lab Demo is coming! While these words reverberated throughout the halls of the Air Force Research Laboratory's (AFRL) Information Directorate in 2010 with less urgency and consternation than Paul Revere's warning in 1775, there was apprehension, nevertheless.

The federal Civil Service has been experiencing performance management reform since the passage of the Pendleton Act in 1883. During the intervening 130+ years, a number of innovations and reforms have been introduced, some of them effective (for example, the Classification Act of 1923), and some of them lacking in productive results (like the now-discarded National Security Personnel System). Many of these innovations were first introduced in the private sector and adopted by the government regardless of success or failure. The latest experiment is the Air Force Laboratory Personnel Demonstration Project, or Lab Demo for short. It introduces several changes never before seen in the Civil Service, such as broadbands, pay for performance, and local control over key human resources processes, and, according to its proponents, has a history of improving work force quality, including employee engagement. This paper examines the effect Lab Demo has on various workforce statistics, and how the results reflect on employee engagement.

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

This paper would not have been possible without the assistance of Dr. Tim Kroecker and Dr. Victoria Goliber, who spent much of their spare time editing this work, and the graciousness of the Colonel Tim Pennington and the now-retired Colonel Dave Blanks, my supervisors who granted every time-off request I submitted in order to pursue my doctorate.

I would also be remiss if I did not mention my fellow cohort members and other classmates. I learned so much and so many other points of view, which I find invaluable not only in my academic pursuits, but my professional endeavors as well.

DEDICATION

This dissertation is dedicated to my wife, Nurdan MacGaffick, and our son, Jeremiah MacGaffick. In her own gentle way, Nurdan encouraged to start both my master's and doctorate programs, imparting her confidence in me and stepping up to handle what I normally would have had I not been attending school. Jeremiah has been a source of motivation for me to attain what I have thus far, and forcing me to set high goals, as it is his sincere desire to surpass them. I wish him well and I hope he does so.

Chapter I

INTRODUCTION

Lab Demo is coming! Lab Demo is coming! While these words reverberated throughout the halls of the Air Force Research Laboratory's (AFRL) Information Directorate in 2010 with less urgency and consternation than Paul Revere's warning in 1775, there was apprehension, nevertheless.

In all actuality, Lab Demo, an experimental personnel management system featuring broadbanding, pay plans, and merit pay, had been in place for AFRL's scientists and engineers since 1997. Since then, additional employees have been added to the project in an effort to demonstrate successful implementation and gain approval. The more recent wave of concern was being felt and expressed by the administrative and support personnel, the latest group of employees to be brought in under Lab Demo's purview. Many of the researchers had expressed displeasure with the system, and, with few other points of reference, office employees looked upon the Lab Demo process with misgivings.

This study will examine whether or not the proposed personnel management system, Lab Demo, has been sufficiently designed, implemented, and modified enough to engender employee engagement among AFRL's workforce. This will be done by examining trends in employees' evaluations, promotions, separations from federal

service, and satisfaction survey results. The separations and survey results will also be compared to the rest of the federal government as a whole, to test whether or not Lab Demo achieves the same results. During the course of this study, we will attempt to answer the following questions:

- Are there trends in how many employees fall into each of the four zones?
- Are there trends in the annual changes in percentages of evaluations in each zone?
- What are the trends in annual promotions?
- What are the trends concerning employees leaving AFRL?
- Are the separation statistics similar to the federal government as a whole?
- Do employees report satisfaction with Lab Demo?
- Are Lab Demo employees as satisfied as the entire federal workforce?

We will use various statistical analyses (correlation, regression, and *t* Tests) to determine whether or not our hypotheses hold up and what the answers are to our seven questions.

What is the Air Force Research Laboratory (AFRL)?

The Air Force Research Laboratory, or AFRL, is the "Air Force's only organization dedicated to leading the discovery, development, and integration of warfighting technologies for our air, space, and cyberspace forces" (Wright-Patterson Air Force Base, 2015). Headquartered at Wright-Patterson Air Force Base, OH, it was formed in 1997 when four previously existing Air Force laboratories and the Office of Scientific Research were merged into one organization. It is now comprised of seven stand-alone technology directorates, the 711th Human Performance Wing, the Air Force Office of Scientific Research, and its own central staff (U. S. Air Force, 2014).

What is Lab Demo?

Lab Demo, formally known as the Air Force Research Laboratory Personnel Demonstration Project, is the Air Force's contribution to Acq Demo, or the Department Of Defense (DOD) Civilian Acquisition Workforce Personnel Demonstration Project. Acq Demo was signed into law by then President Clinton as a part of the Fiscal Year 1996 National Defense Authorization Act. In it, the Secretary of Defense was "encouraged to take such steps as may be necessary to provide for the commencement of a demonstration project, the purpose of which would be to determine the feasibility or desirability of one or more proposals for improving the personnel management policies or procedures that apply with respect to the acquisition workforce of the DOD" (U.S. Congress, 1996, p. 485).

Features of Lab Demo

Beginning with the Pendleton Act of 1883, federal human resources management has seen a succession of legislative acts, each of which introduced innovation, at least as far as the federal government was concerned, to the civil service. Lab Demo has not attained the status of federal decree, yet; however, it does bring to the personnel under its auspices some human resources features heretofore unseen in the Air Force.

Broadbanding. Broadbanding is defined by Office of Personnel Management (OPM) as "the use of a salary structure into fewer broader pay ranges than traditional grades" (OPM, n.d.). One of the advertised benefits of Lab Demo is the flexibility afforded local managers in hiring personnel. To obtain that flexibility, the designers of Lab Demo reduced the 15 General Schedule grades to four. Appendix A lists the broadbands used in the Lab Demo system and their General Schedule equivalents.

Broadbanding has been widely implemented in both the public and private sectors, for reasons such as rewarding performance, improving competitiveness, simplifying classification, delegating authority, reflecting the mission and the nature of the work being performed, and, as stated above, to creating flexibility. There are challenges associated with this move, though, including cost management, communication (fewer distinctions in job evaluations means less communication about roles, hierarchy, and values), performance management, and the linkage of pay to the labor market (Office of Personnel Management, n.d.).

Pay Plans. There are four pay plans under Lab Demo, each of which represents a group of specialties within the federal Human Resources (HR) system. These pay plans replaced the majority of the GS positions throughout AFRL Instead of plugging new employees into narrow job descriptions, based on the exact position number being filled, managers select from a greater array of qualified candidates and simply assign them to the appropriate pay plan. They are: pay plan DR (Engineering and Science); DO (Business Management and Professional); DX (Technicians); and DU (Mission Support). Appendix B breaks down exactly which GS specialties fall under each Lab Demo pay plan.

Contribution-Based Compensation System (CCS). Also known as pay-for-performance or merit pay, CCS is perhaps the most radical departure from the status quo, and quite possibly the most controversial. Figure 1 displays the 2009 compensation for Pay Plan DU. The solid line between the two dotted lines is the standard pay line, or SPL. The SPL represents the target score for all employees. As the amount of basic pay increases, so does the SPL. The upper dotted line is known as the upper rail, and the

lower dotted line is known as the lower rail. The area above the upper rail is known as Zone A, which represents substandard employees. Employee X, whose base pay was about $39,000, finished with a contribution score of 2.00, which is below the target of 3.00. This employee finished above the upper rail, in Zone A, which represents overcompensated or substandard employees. Employee Y, whose basic pay was $21,000, finished with a score of 2.30, well above the standard score of 1.30. This employee is in Zone D, which represents undercompensated employees. Finally, Employee Z, with a basic pay of $27,000, finished with a score of 1.80, right on the SPL. The area between the upper rail and the SPL is Zone B, and that between the SPL and the lower rail is Zone C.

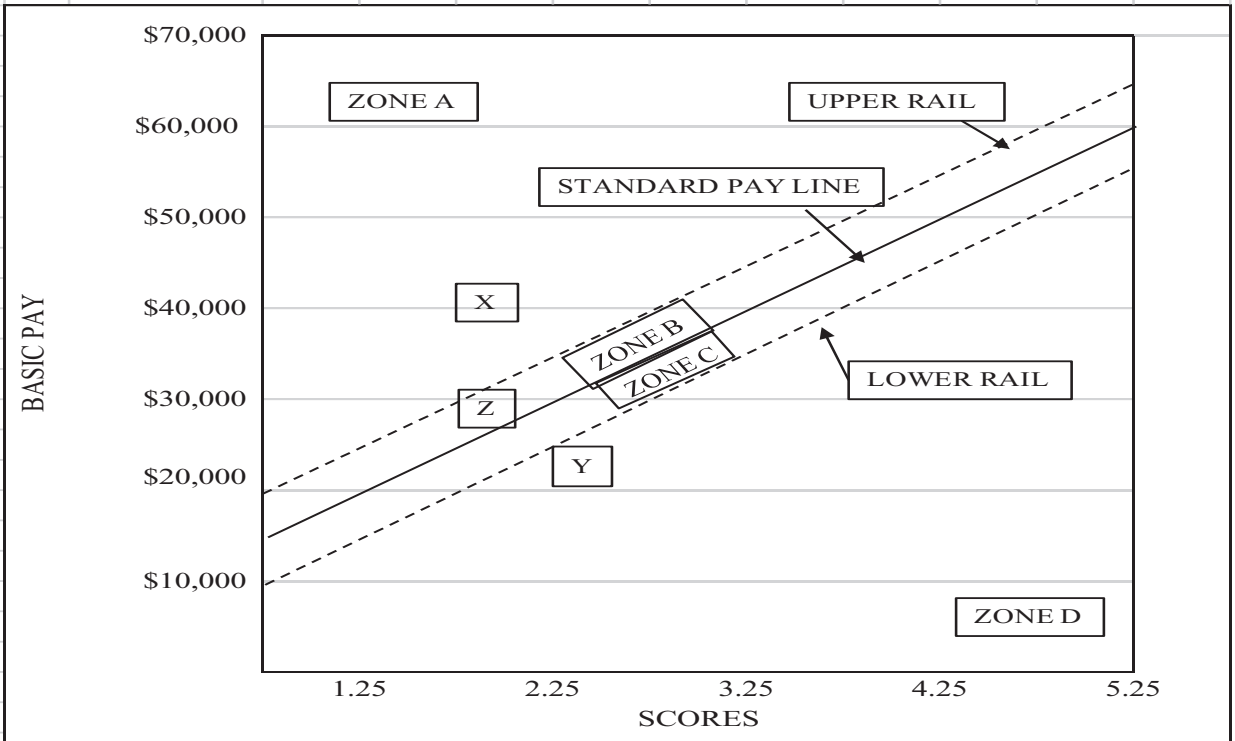

Figure 1. The four zones of Lab Demo personnel evaluations, with examples of employee positioning. Adapted from Air Force Research Laboratory Manual 36-104, Personnel Demonstration Project, September 1, 2011.

Figure 1. Examples of Employee Positioning

The General Schedule (GS) System

The current General Schedule system has been in place since 1949. As stated earlier, it has 15 grades (GS-1 to GS-15). Each of these grades has ten steps, each of which is worth roughly three percent of an employee's salary (OPM, 2014), in addition to other requirements. The progression through these steps is illustrated in Table 1.

Table 1 General Schedule Step Increase Time Requirements

Advancement from...	Requires...
step 1 to step 2	52 weeks of creditable service in step 1
step 2 to step 3	52 weeks of creditable service in step 2
step 3 to step 4	52 weeks of creditable service in step 3
step 4 to step 5	104 weeks of creditable service in step 4
step 5 to step 6	104 weeks of creditable service in step 5
step 6 to step 7	104 weeks of creditable service in step 6
step 7 to step 8	156 weeks of creditable service in step 7
step 8 to step 9	156 weeks of creditable service in step 8
step 9 to step 10	156 weeks of creditable service in step 9

Steps 1 through 3 require 1-year waiting periods; steps 4-6 2 years; and 3 years for steps 7-9; however, after 1 year in one particular grade, an employee may compete for promotion to the next higher grade (for example, GS-7 to GS-8).

During these waiting periods, employees also have three other requirements they must meet. First, his or her job performance "must be at an acceptable level of competence. To meet this requirement, an employee's most recent performance rating of record must be at least Level 3 ("Fully Successful" or equivalent) (OPM, n.d.). Also, he or she must not have received a pay increase equivalent to that which would have resulted from a step increase during the waiting period (Office of Personnel Management, n.d.). Finally, the employee must be occupying a permanent position, meaning a position that by someone whose appointment is not designated as temporary and is does not have

a time limitation of 1 year or less. The term "permanent position" also includes those positions filled by employees promoted on temporary or term bases for a minimum of 1 year (OPM, n.d.).

While this system does provide for regularly scheduled pay increases, there is no real accountability for poor performers. They are paid the same amount and receive the same pay increases as superior performers. As a result of this perceived unfairness, government organizations run the risk of losing their top performers to other levels of government or the private sector, a situation which Lab Demo attempts to resolve. To add to the problem, line managers are limited in the administration of personnel resources, and the regulations governing human resources functions often conflict with what is needed to support the research function.

Another General Schedule defect the developers of Lab Demo seek to fix is the classification system. Its cornerstone is protracted, cumbersome position descriptions, which are then classified through the use of lengthy (the classification standard for this author, a GS-0201 Human Resources Specialist, is 102 pages long), outmoded standards. The Lab Demo system, in which new personnel are hired against locally authored position descriptions, results in a faster hiring process that allows for greater control over the quality of the workforce.

At the other end of the process, in a reduction-in-force, or RIF, both systems actually experience the same difficulties of complex, hard-to-understand rules that do not assign an important enough role to employee contribution (National Archives and Records Administration, 2010, pp. 53076, 53081-53082).

National Security Personnel System (NSPS)

A solution for these human resource issues has been attempted recently, with the now-defunct National Security Personnel System, or NSPS. NSPS was announced with some fanfare and promised many of the same features present in Lab Demo:

> The U.S. Office of Personnel Management will continue as a full partner with the Defense Department during the phase-in of a new human resources system that rewards employee performance, treats workers fairly, yet gives agency managers latitude to manage for results…OPM and DOD have been equal partners in developing the NSPS, and we will remain so in any future adjustments to the basic design…We've built in regulatory flexibility so the department can tailor NSPS to specific needs in staffing and appointing authorities, in classification of occupations, and in managing pay and performance systems…The National Security Personnel System was authorized by Congress and signed into law by President George W. Bush in 2004. It replaces a more than half-century old pay and classification system designed primarily for a clerical work force fighting the Cold War. The NSPS is a contemporary system -- not unlike a new HR system at the Department of Homeland Security -- that links employee pay to performance, adheres to principles of merit in personnel decisions, and gives managers greater flexibility to adjust their operations based on staffing or mission needs… (OPM, 2005).

For all of this promise, however, NSPS proved to be a short-lived system. The first payout exhibited the following issues:

- White employees received higher average performance ratings, salary increases and bonuses than employees of other ethnicities.
- Employees at DOD-level agencies averaged a full percentage point more in payouts more than their counterparts employed at the three military service branches.
- Employees over 40 and employees under 40 received similar ratings, yet those over 40 received more in bonuses and payouts. This, however, might be attributed to those over 40 holding position with more control and responsibility, hence higher salaries.
- Asian American employees received slightly higher overall ratings than their African American counterparts, yet their pay raises and bonuses were significantly smaller.
- Women received larger total payouts than men, even though their ratings were similar on average (Losey, 2008).

These concerns led to NSPS being repealed by the 2010 National Defense Authorization Act on Oct 28, 2009, and employees covered by the system were transitioned back to the "statutory pay system and all other aspects of the personnel system that last applied or would have applied if NSPS had not been established" and that no employee suffer a loss of pay due to the conversion (U.S. Government Accountability Office, 2011). In this same report, the seeds of the next personnel management system were sown: the new performance management system called for in this report would require performance appraisals that bonuses and other actions would link to, and plans for employee training, counseling, and mentoring (U.S. Government Accountability Office, 2011).

Significance of the Study

The importance of the project is twofold. First, it entails a radical departure from the federal civil service system that has been entrenched for decades. It is essentially a

pay-for-performance system, a concept that has engendered a fair amount of controversy. Federal civilian employees, for better or for worse, have long enjoyed great job security and pay raises at regular intervals. These two facts can, depending on the mindset of individual employees, result in complacency resulting from not having to worry about being dismissed or when the next bigger paycheck was due to roll in, as well as dissatisfaction resulting from the knowledge that even with pay raises, the grass is still greener on the private sector side of the fence.

Secondly, the Project is an attempt to address human resources issues throughout the Laboratory, which, while not expressed in so many words, can be traced back to personal motivation and engagement. Unofficially, the Laboratory's mission is to develop technologies designed to help our warfighters perform their assigned missions more efficiently, more effectively, and more safely, and to return home. If ever a DOD function called for its employees to be focused, motivated, and engaged, the Laboratory is it. The overarching goal is the preservation of our warfighters' lives. What could be more important than to protect those who are sworn to support and defend the Constitution of the United States against all enemies, foreign and domestic, and to bear true faith and allegiance to the same (U.S. Government Printing Office, 2012)?

Therefore, employee engagement is a problem, which, if left unsolved, leads to other problems. The literature review will discuss employee engagement and how it affects or is affected by other human resources factors.

This study is being conducted for two reasons. First, in this era of tightening military budgets, Lab Demo represents a higher personnel cost (with bonuses or promotions that are accompanied by pay raises) than does the General Schedule system.

Prior to September 1, 2011, even Zone A employees could expect a pay increase, ranging from as little as zero to as much as the annual civil service pay increase given by Congress (Air Force Research Laboratory, 2008, p. 47). After September 1, 2011, however, those same employees receive no pay increase (Air Force Research Laboratory, 2011, p. 54). Even with this new policy, the only way to incur no additional costs would be to have all employees rated as Zone A. This would be tantamount to admitting human resources flaws of epic proportions and result in administrative nightmares. Second, AFRL is no different from the rest of the government in that it competes with the private sector for talent. Since the public sector lags behind industry in terms of compensation, it must get creative in creating an atmosphere in which people are willing to work in positions that are stigmatized by a reputation of less pay and recognition.

Other Chapters in this Study

Chapter 2 is a literature review covering key concepts in this study, divided into two parts. The first part is related to public administration, and covers literature relating to the administrative responsibility of government, and public vs. private sector administration. After crossing a bridge topic of public service motivation, the review turns to human resources management, and covers employee engagement and merit pay.

Chapter 3 covers the methodology used in this study. Detailed are the three claims set forth in the Federal Register to be examined, as well as the employees that are the subject of this study, the sources of the data to be analyzed, the limitations of this study, the case study procedures, statistical analyses, and the hypotheses for testing.

Chapter 4 shows the data used in this study, the results of the tests conducted on that data, and what those results indicate.

Chapter 5 examines the significance of Chapter 4's results and how they relate to whether Lab Demo is generating higher levels of employee engagement in AFRL. Also discussed will be the future of Lab Demo and what needs to be analyzed in the future, should it be continued.

Chapter II

LITERATURE REVIEW

The proposal for this project developed from professional experience on the part of the author, whose last military duty assignment featured the opportunity to work alongside civilians under NSPS. It is a program evaluation of the Air Force Research Laboratory Demonstration Project, also known as Lab Demo. Its goal is to evaluate the Project's effectiveness at elevating the overall employee engagement of the Laboratory's workforce.

Government Responsibility

"That to secure these rights, Governments are instituted among Men, deriving their just powers from the consent of the governed, --That whenever any Form of Government becomes destructive of these ends, it is the Right of the People to alter or to abolish it, and to institute new Government, laying its foundation on such principles and organizing its powers in such form, as to them shall seem most likely to affect their Safety and Happiness" (National Archives and Records Administration, 2013). While this issue is highly unlikely to result in a new federal government, this short passage from the Declaration of Independence illustrates that the government is accountable to the people, especially in terms of national security, in line with the social contract theory at the heart of our government.

Nevertheless, it wasn't until 1959 that a definition of a government's administrative responsibilities was catalogued. Gilbert stated that the word "responsibility" is a general boundary around a set of values, and that he was attempting to list these values. These values are: 1) responsiveness; 2) flexibility; 3) consistency; 4) stability; 5) leadership; 6) probity (the government should exhibit honesty, integrity, and trustworthiness); 7) candor; 8) competence; 9) efficacy; 10) prudence; 11) due process; and 12) accountability (Gilbert, 1959, pp. 375-378).

So, how is the government performing on this set of values? Rather than reforming the current personnel system, attempts are being made by various agencies to develop alternative systems (such as Acq Demo and Lab Demo), some of which have received approval from the OPM, some of which have not (Lab Demo is in this group). Although the Lab Demo project is well into its second decade without any indication of what will happen, it appears the Air Force is simply not being responsive to indications of the need for policy change. These indications are not the result of such overt actions as strikes or refusal to perform assigned duties; rather, they are the products of the lack of highly qualified candidates to fill vacant positions or the short tenures of scientists and engineers that exit the federal service as soon as possible for the greener pastures of the private sector. One way to compensate would be to offer higher starting salaries, but then this would violate Gilbert's principle of efficacy. Additionally, the fact that problems in the Civil Service were allowed to fester for years shows a clear lack of leadership not only in OPM, but also higher echelons of the Executive Branch (showing a lack of oversight) and Congress (by failing to exercise the power of the purse to allocate resources to fix the problems). Step increases (resulting in small pay raises for individual

employees) are based on the attainment of a satisfactory performance level, and, at the Defense Department, "only about 1 percent of employees receive a rating of unacceptable" (O'Keefe & Yoder, 2011).

Not all responsibility for this untidy situation lies at the feet of those who are involved in policy decisions concerning oversight of the federal workforce, but that majority of it resides there. There is a perhaps less obvious force that appeared in the realm of public administration years ago that can best be introduced by asking the following question.

Is Military Administration Public Administration?

At first blush, it may seem that there is a difference between military administration and public administration. Granted, there are certain rules and requirements that are specific only

to the armed forces, but the same could be said about many other sectors in public administration.

How did this apparent dichotomy come to pass? According to Jefferies, it occurred due to the military's heavy reliance on private business and industry for management techniques, its reputation of uniqueness as a profession, and the lack of interest on the part of public administrators (Jefferies, 1977, p. 321).

Jefferies posits that the focus of public administration study and practice is on the executive branches at all levels of government, and their attempts to formulate and implement domestic public policy. At the federal level, the development and application of foreign policy is neglected, and, since the military is seen as an instrument of that foreign policy, it is ignored (Jefferies, 1977, p. 321). This, then, could be what has led to the military's reliance on private sector management techniques, for better or for worse.

Jefferies, however, reminds us that the armed forces are indeed public agencies, as they are departments within the Executive Branch of the federal government, and, therefore, "fall within the milieu of public administration" (Jefferies, 1977, p. 322). This lack of interaction between the public and military administration seems to have stemmed from the fact that public administration grew up as a child of political science, first concerned with the domestic economy and government agencies, while, at the same time, the military was studied as an instrument of foreign policy, another child of political science; there was little or no interaction between the two siblings (Jefferies, 1977, p. 327).

Further muddying the waters is the fact that the majority of the Air Force Research Laboratory workforce is comprised of civilians; yet, it is still a military organization. It is commanded by a major (two-star) general, who answers to a four-star general, who, in turn, answers to another four-star general, who reports to the Secretary of the Air Force, and so on. Colonels command several of the technical directorates and research sites. Most importantly, its customers are the uniformed men and women who fight our nation's battles, often on the other side of the planet. There is a tendency to forget these facts in an organization predominately staffed by civilians, as well as an inclination to forget that even though they are highly educated researchers, they are also public servants.

If we are to agree with Jefferies' assertion that military administration is public administration, we are still faced with an epic battle of administration between the two sectors.

Public vs. Private Sector Administration

Public administration is saddled with a long history of begging, borrowing, or stealing private sector management techniques, many of which have been abject failures in the private sector before being forced onto the government stage. Ingraham discussed this, citing various assumptions about this method of so-called innovation:

> As one of the anonymous reviewers of this analysis observed, "It is typical for a government to follow private sector examples and buy pokes without looking inside to see if there's really a good pig in there." Other assumptions build on that foundation: working assumptions related to the motivation of public employees (they are just like their private sector counterparts) and the role of financial incentives in motivation and commitment (they are the most important incentive in both cases) clearly assume easy transferability of policies and techniques. Whatever problems may exist with this transferability, the practice of looking to the private sector for public solutions is so well established as to be nearly inviolate (Ingraham 1993, 348-349).

The neglect of the military by those who have developed and guided the study of public administration, and by those who have put what they were taught into practice without giving any indication of what works and what doesn't, is the primary cause of this. For years, the federal government used a system not practiced by the private sector, and with some adjustments, it could still be the one system that would be the ideal solution. Instead, it turned to the private sector with no thought for the fact that private

sector systems are designed with the particular company or firm in mind, which is in business to generate profit, something that is not the intended function of government.

This governmental dependency on private-sector concepts came to a head during the 1980s and 1990s in an era known as New Public Management, or NPM. This phenomenon began in Great Britain during the 1980s and made its way to the rest of Europe, the United States, Australia, and New Zealand. Its widespread adoption clearly indicates a perspective change regarding how governments should be run (Lane, 2003, p. 3).

New Public Management has had its effects on human resources management as well. The Winter Commission of 1993 set forth a number of recommendations to remove barriers to civil service reform. Several of those suggestions have been adopted by the developers of Lab Demo, including emphasizing personnel contributions to organizational objectives, decentralizing personnel policy control, reducing the number of job classes or job titles, simplifying position classification, and implementing broadband compensation (Nigro & Kellough, 2008, p. S51).

While these changes (and the others not listed here) were implemented to differing degrees by the states, the federal government, including the non-acquisition portion of the DOD, largely ignored them. On the other hand, they are all a part of Lab Demo, which is, after all, a two-fold management system.

Lab Demo is a Personnel Management System and a Performance Management System

Lab Demo represents an attempt at a new method of managing federal personnel, one that uses performance management to determine merit pay, promotions, and retention, among other things. In a perfect world, there would be consistency in the way

supervisors perform their evaluator duties; however, there are influences not found in the system that affect how raters rate their subordinates. Grote refers to these factors as "rating errors," and he describes nine possible errors that can negatively influence any rating system. Any one or any combination may have cropped up in Lab Demo. They are: 1) the contrast effect (comparing individuals with each other rather than against job standards); 2) first impression error (tendency to base later information on a positive or negative early impression); 3) halo/horns effect (generalizing from one aspect of someone's performance to all areas of his or her performance); 4) similar-to-me effect (the tendency of a supervisor to rate subordinates similar to home or her more highly than other subordinates; 5) central tendency (the inclination to rate subordinates toward the middle of the scale, even if their performance warrants a higher or lower rating; 6) negative/positive skew (the opposite of central tendency); 7) attribution bias (attributing performance failures to individuals and performance successes to other causes; 8) recency effect (placing more influence on recent minor events than on major events earlier in the reporting period; and 9) stereotyping (generalizing across groups and ignoring individual performance differences) (Grote, 1996, pp. 137-139). These errors will also be discussed further later in Chapter 5.

Why Does the Government use Performance Management?

Our first president tended to appoint government officials based on merit; however, his successors did not completely follow his example, and, when the Jackson administration was in office, the spoils system was the rule rather than the exception. This system resulted in many disgruntled would-be bureaucrats who were not rewarded with the positions they thought they so richly deserved. It was one such man, Charles

Guiteau, who assassinated President Garfield in 1881. Less than 2 years later, the Pendleton Act was passed. It required appointments and promotions by examination; however, there was no appraisal system established. That would not occur until 1912 (OPM, n.d.).

Moynihan and Pandey asked why managers use performance management data. The use of that data, or lack thereof, serves as the best indicator of whether or not performance management is worth the effort. Also, not knowing why it is being used inhibits efforts to successfully utilize performance management (Moynihan & Pandey, 2010, p. 850).

Today, the U.S. government uses performance management for "planning work and setting expectations, continually monitoring performance, developing the capacity to perform, periodically rating performance in a summary fashion, and rewarding good performance" (OPM, n.d.). These almost fit in with the 16 reasons Grote listed for how Fortune 100 companies use performance management data, with one glaring exception. Grote's top reason for the use of this data is "improving work performance" (Grote, 1996, p. 10). Compare it to the government only "continually monitoring performance," and an underlying reason for the malaise suspected of enveloping the federal workforce culture begins to emerge. The popular conception that the federal government is comprised largely of dissatisfied and/or disengaged bureaucrats appears to be reality. If we examine the 2010 Federal Employee Viewpoint Survey, taken in the same year as the AFRL survey we will discuss later, the percentage of positive answers ranges from 15.8% to 97.4%, with an average of 62.5% (OPM, n.d.).

What is Employee Engagement?

Before we proceed further, we need to define the term at the crux of this study. Saks and Gruman tell us that answering this question may be a task more easily said than done. The study of employee engagement appears to have started with the publication of Kahn's 1990 article. Since that time, it has been stated that employee engagement is crucial to organizational success, and that it appears to be on the decline, costing American businesses billions of dollars (Saks & Gruman, 2014, p. 156).

To make this more complicated, Saks and Gruman also state that there are numerous definitions for employee engagement, and that there is even some disagreement about the construct of the name, with some preference for *job engagement*, and some support in the direction of *work engagement* (Saks & Gruman, 2014, p. 156).

This article traces a history of the definition of employee engagement. In it, Saks and Gruman identified three different definitions of employee engagement; however, they did not set forth one of their own.

Macey and Schneider approached the task from several different angles. At times, they offered propositions concerning psychological state engagement; others, behavioral engagement; still others, trait engagement (Macey & Schneider, 2008, p. 3). Rather than going through all 14 of their propositions, we will concern ourselves certain points.

Regarding engagement as a psychological state, discussion has encompassed some form of absorption, attachment, or enthusiasm. Measures of engagement revolved around job satisfaction, organizational commitment, psychological empowerment, and job involvement. Some have thought satisfaction and engagement to be directly linked, but our authors quoted Erickson's 2005 study as being in line with their opinions:

"Engagement is above and beyond simple satisfaction with the employment arrangement or basic loyalty to the employer—characteristics that most companies have measured for many years. Engagement, in contrast, is about passion and commitment—the willingness to invest oneself and expend one's discretionary effort to help the employer succeed" (Macey & Schneider, 2008, p. 7). The first of 14 propositions used to summarize the opinions set forth in this article reads thus: "Proposition 1: Satisfaction when assessed as satiation is not in the same conceptual space as engagement. Satisfaction when assessed as feelings of energy, enthusiasm, and similarly positive affective states becomes a facet of engagement" (Macey & Schneider, 2008, p. 8).

Regarding engagement as commitment, previous studies have offered definitions such as "to be actively committed, as to a cause" and "the extent to which employees commit to someone or something in their organization, how hard they work, and how long they stay as a result of that commitment" (Macey & Schneider, 2008, p. 8). Macey and Schneider offer this viewpoint: "Proposition 2: Organizational commitment is an important facet of the state of engagement when it is conceptualized as positive attachment to the larger organizational entity and measured as a willingness to exert energy in support of the organization, to feel pride as an organizational member, and to have personal identification with the organization" (Macey & Schneider, 2008, pp. 8-9).

Job involvement has also occupied a prominent place in the study of employee engagement. In 2005, Cooper-Hakim and Viswesvaran linked job involvement and job commitment by defining job commitment "as the degree to which an employee psychologically relates to his or her job and the work performed therein" (Macey & Schneider, 2008, p. 9). With their third proposition, our authors agree: "Proposition 3:

Job involvement (including task engagement and job commitment) as traditionally conceptualized and assessed is an important facet of the psychological state of engagement" (Macey & Schneider, 2008, p. 10).

Our final point of discussion from this article is of a behavioral viewpoint. As stated earlier, some studies of employee engagement have approached the subject from a behavioral viewpoint, discussing the subject in terms of "effort," such as "extra effort," "discretionary effort" (the use of more time, brainpower, and energy), or simply "giving it their all" (Macey & Schneider, 2008, p. 14). To encompass all these definitions and more, this proposition is offered: "Proposition 7: Engagement behaviors include innovative behaviors, demonstrations of initiative, proactively seeking opportunities to contribute, and going beyond what is, within specific frames of reference, typically expected or required" (Macey & Schneider, 2008, p. 15).

However, this multi-faceted approach is not without its critics. The University of Toronto's Alan Saks opined that by presenting engagement in behavioral, trait, and state packages, "one is left wondering which of the three should be measured and be the focus of future research" (Saks, 2008, p. 40).

His first argument is from Kahn's 1990 research, which stated, "engagement reflects the extent to which a person is psychologically present in the performance of their work roles" (Saks, 2008, p. 41), or in other words, how they do what they are required to do when they are performing a particular task, and not how they do things not required of them. His example is the August 2006 crash of Comair Flight 5191 at the Lexington, KY, Bluegrass Airport. The aircrew took off from the wrong runway, and the subsequent crash killed 49 of the 50 people on board. Per the National Transportation

Safety Board report, the crewmembers failed to use the available cues and aids to identify their position, failed to crosscheck and verify that they were on the correct runway, and engaged in conversations not pertinent to operations during the taxi. All of these indicate a lack of engagement and contributed to the accident. "Thus, engagement is not about being innovative or doing more and doing things differently; it is about how you do what you are supposed to be doing" (Saks, 2008, p. 41).

The second argument is based on Macey and Schneider's inclusion of adaptive discretionary behavior, which Saks argues is a form of extra-role behavior and might result from engagement. Engaged employees are more apt to initiate change or innovate, and "do things that support organizational effectiveness. Thus, adaptive discretionary behavior is more likely an outcome of engagement, but it is not engagement itself" (Saks, 2008, p. 41).

Saks goes on to cite several works that appear to refute Macey and Schneider's contentions that the meaning and measure of engagement in literature suffer from imprecision and a lack of clarity. He concludes that they "ended up precisely where they began- began—an imprecise definition and a repackaging of other constructs" (Saks, 2008, p. 42).

So now, we have a historical view of the study of the subject, and an overview of what its important components might be, but we still lack a concise definition. For this, we will turn to Marciano. In his book, he begins with what doesn't work in terms of building employee engagement before he provides this definition: "…Although not quite so life-altering, the concept of employee engagement is also about the extent to which one is committed, dedicated, and loyal to one's organization, supervisor, work, and

colleagues. When you're truly committed, motivation becomes a lot less relevant—you're in it for the long haul" (Marciano, 2010, p. 40).

How Does Performance Management Influence Organizational Culture?

Since the General Schedule system only "continually monitors" performance, rather than attempts to improve it, it stands no chance of improving the performance of the federal workforce. This situation extends to all sections of the government, including the Air Force Research Laboratory. Even the most highly PSM-oriented people would find it difficult to remain in such a situation. The GS system has "several major inefficiencies, which hinder management's ability to recruit and retain the best-qualified personnel" (National Archives and Records Administration, 2010, p. 53082). Lab Demo was implemented to reform the culture in AFRL that resulted from the loss of highly capable, highly motivated personnel. Is a performance management system capable of such a massive undertaking? Borrowing from the second and third of Behn's three big questions of public management (eliminating the "how" he began them with), "can public managers motivate public employees (and citizens too) to pursue important public purposes with intelligence and energy," and "can public managers measure the accomplishments of their agencies and of themselves" (Behn, 1995, p. 318)?

Sanger cites several previous works in her study describing how the New York City Department of Finance used performance management to transform itself from "public bureaucracy into a results-based organization" (Sanger, 2008, p. 621), works that posit that leadership and its ability to change an organization's culture are critical to producing lasting changes in how organizations operate (Sanger, 2008, p. 622). Leaders are asked to use culture to change values (creating quality and value rather than creating

efficiency), assumptions and behaviors. Old hierarchal systems must be abandoned in favor of less rigid hierarchies and employees empowered by fewer rules, increased discretion, rewarding of innovation, and tolerance of well-conceived failures (Sanger, 2008, pp. 624-625). Whether or not these have occurred in AFRL depends on who is asked, but, given the apparent improvement in employee performance, movement seems to have appeared in that direction. The previous "laissez-faire" culture has been replaced by one that at least appears to hold employees at all levels accountable. The leadership piece of the puzzle is what proves to me the most challenging aspect. Each directorate in AFRL is led by either a commander or a director, and they are replaced periodically. Since the annual reports encompassed all of AFRL, they failed to take into account the different leadership styles in the different directorates. Thus, it is impossible to determine what effect the leaders had on the engagement of their employees, especially in the area of trust. Yang and Kassekert set forth, among others, these three hypotheses: "trust in leadership is positively associated with job satisfaction;" "when trust in leadership is higher, the relationship between perceived performance orientation and job satisfaction is greater;" and "when trust in leadership is higher, the relationship between innovative culture and job satisfaction is greater" (Yang & Kassekert, 2010, p. 419). All three hypotheses were considered together by examining trust in leadership, performance orientation, and innovative culture. Two plots were developed, one with low performance orientation and one with high. Each also reflected results using high and low trust in leadership numbers. All four plot lines trend upward, so Yang and Kassert's first hypothesis is supported; the second and third are only partially supported, as the

second is dependent on the level of innovative culture and the third relies the level of performance orientation (Yang & Kassekert, 2010, pp. 427, 429).

It should be noted that Yang and Kassekert tied trust in leadership to job satisfaction. Job satisfaction is a component of employee engagement, but the two are not the same. In order to tie these results into employee engagement, we revert to Rich, Lepine, and Crawford, whose results indicated a strong correlation between job satisfaction and employee engagement, with β equaling .29 (Rich, Lepine, & Crawford, 2010, p. 625).

Employee Engagement and Other Factors

Anitha conducted a study on how seven different factors determine an employee's engagement level: 1) work environment; 2) leadership; 3) team and coworker; 4) training and career development; 5) compensation; 6) organizational policies; and 7) workplace well-being. To determine whether or not these factors had a significant influence on employee engagement, she distributed a survey based on a 5-point Likert scale, and asked the respondents to rate each item. Based on a pilot of 60 respondents, the survey instrument was validated, and the factors listed above were found to have reliability factors ranging from .704 to .948. A regression conducted on her data revealed an adjusted r^2 of .672, meaning that these factors account for 67.2% of the variance in an employee's engagement. She then ran a second regression analysis using employee engagement as the independent variable and employee performance as the dependent variable. This analysis resulted in an adjusted r^2 of .597, meaning that engagement had a 59.7% effect on performance. (Anitha, 2014, pp. 311-316). Therefore, according to

Anitha, although other factors are involved, engagement is the single most important component of performance.

Does Organizational Culture Influence Employee Engagement?

Stanislavov and Ivanov conducted interviews at three Bulgarian casinos, each of which had transformed from one leadership style to another. Casino A went from a commanding leadership style focused on control and competitiveness (with a culture described as a machine) to an affiliative style emphasizing collaboration (and a culture thought of as a political system). Self-reported employee engagement remained level, but the Utrecht Work Engagement Scale (UWES) reflected and increase from 52% to 63% (Stanislavov & Ivanov, 2014, p. 30).

The leadership style of Casino B went from democratic/coaching (and a collaborate/compete culture) to pacesetting/commanding (and compete/control culture). The metaphor used to describe the culture changed from brain/culture to domination/machine. Employees self-reported an increase in engagement from 80% to 86%; however, the UWES scale measurement reported a decrease from 88% to 62% (Stanislavov & Ivanov, 2014, pp. 30-31).

Casino C exhibited a change in leadership style from visionary to affiliative, a change in culture from collaborate/create to collaborate/control, and a change in cultural metaphor brain/culture to organism. Self-reported engagement suffered a severe drop (from 92% to 50%), as did the UWES measurement, from 90% to 67% (Stanislavov & Ivanov, 2014, pp. 30-31).

This study shows that different leadership styles engender different cultures and different levels of employee engagement, both perceived and measured.

How Does Employee Engagement Affect Employee Performance?

Rich, Lepine, and Crawford hypothesized that "job engagement is positively related to task performance" (Rich, Lepine, & Crawford, 2010, p. 620). An 18-question survey was distributed to participants, who rated the items on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). The average score was 4.32, indicating a high self-reported level of engagement. Their supervisors were given a 5-item questionnaire concerning job-performance items, using the same scale. For this, the average score was almost as high, 4.27. For this study, the correlation between engagement and performance was given in the form of β, with β = .35 (Rich, Lepine, & Crawford, 2010, pp. 624-625). Engagement predicted performance, and according to this study, that prediction was correct.

Promotions, Procedural Justice Perception, and Employee Engagement

Each position in AFRL is listed on what is called the Unit Manpower Document, or UMD. Among other data, the UMD lists each position's grade (DR-I, for example) and job series (computer engineer, program manager, etc.). Also, each position is authorized to have only one person listed against it. Each combination has but a finite number, so, absent a sizeable exodus of AFRL workforce members, the opportunities for promotions becomes limited. With a workforce that exhibits a higher level of engagement as far as performance, how will it react when fewer promotions are awarded?

Undoubtedly, there are those employees who, at least in their own minds, will have performed above their peers, a belief which may or may not be reflected on their evaluations. With promotions trending downward as limits on each grade draw closer, more of them will be disappointed as they are told that they will not be advanced. The

response from each person in this circumstance will more than likely be a variant of "but that's not fair!"

Such a statement is a commentary on procedural justice perception, defined as "judgments on the degree to which decision making is viewed as just and fair" and influenced by "both decision-making procedures and interpersonal treatment" (Rubin, 2009, p. 127). The procedures behind the decisions have their detractors, to be certain, but they also act as shields against arbitrary personnel actions and "create perceptions of workplace justice, which in turn influence both workplace attitudes and behaviors including satisfaction, commitment, absenteeism, and turnover, individual performance, and organizational performance (Rubin, 2009, pp. 125-126). Her article studies the relation between procedural justice perceptions, trust in management, organizational satisfaction, and turnover intentions. She states that the importance of the study is based on "new legislative authority" allowing the DOD "to redesign its personnel system, including the ability to adjust processes for paying, rewarding, disciplining and removing employees, addressing employee grievances, and conducting labor relations", and that "Assessing the procedural justice perceptions of these employees sets a baseline against which reforms can be evaluated. This is especially significant considering these rule changes may serve as a model for government wide civil service reform" (Rubin, 2009, p. 126). This described the now-defunct NSPS.

For this study, Rubin provided three hypotheses, of which we will concentrate on the second, that there is a positive relationship between federal employees' perceptions of procedural justice and their satisfaction. She used questions from the 2002 Federal Human Capital Survey and their responses (5-point Likert scale) for her data. Since we

are looking at promotions, we will focus on Question 16, which asked whether promotions were based on merit. The responses were a 44.49% agree/strongly agree rate, and a 33.16% disagree/strongly disagree rate. Rubin calculated ρ to be .5405, which supports her hypothesis (Rubin, 2009, pp. 130, 132-135). Therefore, if AFRL's promotion procedures are perceived as being unfair, this can do nothing but lower employee engagement.

Employee Engagement, Public Service Motivation (PSM), and Job-Goal Orientation

Another possible reason for an employee receiving a substandard evaluation is that the mission of AFRL is not very important to them, or their personal and/or professional goals are not aligned with the goals of AFRL and, as a result, their performance suffers. Wright calls into question such an employee's public service motivation and job-goal orientation, stating that "If individuals do not perceive performance objectives as meaningful or important, they have little reason to strive to achieve them" (Wright, 2007, p. 56). The other major part of Wright's study centered on self-efficacy-"the judgment of his or her own "capabilities to organize and execute courses of action required to attain designated types of performances" (Wright, 2007, p. 56). His research disclosed that $r = .61$ for the correlation between work motivation and job-goal importance, but for work motivation and self-efficacy, $r = .25$ (Wright, 2007, p. 58). In other words, job-goal importance (or lack thereof) had more of an effect on motivation, and therefore engagement.

Lab Demo and Employee Satisfaction

Ghosh, Satayawadi, Joshi, and Shadman presented ten factors that influence employee satisfaction: 1) commitment; 2) employee engagement; 3) job satisfaction; 4) compensation; 5) challenging assignments; 6) training and development; 7) work

environment; 8) work life balance; 9) job stress; and 10) organizational culture (Ghosh, Sataywadi, Joshi, & Shadman, 2013, pp. 292-293). Of the 17 questions used to compare Lab Demo and GS non-supervisors, 14 can easily be linked to one of these 10 factors. With the supervisors, 35 questions were used for comparison. The first 17 questions were the same as for non-supervisors, so, again, 14 are linkable to the 10 factors. The remaining 18 questions, strictly supervisory in nature, could be classified under "job stress," as supervision is inherently stressful. With Lab Demo supervisors scoring higher in satisfaction than their GS counterparts, this suggests that job stress can also be rewarding.

Lab Demo and Evaluation Subjectivity

Lab Demo asks its supervisors to judge the contribution of his or her subordinates to the mission, compared to other employees in the same work center. Without standards and hard numbers to judge performance against those standards, it is the epitome of subjectivity and at best guesswork. With this type of a system in use, Grote's "contrast effect" error is unavoidable, since no standards are available, one cannot help but compare subordinates.

The lack of standards also leads to either the "central tendency" and "negative and positive skew" errors described by Grote. In our case, since reports are drifting towards Zone C, this means that raters are marking their workers towards the middle of the scale from wherever they truly should be, and, as a result, the bulk of the evaluations are in the middle two zones. The subjectivity causes problems with the workforce itself as well. Van der Heijden and Nijhof conducted a study on how employees viewed themselves and how their supervisors viewed them, to measures also found in the Lab Demo system.

Their results showed that the employees thought more highly of themselves than what their supervisors thought of them (van der Heijden & Nijhof, 2004, p. 503). This leads us back to the issue of procedural justice perception, as a subordinate may very well think that he is getting a raw deal on his or her report.

Kleiman, Biderman, and Faley conducted a study of employee perceptions concerning a subjective evaluation system. They mailed a questionnaire (with 20 Likert-type questions concerning the appraisal system) to all 510 employees of a multi-unit government agency in the southeastern United States. Their study found that employees were more likely to consider the appraisal system fair and accurate if they believe the supervisor is qualified to rate them, and if the supervisor discusses career outcomes affected by the evaluation, such as pay or promotion (Kleiman, Biderman, & Faley, 1987, pp. 114, 120).

It may never be possible to develop an evaluation system completely devoid of subjectivity; however, limiting it by measuring employees against job standards instead of measuring them against each other, workers will be able to show their performances stack up against those standards and sense that their rating was not some random guess.

Public Service Motivation (PSM)

Public service motivation, or PSM, is "understood as an individual's predisposition to respond to motives grounded primarily or uniquely in public institutions and organizations," with "motives" defined as "psychological deficiencies or needs that an individual feels some compulsion to eliminate" (Perry & Wise, 1990, p. 368). Research on public service motivation is, relatively speaking, a recent phenomenon with its roots in Perry and Wise's 1990 article "The Motivational Bases of Public Service."

Since then, progress in this area has been anything but linear; rather, it has arrived in waves (Perry & Vandenabeele, 2015, p. 692). Like many nascent fields of endeavor, there are achievements that can be touted, as well as challenges that remain and require confrontation.

The first achievement is the spread of research outside the borders of the United States. Survey data has been collected in over two dozen countries, and researchers in 43 countries have used secondary data to study it (Perry & Vandenabeele, 2015, p. 692).

The second achievement is that PSM is a concept born from public administration. Job satisfaction, organizational commitment, and other such areas of study are also researched as facets of public administration, but their origins are in other fields (Perry & Vandenabeele, 2015, pp. 692-693).

The third and final achievement is the recognition of PSM in fields other than public administration, most notably economics. Economic researchers have used public service motivation as a specification when discussing altruism, and as a tool for studying labor market sorting and economic motivation (Perry & Vandenabeele, 2015, p. 693).

Organizational behavior and human resource management scholars also make reference to PSM. They use it to predict employment interests past the agreeableness personality factor, in "prosocial motivation and job design research," and as both a dependent and a moderating variable in the job demands-resources model of employee engagement. Closer to home, political scientists have used PSM to analyze principal-agent behavior, and have studied its link to prosocial behavior exhibited by public and private employees (Perry & Vandenabeele, 2015, p. 693).

There are challenges with PSM, the first of which is problems concerning quality and quality control. The PSM construct's popularity, along with survey-based measurements, has brought many people into the picture. One result of this is the excessive reliance on cross-sectional surveys and their "easy data." However, the initial euphoria resulting from this rapid supply of information is being tempered by professional self-consciousness about the limitations of such surveys, infusions of more quality studies, and increasing applications of "experimental methods and longitudinal research" (Perry & Vandenabeele, 2015, p. 693).

The second challenge is a simultaneous problem of too much measurement on one hand, and, on the other too little. As far as too little measurement is concerned, the PSM scale set forth by Perry in 1996 has not been revised sufficiently to overcome its limitations, as described by Wright in 2008 and Perry and Vandenabeele in 2010. While there have been attempts to resolve this issue, the difficulty of creating scales and the original's adequacy (in spite of its flaws) have limited the success of these efforts (Perry & Vandenabeele, 2015, p. 694).

On the other side of the coin is the amount of attention paid to measurement overall, resulting in too much measurement. Over the last decade, much of the focus has been on the "development of measurement instruments," showing that the effort expended has produced too much data with too little results to show for it (Perry & Vandenabeele, 2015, p. 694).

Where does PSM research go from here? While there have been numerous proposals, there are three set forth as the most promising for the advancement of this concept.

The first proposal is the need to conduct more research on the individual components of public service motivation, rather than the concept as a whole. This would simplify challenges concerning reliability and validity of measures, to make research easier in both single country and multinational studies, bring attention to facets of public service motivation that are important by themselves, such as compassion or self-sacrifice, and increase the prospects for interactions with other fields or disciplines (Perry & Vandenabeele, 2015, p. 695).

The second direction for future PSM research is to study how it is manifest in different regimes, on the front lines. We have much quantitative information, but are lacking in qualitative data. Studies of motives of public servants or moral exemplars would go far towards fixing this deficiency (Perry & Vandenabeele, 2015, p. 696).

The third and final direction PSM research should be steered towards is improving methods to articulate the attraction of making public policy, with an eye toward the view that this dimension demonstrates a person's inclination toward governance and should be part of a global PSM measurement. However, not every scholar believes that a universal measurement can be obtained, but must instead be garnered from individual political entities (Perry & Vandenabeele, 2015, p. 696).

Remaining in the realm of research for the time being, we are going to examine individual PSM concepts. Bozeman and Su posit that PSM's popularity contributed to problems with developing viable concepts, piling one on top of the other while rarely disposing of any. This leads to what they deem a possible "surfeit of concepts," as evidenced by the fact that number of peer-reviewed articles on PSM went from 0 in 1995 to over 20 in 2012. (Bozeman & Su, 2015, pp. 700-701).

Bozeman and Su set forth Gerring's eight criteria for assessing concepts. Using this framework could alleviate the oversupply of concepts and allow researchers to focus on quality rather than quantity. PSM concepts have scored well on five of these criteria (familiarity, resonance, coherence, depth, and theoretical utility) and not so well on two (differentiation and field utility). The eighth, parsimony, is not considered an important criterion as of yet (Bozeman & Su, 2015, p. 705).

To advance the development of PSM as a theory, Bozeman and Su propose two research programs. The first would be, as they describe it, to "sort out the dispositional versus the environmental aspects of PSM" (Bozeman & Su, 2015, p. 706). Much more progress has been made with regard to understanding PSM as a cause, but an even fuller understanding of this would make research on it as an independent variable even more valuable. The second would be to validate PSM in laboratory studies, which have a strong tradition in related concepts (Bozeman & Su, 2015, p. 706).

Even with all these concepts, questions still arise as to what is the true driver of PSM. One proposed answer is a job demands-resources (JD-R) approach. JD-R theory is an organization theory that stems from research into job stress, looking into why some employees lose enthusiasm and burn out. The theory has become very popular and has also been used to predict work engagement and performance. There are six building blocks that summarize JD-R theory (Bakker, 2015, p. 724): 1) the uniqueness of all organizations; 2) the separate impacts job demands and job resources have on employee well-being; 3) the interaction of job demands and job resources leading to indications of employee well-being; 4) the resources employees have access to in order to cope with job demands; 5) the motivation for engaged employees to remain so; 6) the burnout

employees may experience due to high job demands, causing them to accumulate fatigue leading to self-undermining behaviors and higher job demands (Bakker, 2015, pp. 724-725).

Bakker sets forth seven propositions to support the JD-R approach to PSM. It: 1) lessens the positive relationship between daily job demands and exhaustion; 2) dampens both the negative relationship between daily exhaustion and performance and the positive relationship between daily exhaustion and self-undermining behaviors; 3) softens the positive relationship between daily job resources and work engagement; 4) dampens the the positive relationship between daily work engagement and performance, as well as the positive relationship between daily work engagement and job crafting; 5) enjoys a negative relationship with daily exhaustion; 6) enjoys a positive relationship with daily work engagement; and 7) displays a positive relationship with performance (Bakker, 2015, p. 729).

Bakker's argument is that when public administrators begin their careers, they are highly motivated, and that it is this motivation that determines how they cope with their daily job demands and resources. Because of their call to public service, part of their motivation consists of gathering and using their resources to remain engaged and maintain a high level of performance. On the other hand, if their resources are consistently low and demands consistently high, even highly motivated civil servants can lose that motivation and encounter a strengthened loss cycle and weakened gain cycle (Bakker, 2015, p. 730).

Leaving the world of PSM theory and entering the realm of PSM in action, our first task is to try and ascertain why people want to work for the government. After all, it

is common knowledge that the grass is greener on the private sector side of the fence. Thus, we turn now to the subject of motivation, without which any improvement in public employee performance is improbable, if not impossible; let us examine how it interacts with public service. Perry, Hondegem, and Wise, working from three propositions set forth by Perry and Wise 20 years previously, did just that.

Proposition 1: "The greater an individual's public service motivation, the more likely the individual will seek membership in a public organization" (Perry, Hondeghem, & Wise, 2010, p. 683). This goes without saying. People tend to seek out organizations or groups that are compatible with their own dispositions, where they can satisfy their public service desires. The famous words of John F. Kennedy come to mind: "Ask not what your country can do for you, but what you can do for your country." They carry this spirit within them, and they seek kindred spirits.

Proposition 2 says: "In public organizations, PSM is positively related to individual performance" (Perry, Hondeghem, & Wise, 2010, p. 684). Once again, this stands to reason. They cite two previous studies in which PSM was found to positively correlate with both self-reported performance evaluations and actual appraisal ratings. While this proposition was not found to be universally true across all the studies quoted, PSM was deemed to be a "driver of performance" (Perry, Hondeghem, & Wise, 2010, p. 685)

The last of our three propositions has to do with incentives and PSM: "Public organizations that attract members with high levels of public service motivation are likely to be less dependent on utilitarian incentives to manage individual performance effectively" (Perry, Hondeghem, & Wise, 2010, p. 686). This is especially relevant to the subject of this study, the contribution-based compensation system of Lab Demo, as the

final score of an individual's evaluation determines how much of a bonus he or she will receive. If an employee is truly motivated by PSM, is a monetary bonus necessary? It may not be, if our next article is considered credible.

Homberg, McCarthy, and Tabvuma conducted a meta-analysis on the link between PSM and job satisfaction. In the introduction to their study, they cited studies supporting the existence of PSM in Denmark, Malta, South Korea, China, Switzerland, and Australia, as well as multi-national efforts. Also mentioned are studies in PSM's traditional home, public administration, as well as economics and human resources management. However, in spite of all this favorable evidence, the concept of PSM is not without detractors. One study failed to find a significant difference in PSM between the public and private sectors. Another did not find a lower regard for extrinsic rewards on public workers. Still another concurred with the notion of those who enjoy helper others tend to gravitate to the public sector, but failed to find a link between that choice and salary preference (Homberg, McCarthy, & Tabvuma, 2015, pp. 711-712).

For their study, the authors searched several internet-based sources with a variety of search terms, ranging from 1990-2013, as well as some working paper drafts scheduled for presentation at meetings or publication in a work dedicated to PSM. The resulting 79 studies were then whittled down by removing those that did not report a bivariate correlation between PSM and job satisfaction, and those that used only a single-item measure of PSM. This left 28 studies (Homberg, McCarthy, & Tabvuma, 2015, pp. 713-714).

In order to make the data from the different studies comparable, effect sizes were used. Different moderating variables (published vs. unpublished studies, dimensional vs.

global measure, U.S. vs. non-U.S., and whether or not particular jobs offered the opportunity for public service) were established (Homberg, McCarthy, & Tabvuma, 2015, p. 714).

At the end of all this, the study found that a positive correlation between PSM and job satisfaction exists across studies, that it does not suffer from publication bias, and support exists for the concept of public service opportunity (Homberg, McCarthy, & Tabvuma, 2015, p. 718)

If one is to believe that public service motivation exists, like many other human resources-related concepts, it is imperative that those who occupy positions of supervision and leadership ensure that PSM is activated in their subordinates. Pederson opines that the primary focus of research into PSM's positive effects has been cultivation (fostering and sustaining), and wrote his article concentrating on activation (actively engaging an individual's PSM) (Pederson, 2015, p. 734).

His belief in this research's pertinence lies in the fact that while numerous studies have shown that public sector employees demonstrate more PSM than their private sector counterparts, not all civil servants are motivated by the same things. Even those public workers with the highest levels of PSM are occasionally directed by extrinsic factors, such as pay, work-family life balance, or self-achievement needs. Research examining cultivation is good; however, if research into activation is added to the mix, public sectors can use a combination of practices based on both sides to optimize organizational performance. To this end, his article asks if external activation efforts can be used to capitalize on PSM's positive effects, whether or not PSM can be activated by even low-

intensity interventions, and how efforts aimed at PSM activation compare with efforts aimed at other types of internal motivation (Pederson, 2015, p. 735).

Pederson submitted two hypotheses for consideration. The first is based on cultivation-activation effect and expectations grounded in Perry and Vandenabeele's institutional theory. That theory states that while institutions are social constructs present in every societal layer, they do not dictate individual motivation. Instead, the relationship between the institution's content and the individual's motivation is mediated by that individual's self-concept and self-identities. Public sector employees have different self-identities (gender, job, altruistic individual, etc.), and different situations promote salience of different self-identities. Whichever identity is salient at a particular point in time determines how that individual acts. This may help explain why external PSM motivation may set behavioral changes in motion (Pederson, 2015, p. 736).

Additionally, in their 1990 work, Perry and Wise posited that PSM is positively correlated with individual performance. It guides work behavior toward the accomplishment of tasks or goals, partly by increasing commitment, and partly by making workers delve into work marked with task significance. This leads to their first hypothesis, that "the external motivation of PSM increases the amount of time that a person is willing to spend completing a task" (Pederson, 2015, p. 736).

Perry and Wise's second hypothesis is based on self-determination theory (SDT), which helps distinguish between PSM and an individual's need for feelings of self-importance, as well as between intrinsic and extrinsic motivation. In SDT, extrinsic motivation is broken down into four forms:

1. External regulation: external pressure, or behavioral self-regulation with the goal of either obtaining an external reward, or avoiding an external constraint.
2. Introjected regulation: internal pressure, or behavioral self-regulation driven by pride, shame, or needs of self-approval or approval by others.
3. Identified regulation: behavioral regulation guided by personal values or goals.
4. Integrated regulation: values of given activities are identified with by individuals, so much so that it is an internalized part of their behavior and self-identity (Pederson, 2015, p. 737).

PSM falls into one of the last two types, and efforts aimed at its external activation are expected to replicate the positive effects of more self-determined motivation, as opposed to less. The second hypothesis is "the amount of time a person is willing to spend completing a task is greater for the external activation of PSM than for the activation of motivation relating to a need for self-performance" (Pederson, 2015, p. 737).

To test these hypotheses, a survey was administered to 746 law students, who were instructed to complete the survey at the beginning of a 15-minute break sandwiched between 2-hour lectures. Toward the end of the survey, the question "In the near future, you will be invited to participate in a survey about your daily life. How many minutes are you at most willing to spend on completing this survey?" (Pederson, 2015, p. 738) was asked. This exact question was asked of the control group (designated Z^0) as well as the three other groups in the project, but these groups were also given the following additional information:

- Treatment PSM^P (public interest): "Your participation will help ensure the development of society and thus serves the public interest."
- Treatment PSM^C (compassion): "Your participation will help ensure that citizens in need are aided in the best possible way."
- Treatment IEM (introjected extrinsic motivation): "You have been chosen to participate because of your background and special knowledge" (Pederson, 2015, p. 738).

The findings of this experiment confirm the first hypothesis. PSM^P and PSM^C were willing to spend an average of 8 minutes and 29 seconds on the survey, as opposed to 7 minutes and 35 seconds for Z^0. IEM treatment members were willing to spend even more time on the survey, 12 minutes and 53 seconds. On the other hand, the second hypothesis was not supported by the findings. The PSM treatments did not exhibit larger or smaller effects compared to the IEM treatment. Both showed positive effects, but the effect sizes were not statistically significant (Pederson, 2015, p. 739).

Public service motivation can also have a mediating effect on human resources practices and their resulting employee outcomes. Strategically driven "high-performance human resources practices" (HPHRPs) activities aimed at improving organizational function have long been advocated, and, studies have produced a body of evidence supporting the relationship between HPHRPs and performance in both the public and private sectors. However, most of these studies have targeted financial returns, profitability, and service performance; what is lacking are studies focusing on the effect of HPHRPs on employee attitudes and behaviors. So, while HR scholars admit that the workforce is what achieves superior performance, the mechanisms whereby workforce

attitudes and behaviors are affected by HPHRPs are not well understood (Mostafa, Gould-Williams, & Bottomley, 2015, p. 747).

Mostafa, Gould-Williams, and Bottomley examine one such mechanism, PSM, particularly whether it acts "as a motivational mechanism that explains the relationship between HPHRPs and affective organizational commitment and organizational citizenship behaviors (OCBs)" (Mostafa, Gould-Williams, & Bottomley, 2015, p. 748). To do so, they draw on social exchange theory, which states "social exchange comprises actions contingent on the rewarding reactions of others, which over time provide for mutually and rewarding transactions and relationships" (Mostafa, Gould-Williams, & Bottomley, 2015, p. 748). They also draw on process theory, in which the organizational environment influences PSM through socialization, effective job design, and performance feedback (Mostafa, Gould-Williams, & Bottomley, 2015, p. 748).

Outcomes were chosen for these reasons:

- "Affective commitment and OCBs are central and prominent responses to positive social exchange relationships."
- "OCBs are conventionally considered 'discretionary behaviors' that do not form part of the formal employment contract but are a reflection of positive social exchanges."
- "Both affective commitment and OCBs have been linked with enhanced organizational performance and thus have implications that extend beyond the individual" (Mostafa, Gould-Williams, & Bottomley, 2015, p. 748).

Four hypotheses were set forth:

1. "HPHRPs will be positively related to employees' (a) affective commitment and (b) OCBs."
2. "HPHRPs will be positively related to employees' PSM."
3. "HPHRPs PSM will be positively related employees' (a) affective commitment and (b) OCBs.
4. "PSM will mediate the relationship between HPHRPs and employees' (a) affective commitment and (b) OCBs" (Mostafa, Gould-Williams, & Bottomley, 2015, pp. 749-750).

One thousand questionnaires were distributed among various Egyptian public sector employees, of which 671 were returned. Multi-item scales from existing studies were used to measure constructs whenever possible, and all items were measured on 7-point Likert scales ranging from "strongly disagree" (1) to "strongly agree" (7). The five HPHRPs in this study were measured with 20 items in previously published studies. PSM was measured using Perry's 13-item scale, as refined by Giauque et al. Affective commitment was measured on an abridgement of Meyer, Allen, and Smith's scale. OCBs were measured with an abridgment of Lee and Allen's scale (Mostafa, Gould-Williams, & Bottomley, 2015, pp. 750-751).

Overall, the findings supported all four hypotheses, and that how employees view HR practices affects the outcomes of those practices. For managers desirous of improving the social exchanges in their respective public sector organizations, it is imperative that sufficient resources are committed to the implementation of HR practices (Mostafa, Gould-Williams, & Bottomley, 2015, p. 753).

Is Engagement the Same as Motivation?

This is a good place to ask and answer this question, as Marciano answers in the negative by way of a short explanation of what he believes the differences to be. "Engagement is similar to, but not synonymous with, motivation. Engagement refers to an intrinsic, deep-rooted, and sweeping sense of commitment, pride, and loyalty that is not easily altered. In contrast, motivation level is strongly influenced by external factors, especially expectations that certain efforts or accomplishments will yield valued rewards, such as a financial bonus for meeting a quarterly sales objective" (Marciano, 2010, p. 40). From there, he proceeds to teach how to build engagement using the RESPECT (Recognition, Empowerment, Supportive Feedback, Partnering, Expectations, Consideration, Trust) model.

Money as a Motivator

Since we are now on the subject of motivation, it is a good time go to the heart of Lab Demo, the pay-for-performance aspect, and discuss whether or not monetary bonuses truly act as motivators. Campbell, Campbell, and Chia detailed five problems associated with merit pay: 1) measurement issues (can individual performance be measured, and can the performance of two individuals with interrelated jobs be distinguished from each other?) (Campbell, Campbell, & Chia, 1998, p. 132); 2) performance appraisal and feedback issues (employees must believe that the distinctions in pay based on performance are fair) (Campbell, Campbell, & Chia, 1998, p. 133); 3) reward desirability issues (a sufficient amount of funds must be reserved to make the differences in pay meaningful (Campbell, Campbell, & Chia, 1998, p. 134); 4) "system noise" issues (such as the timing of the merit pay and inconsistencies in the system) (Campbell, Campbell, & Chia, 1998, p. 134); and 5) unintended consequences, or costs that abrogate any benefits

obtained by the use of merit pay (Campbell, Campbell, & Chia, 1998, p. 134). The disregard or mishandling of one or any combination of these can result in a breakdown of the system, leading to a disgruntled workforce and a human resources nightmare.

Belle and Cantarelli conducted a study of performance-related pay with executives in the Italian government. They noted that an emphasis on it had not abated since the 1980s, when it was first adopted, even though it had inconsistent results on the performance of civil servants (Belle & Cantarelli, 2015, pp. 99-100).

Belle and Cantarelli investigated how monetary rewards affected the efforts of civil servants and how other motivations moderated that relationship (Belle & Cantarelli, pp. 104-105). They distributed a survey in 9 of 13 central departments in Italy's central government by one manager within each department. All managers were kept unaware of the statistical operations to be performed on the collected data and the hypotheses: 1) civil servant's efforts are not enhanced by monetary incentives; 2) civil servants with higher levels of intrinsic motivation at the baseline experience a weaker effect from monetary incentives; 3) civil servants with higher extrinsic levels of motivation at the baseline experience stronger effects from monetary incentives; and 4) the level PSM among civil servants does not affect the effect of monetary incentives (Belle & Cantarelli, 2015, pp. 101-103).

All four of the hypotheses were supported by the resulting data. The authors claim that this study provides "much needed experimental evidence of the effects of financial incentives on job effort in public sector organizations" (Belle & Cantarelli, p. 114), and is at least a start in the process of filling the void that leads to the prevailing attitude that merit pay does not deliver on its promise.

While Belle and Cantarelli's study was conducted in Italy, another was conducted in the United States 2 years earlier. However, Stazyk asked only one question, "Do employees in cities with variable pay systems demonstrate lower levels of public service motivation and job satisfaction than employees without performance-related pay?" (Stazyk, 2013, p. 252).

Stazyk, like Belle and Cantarelli, refers to much negative argument related to merit pay: it is difficult to implement (Stazyk, 2013, p. 254); public organizations lack the resources to provide large-enough remuneration to permanently affect employee behavior and organizational culture sufficiently to lead to long-lasting performance gains (Stazyk, 2013, p. 254); it is difficult to design an manage a merit pay system that accurately gauges and links employee performance to rewards, and, without a standardized and clearly designed system, employees tend to believe they are not being assessed and compensated fairly (Stazyk, 2013, p. 254); merit pay has failed to deliver intended results, "especially with respect to employee motivation, commitment, and productivity" (Stazyk, 2013, p. 254); merit pay is thought to cause employees to displace goals, "ignore due process outcomes," and disregard "other relevant organizational values" (Stazyk, 2013, p. 254); variable pay systems assume knavish behavior and action on the part of public servants, which are best controlled by extrinsic means (Stazyk, 2013, p. 255); these systems overlook the knightly behaviors of many government employees (Stazyk, 2013, p. 255); attempts to control the aforementioned knavish behaviors may actually reduce public service motivation and influence civil servants to work against organizational concerns for personal gain (Stazyk, 2013, p. 255).

Much of these criticisms of merit pay stem from what is known as motivation "crowding" theory, which demonstrates that "financial and performance-based incentives can reduce intrinsic motivation" and "might affect one's civic behavior" (Stazyk, 2013, p. 255). Simply put, incentives can increase one's extrinsic motivation, but, at the same time, decrease his or her intrinsic motivation.

Stazyk used ordered logistic regression to regress job satisfaction on "public service motivation, variable pay, reward satisfaction, goal and role clarity, and human resources red tape variables" (Stazyk, 2013, p. 260). His experiment produced two main results: first, an employee will probably have higher job satisfaction in cities with variable pay systems rather than those with standard pay systems, and that public service motivation does matter; and, second, employees are less likely to exhibit lower levels of job satisfaction under variable pay systems, and when they have higher public service motivation (Stazyk, 2013, pp. 262-264).

Returning to Italy, Belle also conducted an experiment using motivation crowding theory. He randomly assigned 300 nurses into 12 groups, each with 25 members. Each individual nurse was assigned to one of three reward types: fixed pay, fixed pay with a monetary reward based on performance, or fixed pay with a symbolic reward. The nurses were part of a project aimed at assembling surgical kits and readying them for shipment to a former war zone, and the portion of the project pertinent to this experiment was four hours long (Belle, 2015, p. 233).

The fixed pay group was told that they would be paid their normal wage; the nurses assigned to the merit pay group were promised a decreasing-payment format reward, with bonuses decreasing at certain thresholds as more kits were produced. The

symbolic reward group was promised a certificate outstanding contribution (Belle, 2015, p. 233).

The participants were also divided into two additional categories: disclosed (each participant's performance would be posted on a bulletin board in the staff room) or secret (each participant would be notified of their performance, but it would be anonymous and would not by disclosed publicly). Finally, half of the nurses met with a beneficiary from the target area, and the other half did not (Belle, 2015, pp. 233, 235).

Performance was measured as the number of correctly assembled surgical kits each member produced during each individual's 3-hour shift. Two hypotheses were tested:

1. "The visibility of rewards will moderate the effect of monetary rewards on job performance such that the monetary rewards will have a stronger effect when they are secret and a weaker effect when they are visible."

2. "A perceived prosocial impact will moderate the interaction between monetary rewards and reward visibility such that there will be a stronger (negative) interaction for activities with a higher perception of prosocial impact. Thus, there will be a three-way interaction between monetary rewards, reward visibility, and perceived prosocial impact" (Belle, 2015, pp. 232-233).

The following results were observed:

1. The disclosure of performance and pay data had a strong effect on fixed-pay nurses, but not on the merit-pay group.

2. Merit pay greatly enhanced the performance of nurses in the secret condition, but only marginally so for those in the open condition.

3. Negative interaction between merit pay and visibility was stronger for those nurses who met with the beneficiary.

4. The interaction or merit pay and beneficiary contact was stronger in the open condition than in the secret condition (Belle, 2015, pp. 236-237).

Summary

With an $18 trillion-dollar debt, the federal government faces ever-increasing pressure to spend its money wisely, resulting in cutthroat competition for funding, which AFRL must take part in. Research itself is expensive, not to mention those conducting the research, and research programs are subject to review by the Scientific Advisory Board and the Investment Strategy Review. Failure to satisfy the members of either of these panels will lead to the cancellation of programs, no matter how much progress has been made. Programs have been reduced or even eliminated, with employee morale and engagement being reduced alongside them. Even more mundane facets of military existence suffer. The Information Directorate has been requesting a security fence for a decade, a request which has never ascended to higher than third on the priority list. With the grass at least appearing to be greener on the private sector side of the fence, the pressure is on for all DOD branches to recruit and retain the most highly qualified scientists and engineers. It has resorted to a personnel management system, the main component of which is pay-for-performance, that may or may not deliver on what its proponents have promised, and can lead to payouts that can cause salaries to be a larger portion of the budget, thanks to promotions and bonuses. The question is, does Lab Demo provide the American taxpayer with enough bang for the buck?

Chapter III

METHODOLOGY

Case Study

This case study examines indicators of employee engagement among various segments of the federal workforce, namely, the entire workforce covered by the Lab Demo project, AFRL's General Schedule employees and the entire federal workforce itself, with different segments covered in each of the three below-listed research questions.

Research Questions

Lab Demo was originally designed by the Air Force, with participatory assistance from the DOD and OPM. Their aims were to "demonstrate that the effectiveness of DOD laboratories can be enhanced by allowing greater managerial control over personnel functions and, at the same time, expanding the opportunities
available to employees through a more responsive and flexible personnel system" and correct "several major inefficiencies, which hinder management's ability to recruit and retain the best-qualified personnel" (National Archives and Records Administration, 2010, pp. 53081-53082).

The supporters of Lab Demo point to four claims in the Federal Register touting its successes: "(a) Increased quality of the workforce and the Laboratory products they produce; (b) increased timeliness of key personnel processes; (c) trended workforce data that reveals increased retention of 'excellent contributors' and increased separation rates

of 'poor contributors;' and (d) increased employee satisfaction with the Laboratory" (National Archives and Records Administration, 2010, p. 53082). To evaluate the effectiveness of this program, three human resources management-related questions were asked, each of which has a connection to employee engagement.

Has employee performance improved overall? Research Question 1 was whether or not employee performance has improved overall, as reflected by evaluation scores and promotion rates.

Anitha lists workplace environment (with the giving of positive feedback and the freedom to voice personal concerns), leadership (with the responsibility to communicate that the efforts of the rank-and-file employees are critical for organizational success), compensation, and organizational polices as four of the factors that determine an employee's level of engagement (Anitha, 2014, pp. 311-312). These four factors all tie into the Lab Demo project. Are the employees comfortable with the workplace environment? Do they receive positive feedback? Do they believe they have the freedom to air their concerns? Are they of the opinion that their work is important? Do they believe that they are fairly compensated? Are they happy with organizational policies? If the answers to these questions are in the affirmative, their level of engagement will increase, as well as will their evaluations.

The hypotheses examined in answering this research question were:

$H_0(1)$: Employee performance overall has not improved significantly since the introduction of Lab Demo.

$H_1(1)$: Employee performance overall has improved significantly since the introduction of Lab Demo.

$H_0(2)$: There is no strong correlation between promotions and evaluations meeting or exceeding standards.

$H_1(2)$: There is strong correlation between promotions and evaluations meeting or exceeding standards.

The data for these analyses were derived from the annual Lab Demo reports from 1997-2011, and the entire population of Lab Demo employees was examined.

In answering this research question, the first test conducted was a correlation analysis for Pearson's *r* on the total number of reports and the numbers of reports that fall into each zone. Strong direct or strong inverse correlations would indicate a pattern of consistency as to how subordinates are rated. The independent variable is the total number of reports, and the dependent variables are the number of reports in each zone.

Next, a series of four regression analyses, one for each of the four zones and the percentage of the total reports in each of them were conducted. Higher values for Pearson's *r* indicated a higher level of consistency in grading over the years spanned by the project, and less room for evaluator errors. For each of these regressions, the years 1997-2011 were the independent variable, with the percentages of reports in the zone for each of those years being the dependent variable.

These four regressions were followed by another, conducted on the overall average score for each year. Any increase or decrease in workforce quality reflected in how evaluations are distributed throughout the four zones should also reflect in the overall average score for each year. As in the previous regression analyses, the years covered are the independent variable, but the dependent variable are the average scores.

The final test conducted in determining the direction of workforce quality based solely on evaluations was a correlation analysis for Pearson's *r,* with the desired result being the discovery of any correlation between the percentages of reports in each zone, along with the percentage of reports in each Zone and the average score.

Turning to the factoring in of promotions, the first test conducted was a regression analysis on the rate of satisfactory evaluations (Zones B, C, and D). Zone A was excluded for the obvious reason that unsatisfactory evaluations will not result in an employee being promoted. This test is an attempt to determine the level of consistency in the percentage of satisfactory evaluations. The years of the reports is the independent variable, and the percentage of satisfactory reports is the dependent variable.

The next test was also a regression analysis, this time for the promotion rate, again looking for consistency over the period of the project. The years of the project are the independent variable, and the promotion rates the dependent variable.

Finally, a correlation analysis for Pearson's *r* was conducted, between the percentage of satisfactory evaluations and promotion rates. If there is any consistency to the promotion process at all, an important component of employees' perception of it, we will discover a strong correlation between the number of satisfactory employee evaluations and the number of promotions. If not, AFRL runs the risk of disgruntled workers with the notion that no matter how well they do, the promotion process is rigged.

Are more substandard employees separated, and outstanding employees retained? Research Question 2, pertaining to claim (c), that trended workforce data shows an increased separation of substandard employees and retention of outstanding contributors, is whether or not excellent contributors are increasingly retained and poor contributors

are increasingly separated. Not only will this confirm or refute claim (c), but it also will add additional information to the answer for Research Question 1. Fewer poor contributors and more outstanding contributors would, at least in theory, increase the quality of the workforce. To answer this question, the following hypotheses were tested:

$H_0(3)$: Substandard contributors were not separated at a significantly higher rate.

$H_1(3)$: Substandard contributors were separated at a significantly higher rate.

$H_0(4)$: Outstanding contributors were not retained at a significantly higher rate.

$H_1(4)$: Outstanding contributors were retained at a significantly higher rate.

The data source is once again the annual Lab Demo reports from 1997-2011. Instead of using the entire population of Lab Demo employees, only those whose evaluations fall into Zone A (the poor contributors) and Zone D (the outstanding contributors) were used.

The test consisted of two correlation analyses for Pearson's *r*; one for those is Zone A, and the other in Zone D. In addition, two regression analyses were conducted on the same sets of data, and the results compared to determine if substandard employees are being separated or superior performers are being retained more consistently.

For the correlation analyses, the number of reports is the independent variable, and the number of people separated is the dependent variable. If both of the null hypotheses are accepted, it would indicate that the number of people separated (in the case of Zone A employees) or retained (Zone D employees) had nothing to do with the number of reports generated (obviously dictated by the size of the workforce) and therefore would not be "trended" in that sense. For the regression analyses, the years of

the reports were the independent variable, and the separation rates were the dependent variable.

Do Lab Demo employees exhibit more satisfaction? Research Question 3, regarding claim (d), is whether or not employees exhibit more satisfaction working at AFRL. The expression of this satisfaction can be expressed in different forms, including retention and positive responses to job satisfaction surveys. To answer this research question, the following hypotheses were tested:

$H_0(5)$: There is no significant difference between Lab Demo separation rates and civil service separation rates.

$H_1(5)$: There is a significant difference between Lab Demo separation rates and civil service separation rates.

$H_0(6)$: Lab Demo non-supervisors do not exhibit significantly higher job satisfaction than their General Schedule counterparts.

$H_1(6)$: Lab Demo non-supervisors do exhibit significantly higher job satisfaction than their General Schedule counterparts.

$H_0(7)$: Lab Demo supervisors do not exhibit significantly higher job satisfaction than their General Schedule counterparts.

$H_1(7)$: Lab Demo supervisors do exhibit significantly higher job satisfaction than their General Schedule counterparts.

For this research question, there were multiple data sources. The data for H(5) were derived from the annual Lab Demo reports covering 2001-2012 (for Lab Demo data), and the Bureau of Labor Statistics web site (for federal government workforce data). For H(6) and H(7), the data source was the 2010 Employee Satisfaction Survey,

and the populations examined were both the pool of Lab Demo employees (n = 3289) and the pool of General Schedule employees (n = 1682) employed by AFRL in 2010. Each of the two labor pools were separated into two categories, supervisory and non-supervisory.

For H(5), the independent variable was the span of years covered in the study (limited due to the availability of Bureau of Labor Statistics data), and the dependent variable was the separation rate. For H(6) and H(7), the labor pool and supervisory status was the independent variable, and the positive response rates was the dependent variable.

Engaged, satisfied employees tend to remain with their employers, so the first analysis conducted was a comparison between Lab Demo employees that separate and those that separate from the entire federal government. To test this, a *t* test was conducted to see how much difference there is in the separation rates of the Lab Demo workforce and the federal government.

To have the null hypothesis accepted in this case would indicate that the extra money spent on promotions and bonuses under the Lab Demo project was not enough of a factor to increase retention rates, especially among highly-educated, highly-qualified, and, therefore, highly expensive personnel.

Employees also quite often express their satisfaction levels by answering surveys. In 2010, when Lab Demo was 13 years old, data was recorded for both Lab Demo and AFRL's General Schedule employees. Two *t* tests (between Lab Demo/General Schedule supervisors and between Lab Demo/General Schedule non-supervisors) were conducted to see which labor pool scored higher in job satisfaction.

A comparison of responses from the original Lab Demo survey and the 2010 survey displays an undoubted increase in positive responses in all areas; the question is whether or not that increase was large enough to match the job satisfaction of the General Schedule employees, since those managing the Lab Demo project seek to gain approval from the OPM for implementing across all personnel in AFRL, completely substituting it for the General Schedule system. To accept both of the null hypotheses here would indicate that in the then 13-year lifespan of the Lab Demo project, all of the planning, effort, time, and money spent was simply not enough to successfully accomplish that goal.

Data Reliability and Validity

All of the data used in this study was previously published and publicly available government data, so there were no concerns with reliability or validity.

Study Limitations

Air Force Policy concerning the conducting of surveys amongst its personnel is very restrictive and any request to do so must meet an exacting set of standards before it is approved. The Air Force Survey Office, located in the Air Force Personnel Center at Randolph Air Force Base, Texas, disapproved this researcher's request. Therefore, this study relies on the small set of data outlined above.

Summary

The General Schedule system has been in use since 1949, with very few alterations. The failure to make changes led to complacency throughout the federal work force, including those personnel employed by AFRL. Lab Demo is an attempt to fix problems in this small corner of the civil service, especially those involving retention,

since the government is at a compensatory disadvantage when competing with the private sector for highly-educated scientific talent.

The annual reporting on how the system is working reportedly supports the claims by those who manage it. Beyond the reports, beneath the numbers, lies the truth of just how effective the system is in solving the personnel problems that run rampant throughout AFRL. Conducting the tests described above discovered that truth. If the null hypotheses for Research Question 1 was accepted and the null hypotheses for Research Questions 2 and 3 were rejected, it can be shown that Lab Demo system was not designed or implemented to address the employee engagement problem, and that the extra money it has cost the American taxpayer has not been wisely utilized, as under-engaged employees will continue to exhibit the same behaviors as they did under the General Schedule system.

Chapter IV

RESULTS

The results and observations annotated in this chapter are designed to be analytical, allowing for explanation of the author's conclusions and allowing the reader to draw his or her own. While these findings are not representative of all factors that contribute to employee engagement (or motivation, for that matter), they are intended to contribute to further study and improvements to the Lab Demo project, with the ultimate objective being the needed approval from the OPM to make it the official system of AFRL. Otherwise, Lab Demo may meet the same fate as the National Security Personnel System.

This chapter contains an overview of the findings, including how evaluations, scores, promotion rates, and retention rates are affected over time, how evaluations and scores, evaluations and promotions, and evaluations and retention interact, how AFRL employees compare with the federal government as a whole, and how AFRL's Lab Demo employees stack up against their General Schedule counterparts. Finally, employee satisfaction is measured and discussed.

Research Question 1

Hypothesis 1. A total of seven tests were conducted to test our first null hypothesis, $H_0(1)$, that employee performance overall had not improved significantly since the advent of Lab Demo.

Table 2. Correlation Analysis Between Total Reports and Reports in Each Zone

	A	*B*	*C*	*D*	*AVERAGE OCS SCORE*
A	1				
B	0.79960503	1			
C	-0.7750468	-0.9176445	1		
D	0.70574904	0.85284874	-0.9891247	1	
AVERAGE OCS SCORE	-0.6519535	-0.8194655	0.60312645	-0.5011851	1

Table 2 displays the results of the correlation analysis for Pearson's *r* between the total number of reports and the number of reports that are in each zone. The varying results for Pearson's *r*, ranging from a strong inverse correlation (between Total Reports and Zone D) to a strong direct correlation (between Total Reports and Zone C) suggests there is little in the way of a consistent relationship between total reports and how they are distributed among the four zones. This could mean that either Lab Demo supervisors evolved as they gained experience in the system, or that any of Grote's nine common rater errors crept into the picture. Strong enough correlations or inverse correlations to show consistent patterns in the distribution of evaluations across the four zones could not be found, forcing a deeper dive into the data recorded for each zone.

To further examine the trends of Lab Demo evaluations, four regression analyses were conducted, one for each of the four zones, for determining the direction in which each zone trended, and whether or not the trends were significant. In all of these analyses, $\alpha = .05$. The first, examining Zone A, is displayed in Table 3.

Table 3. Regression Analysis for Zone A Reports

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.622136095
R Square	0.38705332
Adjusted R Square	0.339903576
Standard Error	0.007089677
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.000412614	0.000412614	8.209022624	0.013266223
Residual	13	0.000653426	5.02635E-05		
Total	14	0.00106604			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	2.441812857	0.849075156	2.875850083	0.012996456	0.607497504	4.276128211	0.607497504	4.276128211
YEAR	-0.001213929	0.000423689	-2.865139198	0.013266223	-0.002129253	-0.000298604	-0.002129253	-0.000298604

In this analysis, Pearson's $r = .62$, indicating a direct, but not a strong, correlation between how the system has evolved over time and the percentage of Zone A reports. The coefficient is -.0012, indicating a downward trend in Zone A reports. That, combined with the P-value being less than α, indicated that this regression can be used to support the rejection of the null hypotheses. The downward trend in unsatisfactory reports alone is enough to indicate an improvement in workforce quality, since there is no other direction in which to go but upward. Since $P < \alpha$, this was significant.

Table 4. Regression Analysis for Zone B Reports

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.718725555
R Square	0.516566424
Adjusted R Square	0.479379226
Standard Error	0.021999294
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.0067228	0.0067228	13.89097456	0.002535852
Residual	13	0.006291596	0.000483969		
Total	14	0.013014396			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	9.92254	2.634683481	3.76612222	0.002354149	4.230652388	15.61442761	4.230652388	15.61442761
YEAR	-0.0049	0.001314709	-3.727059775	0.002535852	-0.007740257	-0.002059743	-0.007740257	-0.002059743

Continuing to Table 4, examining Zone B reports, Pearson's *r* was slightly higher than for that of Zone A (.72), which at least places it in what some describe as a strong correlation; however, there was over one-fourth of the variance in reports explained by factors other than system evolution. Like Zone A, Zone B reports have a downward trend, with the coefficient being -.0049, and since the P-value (.0025) < α, this was significant. Again, this analysis was used to support the rejection of the null hypothesis. Even though Zone B reports are considered satisfactory, the fact that they exhibited a downward trend telsl us that they could either go downward to Zone A, or upward to Zone C or Zone D. Since Zone A reports also exhibit a downward trend, the same trend in Zone B means that they are only going upward.

Table 5. Regression Analysis for Zone C Report

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.921135477
R Square	0.848490566
Adjusted R Square	0.836835994
Standard Error	0.054708181
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.217899004	0.217899004	72.80323801	1.09677E-06
Residual	13	0.038908806	0.002992985		
Total	14	0.256807809			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-55.14103619	6.551971282	-8.415945952	1.2783E-06	-69.29570959	-40.98636279	-69.29570959	-40.98636279
YEAR	0.027896429	0.003269439	8.532481351	1.09677E-06	0.020833235	0.034959622	0.020833235	0.034959622

Table 5, which recorded the regression analysis for Zone C reports, showed a couple of major differences. First, Pearson's *r* jumped all the way to .92, showing the first very strong direct correlation between the system evolution over time and the percentage of Zone C reports. Second, this analysis also yielded the first positive coefficient, .0278, indicating a zone exhibiting an upward trend in reports. Since P (.000001) < α, this was a significant finding. Therefore, this analysis can be used to support a rejection of the null hypothesis. Since this zone was considered the upper half of the satisfactory zone, it garnered a larger share of the reports, and both of the zones below it exhibited diminished proportions of reports, it can be stated that this zone absorbed what Zone A and Zone B lost, thus contributing to at least the perception of an improved workforce.

Table 6. Regression Analysis for Zone D Reports

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.956880735
R Square	0.91562074
Adjusted R Square	0.909130028
Standard Error	0.030686951
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.132840577	0.132840577	141.0662959	2.3641E-08
Residual	13	0.012241957	0.000941689		
Total	14	0.145082533			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	43.77451619	3.675136341	11.91099108	2.28485E-08	35.83486683	51.71416555	35.83486683	51.71416555
YEAR	-0.021781429	0.001833896	-11.87713332	2.3641E-08	-0.02574332	-0.017819537	-0.02574332	-0.017819537

Finally, there is the regression analysis conducted on reports in Zone D, the superior performers, displayed in Table 6. In this test, Pearson's *r* was a .96, a very strong direct correlation leaving little room for exterior influences. However, like Zone A and Zone B, Zone D exhibited a downward trend, as illustrated by its coefficient, -.0218. Automatically, this cannot be used to support a rejection of the null hypothesis, as the diminishing pool of superior performers cannot be construed as an improvement in workforce quality. However, with P (2.3641×10^{-8}) < α, this was a significant finding.

Table 7. Regression Analysis for OCS Scores

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.352512175
R Square	0.124264834
Adjusted R Square	0.05690059
Standard Error	0.060721245
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.006801429	0.006801429	1.844670515	0.197509149
Residual	13	0.047931905	0.00368707		
Total	14	0.054733333			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-6.490190476	7.272108979	-0.892477065	0.388359247	-22.20062678	9.220245828	-22.20062678	9.220245828
YEAR	0.004928571	0.003628788	1.35818648	0.197509149	-0.002910949	0.012768092	-0.002910949	0.012768092

Even though the regression analyses of the four zones were finished, there was still more to examine. The next regression conducted was on the overall average evaluation scores (known as OCS scores). Any change in workforce quality should be reflected in the scores, regardless what direction the former trends. Table 7 above lists the results of this regression analysis. Of particular interest is the Pearson's *r* of .35. This indicated that influences outside those intended to be a part of this evaluation system have more of an influence on OCS scores than on the distribution of reports. However, with a coefficient of .0049, OCS scores did exhibit an upward trend, but, with a P-value (.1975) larger than α, this finding was not significant. Therefore, this analysis cannot be used to support a rejection of the null hypothesis.

Table 8. Correlation Analysis Between Report Distribution and Average OCS Scores

	A	*B*	*C*	*D*	*AVERAGE OCS SCORE*
A	1				
B	0.79960503	1			
C	-0.7750468	-0.9176445	1		
D	0.70574904	0.85284874	-0.9891247	1	
AVERAGE OCS SCORE	-0.6519535	-0.8194655	0.60312645	-0.5011851	1

In the last test solely dedicated to determining acceptance or rejection of $H_0(1)$, correlation analysis for Pearson's *r* was conducted to discover how much correlation exists between the percentage of reports in each Zone and the average OCS score. Not surprisingly, the only zone with a direct correlation, albeit not a particularly strong one (r = .60) was the only zone with an upward trend, Zone C. Zone A (r = -.65), Zone B (r = -.82), and Zone D (r = -.50), with their downward trends, each have an inverse correlation with average OCS scores. With Zone C having enough of a portion of the reports to move average OCS scores upward despite the other three zones pulling them downward, this test can be used to support a rejection of the null hypothesis.

For this hypothesis, five out of the seven tests conducted supported a rejection of the null hypothesis, each with significant results; therefore, the null hypothesis was rejected, and a statement can be made that employee performance improved significantly from 1997-2011.

Hypothesis 2. The next hypothesis, $H_0(2)$, that there is no strong correlation between promotions and evaluations meeting or exceeding standards, required only two regression analyses and a correlation analysis for Pearson's *r*.

Table 9. Regression Analysis for Satisfactory Evaluations

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.622129054
R Square	0.38704456
Adjusted R Square	0.339894142
Standard Error	0.007090475
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.000412692	0.000412692	8.208719525	0.013267569
Residual	13	0.000653573	5.02748E-05		
Total	14	0.001066265			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-1.442042148	0.849170767	-1.698176862	0.113266337	-3.276564057	0.39247976	-3.276564057	0.39247976
YEAR	0.001214043	0.000423737	2.865086303	0.013267569	0.000298615	0.002129471	0.000298615	0.002129471

The first regression, recorded in Table 9, was conducted on the percentage of performance reports that met or exceeded standards, in other words, those in Zone B, Zone C, and Zone D. Again, since Zone A reports reflected substandard performance, they would not lead to promotions and were therefore omitted. Since the regression analysis in Table 2 already revealed a downward trend in Zone A reports, it was no surprise that this analysis of the other three zones showed an upward trend, as demonstrated by the coefficient (.0012). With a slight upward trend in satisfactory reports, a corresponding upward trend in promotions would not seem unreasonable. However, in Table 10, the analysis trended opposite to what was expected.

Table 10. Regression Analysis for Promotion Rates

Regression Statistics	
Multiple R	0.332630505
R Square	0.110643053
Adjusted R Square	0.04223098
Standard Error	0.025674721
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.001066112	0.001066112	1.617303034	0.225748942
Residual	13	0.008569487	0.000659191		
Total	14	0.009635599			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	3.981838298	3.074860634	1.294965455	0.217858922	-2.660994238	10.62467083	-2.660994238	10.62467083
YEAR	-0.001951293	0.001534358	-1.271732296	0.225748942	-0.005266072	0.001363486	-0.005266072	0.001363486

This second regression showed a downward trend in the promotion rate, with a coefficient of -.0020. At this juncture, one could already wonder how much the evaluation scores tie into promotions. There are limits to promotions, as each position in the Laboratory is assigned a particular grade, and higher grades occupy fewer positions than do lower grades. However, if the number of promotions is limited due to prior promotions, suspicions of politics could make their way through the workforce and force engagement and motivation levels downward.

Table 11. Correlation Analysis Between Satisfactory Evaluations and Promotion Rates

	% SATISFACTORY	*PROMOTION RATE*
% SATISFACTORY	1	
PROMOTION RATE	0.381892446	1

With trends headed in opposite directions, as these two are, there is no wonder to the result of the correlation analysis for Pearson's *r*, with *r* being .38, as shown in Table 11. With this result, it confirms that either the maximum number of promotions was close to being met or that promotions had to do with more than just superlative evaluations. At any rate, these results lead to accepting the null hypothesis in this case, that there wqs no significant correlation between satisfactory evaluations and promotion rates.

Research Question 2

For this research question, the aim was to ascertain whether the data revealed a trend of substandard employees being separated at an increased rate, and whether it reveals a trend of outstanding contributors being retained at an increased rate. If either one of these trends held true, the workforce quality would increase; to have both trends in existence raise the quality even more. For these tests, $\alpha = .05$.

Attention is first turned to the substandard contributors. Even though one substandard report is not enough to fire an employee, it does result in that employee being put on a performance improvement plan, and, if sufficient progress is not attained, that person is removed from the workforce. Even though it is possible to remain on staff with one Zone A report and not be separated for 2 or more years, Zone A still experienced the highest separation rates of the four zones.

Table 12. Correlation Analysis Between Zone A Reports and Separations

	ZONE A	*SEPARATIONS*
ZONE A	1	
SEPARATIONS	0.906820158	1

Hypothesis 3. The first test in determining acceptance or rejection of $H_0(3)$, that substandard employees are not being separated at a significantly higher rate was a correlation analysis for Pearson's *r*, to determine how much correlation existed between the number of employees whose evaluations were in Zone A and the number of those employees who left AFRL. As demonstrated in Table 12, the correlation was very strong ($r = .91$). As stated before, people whose evaluations landed in Zone A experienced the highest separation rate of all four zones, so this is not a surprise. The surprise, however, was the continual occurrence of Zone A evaluations, even with a high rate of separation. Even though percentage of reports in this zone trended downward, and the separation rate trended upward, as seen in Table 13, there were still people in it. So, either many of the new hires into AFRL fall into Zone A, or people who improved Zone B or Zone C were at least partially replaced by people who fell out of those zones.

Table 13. Regression Analysis for Zone A Separation Rates

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.152127766
R Square	0.023142857
Adjusted R Square	-0.052
Standard Error	0.135683455
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.00567	0.00567	0.307984791	0.588341482
Residual	13	0.23933	0.01841		
Total	14	0.245			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-8.628	16.24974705	-0.530962111	0.6043983	-43.73344421	26.47744421	-43.73344421	26.47744421
YEAR	0.0045	0.008108637	0.554963774	0.588341482	-0.013017646	0.022017646	-0.013017646	0.022017646

The next test conducted was a regression analysis on the separation rates for Zone A personnel. As Table 13 shows, the positive coefficient (.0045) reflected an upward trend; however, since the P-value (.5883) is higher than α, it must accepted that the null hypothesis that substandard contributors were not separated at a significantly higher rate. Even though the rate was higher, it was not significantly so. This was either a result of employees being given a chance to improve, or by them actually improving to a more satisfactory level.

Table 14. Correlation Analysis Between Zone D Reports and Retention

	ZONE D	*RETENTION*
ZONE D	1	
RETENTION	0.999737446	1

Hypothesis 4. For the test on $H_0(4)$, that outstanding contributors were not retained at a significantly higher rate, the process used to test the previous hypothesis was duplicated. The correlation analysis for Pearson's *r* revealed correlation of almost 1.00 (.9997) between outstanding reports and retention of those employees, as Table 14 shows. However, the fact that it is not exactly 1.00 indicates that a small percentage of Zone D personnel were leaving AFRL. This calls into question their level of engagement, or, more probably, their motivation. Higher-performing scientists and engineers tend to be more well-known in the research community, and headhunters from both the private and public sectors take notice. Corporate recruiters can and often do offer compensation packages that governments cannot compete with. If that is what motivates an employee to leave government service, what does that say about his or her level of engagement? Government recruiters may not offer the same pay and/or benefits, but the prestige of working on a more important or a more publicized project may be enough to convince a researcher to leave one public sector position for another. Such an employee may be described as engaged as far as working for the government, but it appears that it is more of a motivation, and that such an employee is not as committed to his actual work center or colleagues.

Table 15. Regression Analysis for Zone D Retention Rates

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.648234147
R Square	0.42020751
Adjusted R Square	0.375608088
Standard Error	0.024531508
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.00567	0.00567	9.421815083	0.008957509
Residual	13	0.007823333	0.000601795		
Total	14	0.013493333			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	9.958666667	2.937946967	3.389668628	0.004837408	3.611618124	16.30571521	3.611618124	16.30571521
YEAR	-0.0045	0.001466038	-3.069497529	0.008957509	-0.007667183	-0.001332817	-0.007667183	-0.001332817

Again, as in the testing of the previous hypothesis, there was a "however." The regression analysis in Table 15 showed a negative coefficient (-.0045), indicating a downward trend in the retention of outstanding contributors. Even though the P-value (.0090) < α, the null hypothesis must be accepted because the retention rate was falling.

Research Question 3

The focus for this research question was whether AFRL's Lab Demo employees exhibited satisfaction in the form of retention rates and survey responses. Attention will first be turned to retention. Again, as in the previous two research questions, $\alpha = .05$.

Hypothesis 5. To analyze $H_0(5)$, that there is a significant difference between the separation rates of Lab Demo employees and the federal government as a whole, a *t* test was conducted to calculate the average separation rate of the two labor pools over a 12-year span.

Table 16. *t* Test Comparing Separation Rates

t Test: Paired Two Sample for Means

	LAB DEMO	*FEDERAL*
Mean	0.061666667	0.074666667
Variance	0.000415152	0.000146606
Observations	12	12
Pearson Correlation	-0.074926702	
Hypothesized Mean Difference	0	
df	11	
t Stat	-1.840430639	
P(T<=t) one-tail	0.046408885	
t Critical one-tail	1.795884819	
P(T<=t) two-tail	0.092817769	
t Critical two-tail	2.20098516	

It was seen that while Lab Demo employees average over a one and one-quarter percent lower separation point over the span of observations, the P-value of .0928 caused a rejection of the null hypothesis, and state that while Lab Demo employees did have a higher rate of retention, it was not statistically significantly so.

Hypothesis 6. Now, employee satisfaction will be addressed, starting with the non-supervisors. In 2010, an employee satisfaction survey was conducted and the results were broken between Lab Demo employees and General Schedule employees. At this time, the Lab Demo had a 13-year window of opportunity for changes to be made to the system, and for disgruntled employees to leave. For this hypothesis, non-supervisors were compared to non-supervisors, and supervisors to supervisors, because each had the same questions asked of them. Appendix C lists the questions asked of non-supervisors, and Appendix D lists those of supervisors. A *t* test was conducted on the rate of satisfied responses given for each question, both for Lab Demo and General Schedule non-supervisors. Table 17 displays the results.

Table 17. *t* Test for Satisfied Response Rates, Non-supervisors

t Test: Paired Two Sample for Means

	2010 LAB DEMO NON-SUPERVISOR	*2010 GS NON-SUPERVISOR*
Mean	0.539411765	0.606470588
Variance	0.050980882	0.039724265
Observations	17	17
Pearson Correlation	0.873107038	
Hypothesized Mean Difference	0	
df	16	
t Stat	-2.511262995	
P(T<=t) one-tail	0.011571755	
t Critical one-tail	1.745883676	
P(T<=t) two-tail	0.023143511	
t Critical two-tail	2.119905299	

The questions asked of the non-supervisors pertained to how the employees felt about their jobs, their supervisors, their compensation, certain HR-related processes, and their respective units. The *t* test results in Table 17 showed that General Schedule non-supervisors averaged almost seven percentage points higher in satisfied answers than their Lab Demo counterparts, by itself enough to accept the null hypothesis, $H_0(6)$, that Lab Demo non-supervisors do not exhibit significantly higher job satisfaction than General Schedule non-supervisors.

Hypothesis 7. The same process was conducted for supervisory personnel. In addition to the questions asked of the non-supervisors, the supervisors were asked questions concerning their supervisory roles in their respective systems. A *t* test was conducted on the satisfied responses of all questions, the results of which are shown in Table 18.

Table 18 *t* Test for Satisfied Response Rates, Supervisors

t Test: Paired Two Sample for Means

	2010 LAB DEMO SUPERVISOR	*2010 GS SUPERVISOR*
Mean	0.596764706	0.546470588
Variance	0.045677094	0.072278075
Observations	34	34
Pearson Correlation	0.70560396	
Hypothesized Mean Difference	0	
df	33	
t Stat	1.527293111	
P(T<=t) one-tail	0.068108589	
t Critical one-tail	1.692360309	
P(T<=t) two-tail	0.136217177	
t Critical two-tail	2.034515297	

Even though Lab Demo supervisors averaged just over a 5% higher satisfaction rate than General Schedule supervisors, the P-score of .1362 > α, so it was not significantly higher. Therefore, $H_0(7)$ must be accepted, and it must be stated that Lab Demo supervisors do not exhibit significantly higher job satisfaction than their General Schedule counterparts.

The results of the tests showed that Lab Demo has been successful in some areas, and less than successful in others. In its attempt to attract and maintain an engaged (or at least motivated) workforce, AFRL has managed to improve workforce performance, but it still has some work to do in terms of rewarding that improved performance, retaining the employees that have raised their performance level, and generating satisfaction among the employees that do elect to stay on board. All of these will be discussed in the next chapter.

Chapter V

DISCUSSION

Since 1949, the federal government has, in one form or another, utilized the General Schedule system. The extended usage of this system has resulted in inflated evaluation ratings which all but guarantee the regularly-scheduled within-grade step promotions and the resulting pay raises, whether or not an employee demonstrated a corresponding lift in performance or production. Knowing that these promotions were almost certain, only those employees endowed with high levels of public service motivation and/or employee engagement would perform at the level normally required for such rewards, and they might only do so for a limited time, after seeing their lower-performance colleagues being given the same benefits with much less effort. Since the mission of AFRL is so critical to the continued success of the men and women who fight America's wars, this is not the type of employee needed to ensure that our warfighters can perform their missions and come home safely.

Lab Demo represents the Air Force's contribution to Acq Demo, an attempt by the Clinton/Gore administration to ameliorate human resources management within the DOD's acquisition community, especially in terms of talent and performance management. Efforts to analyze the effectiveness of AFRL's share have included data from the annual reports from 1997-2010 and the 2010 employee satisfaction survey. Data from the Bureau of Labor Statistics and OPM were analyzed to determine whether

or not the desired changes in employee performance, retention rates, and employee satisfaction were realized, and whether or not they were significant enough to justify the increased cost to the American taxpayer. Some of the test results reflected the significant changes desired by Lab Demo proponents; however, others reflected either results too insignificant to justify the time and money poured into the project, or results completely opposite from what was desired. Nevertheless, this research project will not result in a call to completely scrap this system and go back to the old way and risk undoing whatever good has come about. To do so would betray our men and women in uniform, and would also fail to meet at least some of Gilbert's foundational concepts: 1) responsiveness (prompt acquiescence by the government for policy change demands); 2) flexibility (making an administering policy as to account for individual, group, or situational differenced or local concerns); 3) leadership (the demand that government take the initiative rather than merely respond); 4) candor (policy should be open to public scrutiny); competence (policy should be made and administered according to recognized standards or disciplinary standards if available); 5) efficacy (government performance should be efficient, timely, thorough, and conservatively priced); and 6) accountability (Gilbert, 1959, pp. 375-378).

What Happened at AFRL?

This study was initiated with the intent of statistically indicating whether Lab Demo has had a positive effect on employee engagement at AFRL. Employee engagement was addressed by asking three research questions concerning employee performance, employee retention, and employee satisfaction.

The first issue addressed was whether AFRL witnessed significantly higher employee performance because of the Lab Demo project. The question was approached from two angles, employee evaluations and promotions.

Employee evaluations showed promise. Table 3 displayed a decreasing trend in Zone A (unsatisfactory) reports (regression coefficient = -.0012) and Table 4 did likewise with Zone B (low satisfactory) reports (regression coefficient = -.0049). However, Table 6 also showed a downward trend in Zone D (outstanding) reports, with a regression coefficient of -.0218. The only zone to trend upward was Zone C (high satisfactory reports) in Table 5, with a regression coefficient of .0279. Finally, the regression analysis conducted on OCS scores in Table 7 also showed an increasing trend (regression coefficient = .0049). So, considering Rich, Lepine, and Crawford's conclusion that engagement at least partially determines performance (Rich, Lepine, & Crawford, 2010, p. 625), we could tentatively agree with that finding.

However, other data in the findings dampens that agreement. The correlation analysis between the total reports and zone distribution in Table 2 resulted in values for Pearson's *r* ranging from -0.90 to .79, indicating a lack of consistency in how raters evaluated their subordinates from year to year. Also, the only positive correlations in Table 2 are those involving Zone C, indicative of the fact that Zone C was the only zone to show an increase in its share of reports. Furthermore, the regressions for Zone A and Zone B reports yielded *r* values of .62 and .72, respectively, and they dwarf the Pearson's *r* for the egression on OCS scores, .35, leaving much room for outside factors to enter the evaluation process. Finally, the only zone to show a positive correlation between it and overall average OCS scores was Zone C ($r = .60$) as shown in Table 8, alerting us to the

very real possibility that the increase in average OCS scores was due to the outflow of Zone A and Zone B reports to Zone C. Put together, these issues show that any of Grote's nine rater errors could have made its way into the process, especially central tendency.

Turning now to the promotion viewpoint, the statistical analysis was approached by comparing the rate of satisfactory evaluations with the rate of promotions. Table 9, a regression analysis of satisfactory reports, trends upwards (regression coefficient = .0012), which should come as no surprise as unsatisfactory reports, which showed a downward trend to begin with, were filtered out. On the other hand, Table 10, a regression analysis of promotion rates, shows a downward trend of promotion rates (regression coefficient = .0020). and Table 11, a correlation analysis between satisfactory evaluations and promotions shows a weak positive correlation between the two (r = .38). This data shows that the opportunities for advancement have lessened over the years of Lab Demo, and could rightly call into question the fairness of the promotion process, in line with Rubin's research. With opportunities on the decline, those who do get selected for promotion will be severely derided by those who do not, due to a dim view of the process. Thus, an increasing number of employees will experience a decreased level of engagement.

In 2010, an employee satisfaction survey was administered in AFRL, and one of the questions was "satisfied with opportunities for advancement." The percentage of respondents that answered with "strongly agree" or "agree" was broken down into four categories, each with shockingly low numbers: Lab Demo supervisors (66%), Lab Demo non-supervisors (36%), General Schedule supervisors (50%), and General Schedule non-

supervisors (35%). These results resemble the results of Rubin's survey, and indicate a low perception of the fairness in the promotion process, which, at least for this question, leads to low job satisfaction, which, in turn, leads to lower employee engagement and lower employee performance.

The second question centered on employee retention, as engaged employees rend to remain with their employers, and examined whether Zone A and Zone D employees were separated or retained at significantly higher rates. The Federal Register stated that Lab Demo contributed to increased separation of poor contributors, and both the correlation analysis between Zone A reports and separations in Table 11 (r = .91) and the regression analysis in Table 13 (coefficient = .0045) bear this out. However, without knowing which of Anitha's seven factors affect each individual employee and how to handle each situation, it will be difficult at best to eliminate the distribution of reports in Zone A. To better cope with these employees, to either minimize their number in the first place or to better help them transition upward, it may be best to understand why they end up in the bottom tier in the first place.

It may be easy to understand why substandard employees leave an organization, because many of them do so involuntarily; however, this is not the case with outstanding contributors. The correlation analysis between Zone D reports and retention in Table 14 shows an almost direct correlation (r = .9997) in retention of outstanding employees in AFRL; however, the regression analysis in Table 15 shows a downward trend (coefficient = -.0045) in their retention. More confusing than the departure of poor contributors is why outstanding contributors leave organizations they have enjoyed success in. If they do so because they do not feel they are adequately compensated, it may be a question of

procedural justice perceptions, referring again to Rubin's study. Another argument along the same line is the decreasing number of evaluations in Zone D, as more reports end up in Zone C. However, a more apparent reason might be the lack of satisfaction with pay, especially when private sector recruiters come calling, as the grass is greener on the private sector side of the fence.

Table 16, a *t* test comparing the separations of Lab Demo employees and federal employees in total, shows that, Lab Demo employees left their positions at a slower rate than the federal workforce as a whole, (6.2% annually as opposed to 7.5% annually), albeit not significantly so. Does this mean that the Lab Demo work force was more satisfied overall with their employment situation than the rest of the federal workers? A search for a legitimate answer to this question would be cause for satisfaction and exit surveys, which will be addressed later. Attention to public service motivation early in employees' tenures would be the first step in addressing this issue, with Perry, Hondeghem, and Wise's three propositions being particularly important. "The greater an individual's public service motivation, the more likely the individual will seek membership in a public organization" (Perry, Hondeghem, & Wise, 2010, p. 683) suggests that those who seek employment with AFRL have at least some desire to work in the public sector. "In public organizations, PSM is positively related to individual performance" (Perry, Hondeghem, & Wise, 2010, p. 684) theorizes that those with higher levels of PSM will perform at higher levels while fulfilling the AFRL mission. "Public organizations that attract members with high levels of public service motivation are likely to be less dependent on utilitarian incentives to manage individual performance effectively" (Perry, Hondeghem, & Wise, 2010, p. 686) would, however, seem to

contradict the applicability of the first two propositions to AFRL. The bonuses given to those employees in Zones B, C, and D figure quite prominently in compensation packages, a major recruiting tool for any employer.

The last research question compared the satisfaction between Lab Demo and General Schedule workforce members. Table 17, a *t* test comparing satisfaction rates between Lab Demo and General Schedule non-supervisors, and Table 18, another *t* test doing the same for supervisors, explore whether Lab Demo had increased employee satisfaction among its workforce over that of the AFRL GS employees in the first 13 years of the program. Among non-supervisory employees, Lab Demo had not succeeded in doing so, mustering a 54% satisfaction rate compared to 61% for General Schedule; among supervisors, the opposite held true. Lab Demo satisfaction was 60% as opposed to 55% for General Schedule. Considering the 10 factors set forth by Ghosh, Satayawadi, Joshi, and Shadman, they appear to affect the four different segments of the AFRL population differently. Instead of developing blanket solutions for problems, the better solution looks to be solving these problems by addressing the different groups individually.

Why Does All This Matter?

Lab Demo represents an attempt to improve human resources management in AFRL, and in some ways it has demonstrated itself to be an improvement over the time-honored General Schedule system. However, as indicated by the acceptance of some of the null hypotheses in some of our statistical analyses, and the lack of significance in the findings of others, it is not without its flaws, perhaps the most glaring of which is its subjectivity.

It may never be possible to develop an evaluation system completely devoid of subjectivity; however, limiting it by measuring employees against job standards instead of measuring them against each other, workers will be able to show their performances stack up against those standards and sense that their rating was not some random guess.

The other big weakness of the Lab Demo system is the employees whose evaluations fall into Zone A. While not a part of the system design, these employees represent potential in danger of being wasted should they leave AFRL. Dealing with them would take a two-pronged approach, one centered on prevention, the other on rehabilitation.

The prevention initiative would center on employee satisfaction surveys, to be given at least annually, before midterm feedbacks. These surveys would need to be carefully constructed so as to address those areas mentioned earlier that predict employee satisfaction, and therefore employee engagement. These surveys should be reviewed by somebody outside the employee's rating chain, and any responses that are cause for concern would be relayed to the supervisory chain and addressed during the midterm feedback session.

The rehabilitation aspect, first and foremost, depends on the employee's willingness to go through that process. If not, then he or she could be shown the door. If so, then a questionnaire should be given to the employee, with questions addressing potential reasons that could lead to poor employee performance. The responses to the questionnaire then would form the basis of the resulting performance improvement plan. Successful completion of the plan would lead to the opportunity for continued employment at AFRL.

What's Needed?

First and foremost, the AFRL workforce must be reminded that it is a military organization, and accompanying that fact are certain expectations, such as the mindset that the civilians in AFRL are not only researchers but also public servants. However, those charged with leading and managing AFRL also need to be reminded that it is a component of the DOD, and that civilian management techniques are not always suitable for the military.

As stated before, the small amount of data is a weakness in this study. To continue this study with the same analyses could only be accomplished if AFRL were to publish the results of all pay bands in its annual reports, which has not been done since 2011. Another possibility would be to continue to analyze the reports on the two pay bands still reported annually, DR and DO.

Future efforts should also look at employee satisfaction surveys, a tool that has not been utilized often. One was conducted in 1996, as Lab Demo was starting up; another was conducted in 2005, and the last one in 2010. These surveys are invaluable tools for gauging how the workforce views its situation, and, given the rate of turnover and changes within the organization, should be conducted more often than has happened.

Another tool that needs to be more heavily utilized is the separation survey. Employees do leave AFRL, for various reasons. These surveys could be analyzed in an attempt to find trends that AFRL has the authority to change in an attempt to eliminate the common causes for employee separation.

Should the situation not improve, or if OPM does not grant permission to permanently implement Lab Demo as the personnel management system for AFRL,

serious consideration needs to be given for dropping the Project altogether, and converting all employees to DOD's new personnel management system, New Beginnings. It would provide consistency across more of the federal workforce and save the taxpayers the money that went towards bonuses under Lab Demo.

As of the writing of this paper, a presidential race has resulted in the election of someone who has announced his attention to freeze federal hiring and reduce the federal workforce by attrition. AFRL has experienced hiring freezes in the past, and has managed to find a way to hire new employees nonetheless. However, each freeze is different, and should the next one be absolute, it will be imperative for AFRL to retain as many of its experienced scientific experts as possible. Will the workforce be sufficiently engaged, or at least motivated, to stay on board, or will the lure of higher-profile, higher-paying jobs be enough to lure researchers into private sector pastures?

AFRL, like the rest of the DOD, is a part of the U.S. government. With all of the controversies occurring in government, there have calls for change. However, that change will not occur until the people in the government are changed, and that will not happen until the government is transformed into a place that people actually want to work in.

REFERENCES

Air Force Research Laboratory. (2008, July 1). *Air Force Research Laboratory Manual 36-104, Laboratory Personnel Demonstration Project*, 1-123.

Air Force Research Laboratory. (2011, September 1). Air Force Research Laboratory Manual 36-104, Laboratory Personnel Demonstration Project. 1-141.

Anitha, J. (2014). Determinants of Employee Engagement and Their Impact on Employee Performance. *International Journal of Productivity and Performance Management, 63*(3), 308-323.

Bakker, A. B. (2015, September/October). A Job Demands-Resources Approach to Public Service Motivation. (H. G. Rainey, Ed.) *Public Administration Review, 75*(5), 723-732.

Behn, R. D. (1995, July/August). The Big Questions of Public Management. *Public Administration Review, 55*(4), 313-324.

Belle, N. (2015, March/April). Performance-Related Pay and the Crowding Out of Motivation in the Public Sector: A Randomized Field Experiment. *Public Administration Review, 75*(2), 230-241.

Belle, N., & Cantarelli, P. (2015, June). Monetary Incentives, Motivation, and Job Effort in the Public Sector: An Experimental Study with Italian Government Executives. *Review of Public Personnel Administration, 35*(2), 99-123.

Bozeman, B., & Su, X. (2015, September/October). Public Service Motivation Concepts and Theory: A Critique. *Public Administration Review, 75*(5), 700-710.

Campbell, D. J., Campbell, K. M., & Chia, H.-B. (1998). Merit Pay, Performance Appraisal, and Individual Motivation: an Analysis and Alternative. *Human Resource Management, 37*(2), 131-146.

Ghosh, P., Sataywadi, R., Joshi, J. P., & Shadman, M. (2013). Who Stays With You? Factors Predicting Employees' Intention to Stay. *International Journal of Organizational Analysis, 21*(3), 288-312.

Gilbert, C. E. (1959, August). The Foundations of Administrative Responsibility. *The Journal of Politics, 21*(3), 373-407.

Grote, R. C. (1996). *The Complete Guide to Performance Appraisal.* New York, NY: American Management Association.

Homberg, F., McCarthy, D., & Tabvuma, V. (2015, September/October). A Meta-Analysis of the Relationship between Public Service Motivation and Job Satisfaction. (M. McGuire, Ed.) *Public Administration Review, 75*(5), 711-722.

Ingraham, P. W. (1993). Of Pigs in Pokes and Policy Diffusion: Another Look at Pay-for-Performance. *Public Administration Review, 53*(4), 348-356.

Jefferies, C. L. (1977). Public Administration and the Military. *Public Admnistration Review, 37*(4), 321-333.

Kleiman, L. S., Biderman, M. D., & Faley, R. H. (1987, Winter). An Examination of Employee Perceptions of a Subjective Performance Appraisal System. *Journal of Business and Psychology, 2*(2), 112-121.

Lane, J.-E. (2003). *New Public Management.* New York: Routledge.

Losey, S. (2008, August 11). *Is DOD's New Pay System Fair?* Retrieved July 16, 2015, from Federal Times:

http://archive.federaltimes.com/article/20080811/DEPARTMENTS01/111122001/1023/DEPARTMENTS01

Macey, W. H., & Schneider, B. (2008, 2008). The Meaning of Employee Engagement. *Industrial and Organizational Psychology, 1*, 3-30.

Marciano, P. L. (2010). *Carrots and Sticks Don't Work: Build a Culture of Employee Engagements with the Principles of RESPECT.* New York: McGraw-Hill.

Mostafa, A. M., Gould-Williams, J. S., & Bottomley, P. (2015, September/October). High-Performance Human Resource Practices and Employee Outcomes: The Mediating Role of Public Service Motivation. *Public Administration Review, 75*(5), 747-757.

Moynihan, D. P., & Pandey, S. K. (2010, October). The Big Question for Performance Management: Why Do Managers Use Performance Information? *Journal of Public Administration Research and Theory, 20*(4), 849-866.

National Archives and Records Administration. (2010, August 30). Federal Register: Part III DOD-Science and Technology Reinvention Laboratory Personnel Management Demonstration Project, Department of the Air Force, Air Force Research Laboratory (AFRL); Notice. 53076-53126. Washington, D.C.: GPO.

National Archives and Records Administration. (2013, May 20). *The Declaration of Independence.* Retrieved July 15, 2015, from http://www.archives.gov/exhibits/charters/declaration_transcript.html

National Archives and Records Administration. (n.d.). *Pendleton Act (1883)*. Retrieved May 23, 2016, from We the People: https://www.ourdocuments.gov/doc.php?flash = true&doc = 48

Nigro, L. G., & Kellough, J. E. (2008, December). Personnel Reform in the States: A Look at Progress Fifteen Years After the Winter Commission. *Public Administration Review, 68*(Supp 1), S50-S57.

Office of Personnel Management. (2005, July 2005). *OPM to Stand with DOD During Implementation of New Pay System and Personnel Rules.* Retrieved October 22, 2015, from Latest News: https://www.opm.gov/news/releases/2005/07/opm-to-stand-with-DOD-during-implementation-of-new-pay-system-and-personnel-rules/

Office of Personnel Management. (2014, December 19). *Pay and Leave Systems*. Retrieved July 16, 2015, from https://www.opm.gov/policy-data-oversight/pay-leave/pay-systems/general-schedule/

Office of Personnel Management. (n.d.). *2010 Employee Survey Results U.S. Office of Personnel Management (OPM).* Retrieved August 2016, from Data, Analysis & Documentation: Employee Surveys: https://www.opm.gov/policy-data-oversight/data-analysis-documentation/employee-surveys/results/2010-employee-survey-results/

Office of Personnel Management. (n.d.). *Broadbanding*. Retrieved October 28, 2015, from Strategic Compensation Policy Center: http://apps.opm.gov/strat/bandtext.cfm

Office of Personnel Management. (n.d.). *Pay Administration*. Retrieved August 17, 2016, from Pay & Leave: https://www.opm.gov/policy-data-oversight/pay-leave/pay-administration/fact-sheets/within-grade-increases/

Office of Personnel Management. (n.d.). *Performance Management Overview & History-Overview*. Retrieved May 23, 2016, from OPM.gov:

https://www.opm.gov/policy-data-oversight/performance-management/overview-history/#url = 5

Office of Personnel Management. (n.d.). *Performance Managment Overview & History-Historical Chronology*. Retrieved May 23, 2016, from OPM.gov: https://www.opm.gov/policy-data-oversight/performance-management/overview-history/#url = Historical-Chronology

O'Keefe, E., & Yoder, E. (2011, May 26). *No More Raises for Poor-Performing Federal Workers?* Retrieved July 20, 2015, from http://www.washingtonpost.com/blogs/federal-eye/post/no-more-raises-for-poor-performing-federal-workers/2011/05/25/AG2awsBH_blog.html

Pederson, M. J. (2015, September/October). Activating the Forces of Public Service Motivation: Evidence from a Low-Intensity Randomized Survey Experiment. *Public Administration Review, 75*(5), 734-746.

Perry, J. L., & Vandenabeele, W. (2015, September/October). Public Service Motivation Research: Achievements, Challenges, and Future Directions. *Public Administration Review, 75*(5), 692-699.

Perry, J. L., & Wise, L. R. (1990, May-June). The Motivational Bases of Public Service. *Public Administration Review, 50*(3), 367-373.

Perry, J. L., Hondeghem, A., & Wise, L. L. (2010). Revisiting the Motivational Bases of Public Service: Twenty Years of Research and an Agenda for the Future. *Public Administration Review, 70*(5), 681-690.

Rich, B. L., Lepine, J. A., & Crawford, E. R. (2010, June). Job Engagement: Antecedents and Effects on Job Performance. *The Academy of Management Journal, 53*(3), 617-635.

Rubin, E. V. (2009, January). The Role of Procedural Justice in Public Personnel Management: Empirical Results from the DOD. *Journal of Public Administration Research and Theory, 19*(1), 125-143.

Saks, A. M. (2008). The Meaning and Bleeding of Employee Engagement: How Muddy Is the Water? *Industrial and Organizational Psychology, 1*, 40-43.

Saks, A. M., & Gruman, J. A. (2014, Summer). What do we Really Know About Employee Engagement. *Human Resource Development Quarterly, 25*(2), 155-182.

Sanger, M. B. (2008, June). Getting to the Roots of Change: Performance Management and Organizational Culture. *Public Performance & Managment Review, 31*(4), 621-653.

Stanislavov, I., & Ivanov, S. (2014, 62). The Role of Leadership for Shaping Organizational Culture and Building Employee Engagement in the Bulgarian Gaming Industry. *Tourism*, 19-40.

Stazyk, E. C. (2013, September). Crowding Out Public Service Motivtion? Comparing Theoretical Expectations with Empirical Findings on the Influence of Performance-Related Pay. *Review of Public Personnel Administration, 33*(3), 252-274.

U.S. Air Force. (2014, December 15). *Air Force Reserach Laboratory*. Retrieved July 9, 2015, from

http://www.af.mil/AboutUs/FactSheets/Display/tabid/224/Article/104463/air-force-research-laboratory.aspx

U.S. Congress. (1996). Public Law 104-106 Feb 10, 1996-National Defense Authorization Act for Fiscal Year 1996. 1-517. Washington, D.C.: U.S. Government Printing Office. Retrieved July 16, 2015, from http://thomas.loc.gov/cgi-bin/query/F?c104:4:./temp/~c104f1GEov:e1664387:

U.S. Government Accountability Office. (2011, April 28). *PERFORMANCE MANAGEMENT: DOD Is Terminating the National Security Personnel System, but Needs a Strategic Plan to Guide Its Design of a New System.* Retrieved October 22, 2015, from http://www.gao.gov/assets/100/97461.pdf

U.S. Government Printing Office. (2012, January 12). *5 U.S.C. 3331 - Oath of Office.* Retrieved July 19, 2015, from http://www.gpo.gov/fdsys/granule/USCODE-2011-title5/USCODE-2011-title5-partIII-subpartB-chap33-subchapII-sec3331/content-detail.html

van der Heijden, B. I., & Nijhof, A. H. (2004, May). The Value of Subjectivity: Problems and Prospects for 360-Degree Appraisal Systems. *International Journal of Human Resource Management, 15*(3), 493-511.

Wright, B. E. (2007, January-February). Public Service and Motivation: Does Mission Matter? *Public Administration review, 67*(1), 54-64.

Wright-Patterson Air Force Base. (2015, June 29). *Air Force Research Laboratory*. Retrieved July 9, 2015, from http://www.wpafb.af.mil/AFRL/

Yang, K., & Kassekert, A. (2010, April). Linking Management Reform with Employee Job Satisfaction: Evidence from Federal Agencies. *Journal of Public Administration Research and Theory, 20*(2), 413-436.

APPENDIX A:

General Schedule/Lab Demo Equivalents

Table A1. Lab Demo Grades and Their General Schedule Equivalents

Broadband	DR	DO	DX	DU
I	7-11	7-11	1-4	1-4
II	12-13	12-13	5-7	5-6
III	14	14	8-10	7-8
IV	15	15	11-12	9-10

This table, taken from the AFRL instruction governing Lab Demo, displays the four broadbands and their equivalent General Schedule grade equivalents. For example, a DR-I, the lowest level scientist or engineer, is equivalent to a GS-7 to GS-11. The four pay plans (DR, DO, DX, and DU) will be explained in Appendix B.

APPENDIX B:

Pay Plans and Occupational Series

These four tables display the four pay plans within Lab Demo, and each of the occupational series that fall within them.

Table B1. DR Pay Plan (Scientists and Engineers)

0101	Social Science	0855	Electronics Engineering
0180	Psychology	0858	Bioengineering & Biomedical Engineering
0190	General Anthropology	0861	Aerospace Engineering
0199	Social Science Student Trainee	0893	Chemical Engineering
0401	General Biological Sciences		
0403	Microbiology	0896	Industrial Engineering
0405	Pharmacology	0899	Engineering & Architecture Student Trainee
0413	Physiology		
0414	Entomology	1301	General Physical Science
0415	Toxicology	1306	Health Physics
0601	General Health Science	1310	Physics
0662	Optometrist	1313	Geophysics
0665	Speech Pathology & Audiology	1320	Chemistry
0701	Veterinary Medical Science	1321	Metallurgy
0801	General Engineering	1330	Astronomy & Space Science
0803	Safety Engineering	1340	Meteorology
0804	Fire Protection Engineering	1370	Cartography
0806	Materials Engineering	1399	Physical Science Student Trainee
0808	Architecture	1515	Operations Research
0810	Civil Engineering	1520	Mathematics
0819	Environmental Engineering	1529	Mathematical Statistics
0830	Mechanical Engineering	1530	Statistics
0840	Nuclear Engineering	1550	Computer Science
0850	Electrical Engineering	1599	Mathematics & Statistics Student Trainee
0854	Computer Engineering		

Table B2. DO Pay Plan (Business Management and Professionals)

0018	Safety & Occupational Health Management	1040	Language Specialist
		1060	Photography
0028	Environmental Protection Specialist	1071	Audiovisual Production
0030	Sports Specialist	1082	Writing & Editing
0080	Security Administration	1083	Technical Writing & Editing
		1084	Visual Information
0099	General Student Trainee	1035	Public Affairs
0110	Economist	1101	General Business & Industry
0130	Foreign Affairs	1102	Contracting
0131	International Relations	1150	Industrial Specialist
0132	Intelligence	1152	Production Control
0170	History	1170	Realty
0201	Human Resources Management	1199	Business & Industry Student Trainee
0301	Miscellaneous Administration & Program	1222	Patent Attorney
		1410	Librarian
		1412	Technical Information Services
0340	Program Management	1601	Equipment, Facilities, & Services
0341	Administrative Officer	1640	Facility Operations Services
0343	Management & Program Analysis	1654	Printing Services
0346	Logistics Management	1670	Equipment Services
0391	Telecommunications	1712	Training Instruction
0399	Admin/Office Support Student Trainee	1730	Education Research
0501	Financial Administration & Program	1740	Education Services
0505	Financial Management	1750	Instructional Systems
0510	Accounting	1910	Quality Assurance
0560	Budget Analysis	2001	General Supply
0599	Financial Management Student Trainee	2003	Supply Program Management
0644	Medical Technologist	2010	Inventory Management
0671	Health System Specialist	2130	Traffic Management
0690	Industrial Hygiene	2210	Information Technology Management
0905	General Attorney	2299	Information Technology Student
0950	Paralegal Specialist		
1020	Illustrating		

Table B3. DX Pay Plan (Technicians)

0181	Psychology Aid & Technician	0698	Environmental Health Technician
0404	Biological Science Technician	0802	Engineering Technical
0642	Nuclear Medicine Technician	0809	Construction Control Technical
0645	Medical Technician	0856	Electronics Technical
0647	Diagnostic Radiologic Technologist	1311	Physical Science Technician
0649	Medical Instrument Technician	1521	Mathematics Technician

Table B4. DU Pay Plan (Mission Support)

0083	Police	0681	Dental Assistant
0085	Security Guard	0899	Engineering & Architecture Student Trainee
0086	Security Clerical & Assistant		
0099	General Student Trainee	0963	Legal Instruments Examining
0199	Social Science Student Trainee	0986	Legal Assistant
0303	Miscellaneous Clerk & Assistant	1101	General Business & Industry
0305	Mail & File	1105	Purchasing
0318	Secretary	1106	Procurement Clerical & Technician
0326	Office Automation Clerical & Assistance	1199	Business & Industry Student Trainee
		1399	Physical Science Student Trainee
0335	Computer Clerk & Assistant	1411	Library Technician
0344	Management & Program Clerical & Assistance	1599	Mathematics & Statistics Student Trainee
0399	Administrative & Office Support Student Trainee	1603	Equipment, Facilities, & Services Assistance
0503	Financial Clerical & Technician	1702	Education & Training Technician
0525	Accounting Technician	2001	General Supply
0540	Voucher Examining	2005	Supply Clerical & Technician
0561	Budget Clerical & Assistance	2102	Transportation Clerk & Assistance
0599	Financial Management Student Trainee	2299	Information Technology Student Trainee
0650	Medical Technical Assistant		
0675	Medical Records Technician		
0679	Medical Support Assistance		

APPENDIX C:

Questions Asked of Non-Supervisors

These questions were asked of both Lab Demo and General Schedule non-supervisors during the 2010 Employee Satisfaction Survey.

Table C1 *Questions asked of both Lab Demo and General Schedule Non-Supervisors*

I understand my unit's mission
I understand how my job relates to my organization's mission
My supervisor understands how my job relates to my organization's mission
All in all, I am satisfied with my job
Work gives me a feeling of personal accomplishment
My supervisor provides effective mid-cycle feedback
My supervisor provides effective end-cycle feedback
I have trust and confidence in my supervisor
In this organization, employees are kept well informed on all issues affecting their jobs
Pay raises are based on contribution to the organization's mission
CCS is administered fairly in my pay pool
Satisfied with opportunities for advancement
Satisfied with pay
AFRL is able to attract high-quality candidates
Competition for jobs is fair
Demonstration RIF procedures are well understood
Demonstration RIF procedures are fair

APPENDIX D:

Questions Asked of Supervisors

In addition to the questions listed in Appendix C, these questions were asked of Lab Demo and General Schedule supervisors concerning their supervisory roles.

Table D1. Questions asked of Lab Demo and General Schedule Supervisors

Corrective actions are taken when employee contribution is inadequate
I have enough authority to hire people with right skills when I need them
I have enough authority to remove people from their jobs if they don't contribute adequately
I have enough authority to advance people
I have enough authority to influence my employee's pay
I have enough authority to set the pay of new hires
The knowledge, skills, and abilities (KSA's) of the most recent candidate I hired were a good match for the job
I understand the process used to fill vacancies
The paperwork needed to fill vacancies is processed in a timely manner
Disciplinary actions within my directorate are avoided because of the paperwork required
My directorate passes off marginal and unsatisfactory workers to others or moves them to positions where they can be ignored
I understand the job classification system currently being used for AFRL Lab Demo
I have enough authority to influence demonstration job classification decisions
Demonstration job classification decisions are approved in a timely manner
It is easy to reassign demonstration employees to other positions within the lab
Negotiations with civilian personnel office over demonstration job classification delay

the hiring process

Overall, AFRL does good job of assigning the right people to right job

APPENDIX E:

Institutional Review Board Exemption

Valdosta State University Graduate School

Institutional Review Board Oversight Screening Form for Graduate Student Research

Project Title: The Effects of Lab Demo on Employee Engagement

Name: John J. MacGaffick

Faculty Advisor: Dr James W Peterson

Department: Political Science

E-mail: jjmacgaffick@ valdosta.edu

Telephone: (315)832-0178

Please indicate the academic purpose of the proposed research:

- ☒ Doctoral Dissertation
- ☐ Master's Thesis
- ☐ Other:

1. ☐ YES ☒ NO **Will you utilize *existing identifiable private* information about living individuals?** "Existing" information is data that were previously collected for some other purpose, either by the researcher or, more commonly, by another party. "Identifiable" means that the identities of the individuals can be ascertained by the researcher by name, code number, pattern of answers, or in some other way, regardless of whether or not the researcher needs to know the identities of the individuals for the proposed research project. "Private" information includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place or information provided for specific purposes that the individual can reasonably expect will not be made public (e.g., a medical record or student record).

   Note: If you are using data that: (1) are publicly available; (2) were collected from individuals anonymously (i.e., no identifying information was included when the data were first collected); (3) will be de-identified before being given to the researcher, (i.e., the owner of the data will strip identifying information so that the researcher cannot ascertain the identities of individuals); or (4) do not include any private information about the individuals, regardless of whether or not the identities of the individuals can be ascertained, your response to Question 1 should be NO.

2. ☐ YES ☒ NO **Will you *interact* with individuals to obtain data?** "Interaction" includes communication or interpersonal contact between the researcher and the research participant, such as testing, surveying, interviewing, or conducting a focus group. It does not include observation of public behavior when the researcher does not participate in the activities being observed.

3. ☐ YES ☒ NO **Will you *intervene* with individuals to obtain data?** "Intervention" includes manipulation of the individual or his/her environment for research purposes, as well as using physical procedures (e.g., measuring body composition, using a medical device, collecting a specimen) to gather data for research purposes.

If you answered YES to ANY of the above questions, your research is subject to Institutional Review Board oversight. Please discard this form and complete and submit an IRB application. Do not begin your research until your application has been reviewed by the IRB and you are informed of the outcome of the review.

◆◆◆◆◆◆◆◆◆◆◆◆

If you answered NO to ALL of the above questions, your research is not subject to Institutional Review Board oversight. Stop here, sign below, secure your faculty advisor's signature, and submit this form to the Graduate School. Please remember that, even though your project is not subject to IRB oversight, you should still observe ethical principles in the conduct of your research.

STUDENT CERTIFICATION: I certify that my responses to the above questions accurately describe my proposed research.

Student's Signature: _____ **Date:** March 6, 2016

FACULTY ADVISOR CERTIFICATION: I have reviewed the student's proposed research and concur that it is not subject to Institutional Review Board oversight.

Faculty Advisor's Signature: _____ **Date:** 4-10-17