REPORT

Determinants of Productivity for Military Personnel

A Review of Findings on the Contribution of Experience, Training, and Aptitude to Military Performance

Jennifer Kavanagh

Prepared for the Office of the Secretary of Defense

TR-193-OSD

RAND NATIONAL DEFENSE RESEARCH INSTITUTE

The research described in this report was sponsored by the Office of the Secretary of Defense (OSD). The research was conducted in the RAND National Defense Research Institute, a federally funded research and development center supported by the OSD, the Joint Staff, the unified commands, and the defense agencies under Contract DASW01-01-C-0004.

Library of Congress Cataloging-in-Publication Data

Kavanagh, Jennifer, 1981–
Determinants of productivity for military personnel : a review of findings on the contribution of experience, training, and aptitude to military performance / Jennifer Kavanagh.
p. cm.
"TR-193."
Includes bibliographical references.
ISBN 0-8330-3754-4 (pbk. : alk. paper)
1. United States—Armed Forces—Personnel management. 2. Productivity accounting—United States. I.Title.

UB153.K38 2005
355.6'1—dc22

2005003667

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Published 2005 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
201 North Craig Street, Suite 202, Pittsburgh, PA 15213-1516
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

PREFACE

This report discusses the primary literature and empirical findings related to three major factors that affect military personnel productivity: experience, training, and ability. It represents a portion of a larger research project concerned with the setting of retention requirements for the armed forces. The study responds to the question of the optimal experience and skill mix for the current armed forces, a question that is of increasing relevance to manpower planners as technology develops rapidly and as national security concerns evolve. This literature review is intended to serve as a point of departure for a discussion of issues relating to the performance benefits of experience, training, and innate ability and also as a summary of the research already completed in this area. The report will be of particular interest to policymakers and planners involved in the manpower requirement determination and personnel management processes as well as to participants in the training and recruiting aspects of force shaping. This Technical Report will eventually be incorporated into a larger publication that will include a more complete description of the project's objectives, findings, and recommendations.

This research was sponsored by the Office of Military Personnel Policy and was conducted for the Under Secretary of Defense for Personnel and Readiness. It was conducted within the Forces and Resources Policy Center of the RAND National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the unified commands, and the defense agencies. Comments are welcome and may be addressed to Jennifer Kavanagh, RAND Corporation, 1776 Main Street, Santa Monica, California 90407, or Jennifer_Kavanagh@rand.org. For more information on RAND's Forces and Resources Policy Center, contact the Director, Susan Everingham. She can be reached at the same address, by e-mail: susan_everingham@rand.org, or by phone: 310-393-0411, extension 7654. More information about RAND is available at www.rand.org.

CONTENTS

TABLES

SUMMARY

The literature describing the determinants of military personnel productivity offers an empirical perspective on how experience, training, and individual aptitude affect personal and unit performance. It also provides insight into the determination of the optimal skill and experience mix for the armed forces. The relationship between personnel productivity and each of these determinants is important because it affects the personnel development processes of the armed forces and ultimately contributes to overall force readiness and capability.

Although this issue appears relatively straightforward, a deeper analysis reveals several challenges. First, it is important to note that the military carries out many different activities, ranging from combat to more technical operations, each of which may require a different experience mix or a different amount of training. For example, technical positions, such as communications or radar operations, may benefit from having a large number of highly proficient personnel, whereas administrative occupations may exhibit lower returns to additional training and experience. A second challenge is the difficulty of defining the proper unit of output for measuring productivity. There are several possible choices including supervisor ratings, which are more subjective, or individual task performance scores, which measure the accuracy or success of personnel on specific activities. Both of these are acceptable measures, but neither is able to capture the full meaning of productivity. Importantly, the choice of an output measure is related to the definition and measurement of experience more generally.

The majority of studies concerning the relationship between productivity and experience, training, or aptitude find that each of these three factors contributes significantly to personnel productivity. As one example of the effect of experience on productivity, Albrecht (1979) uses supervisor ratings taken at four separate points during individual careers to determine how the productivity of first-term personnel differs from that of careerists. He finds that careerists are from 1.41 to 2.25 times as productive as first-term personnel. Most

studies confirm the basic results of this study, although there is some discrepancy over the actual quantitative effect of experience. Furthermore, it is important to remember that, as mentioned above, the size of the experience differential is likely to vary based on the nature and requirements of a given occupation.

Additional training has also been found to consistently affect productivity of personnel. Training appears to be significant as a source of skill acquisition, knowledge building, and capability development. Many studies suggest that it is the accumulation of training over a lifetime that has the largest effect on individual performance, rather than simply training in the previous six months. In order to study this effect, Hammon and Horowitz (1990) look at how additional hours of training, both short-term and long-term, affect performance on several different tasks, including marine bombing, carrier landings, and air-to-air combat. They find that positive performance effects result from additional training in each of these activities. In the carrier landing exercise, for example, individuals were scored on a seven-point scale, ranging from dangerous to excellent. The effect of a career decrease in training hours of 10 percent led to a 10 percent increase in the number of unsatisfactory landings, from 14 percent to 24 percent of the total, and a 5 percent decrease in the number of excellent landings, to 28 percent of flights. These results imply that additional training can improve proficiency, reduce performance error, and lead to a higher technical skill level among personnel.

A final determinant of personnel productivity that will be discussed in this report is Armed Forces Qualification Test (AFQT) score as a measure of individual ability. A representative study of the effect of AFQT on performance was conducted by Winkler, Fernandez, and Polich (1992). Their study looks at the relationship between AFQT and the performance of three-person teams on communications tasks, including making a system operational and troubleshooting the system to identify faults. They find a significant relationship between the group's average AFQT score and its performance on both activities. On the first task, they find that if the average group AFQT is lowered from the midpoint of category IIIA to the midpoint of category IIIB, the probability that the group will successfully operate the system falls from 63 percent to 47 percent. Similar results are found for the troubleshooting task; the probability that a group would identify three or more faults falls drastically as average AFQT score fell. Another important observation is that the effect of AFQT is additive, meaning that each additional high-scoring team member increases the overall performance of the team. This is particularly important in the military context, given the number of group-centered tasks the armed forces are required to complete.

The results of these studies have several important implications for manpower requirement determination processes and the future development of the armed forces. First, in certain occupations--highly technical ones for example, where returns to experience are very high--a shift to a more senior force could be cost-effective, despite the fact that senior personnel must be paid higher wages and given larger compensation packages than their more junior counterparts. This may not be true in other occupations where technical expertise and experience are less important for performance. Second, military transformation[1] and the integration of technological advances into the armed forces have a profound effect on the appropriate skill and experience mix for the armed forces as well as on the returns to experience and training. Despite this rapid evolution, the majority of literature on this topic is fairly old and outdated. This suggests that issues relating the determinants of personnel productivity should be reevaluated in the context of transformation and the developments associated with it.

A more advanced understanding of the production of military activities would be valuable to the readiness of the armed forces, the effectiveness of the manpower requirement determination process, and the recruitment and retention programs used by each of the services. Additional evidence on the relationships among personnel productivity,

---

[1] Transformation refers to the evolution and development of the military in the face of technological and national security environment changes. It includes the goal of making the force more agile and deployable.

experience, training, and ability would also allow policymakers and planners to pursue multiple, even competing objectives while also addressing technological and environmental changes that could affect the nature of their optimal structure. This report offers a framework for thinking about these issues by describing how previous research contributes to understanding the effects of personnel experience, training, and aptitude on productivity and performance.

ACKNOWLEDGMENTS

The author would like to thank James Hosek and John Romley for their assistance and advice throughout the writing of this report and Michael Polich and Craig Moore for their constructive reviews. Major Harvey Johnson and Gwen Rutherford, from OSD (P&R), also contributed their expertise to the report.

1. INTRODUCTION

The study of personnel characteristics, including aptitude, training, and experience, and their relationship with individual and unit performance is not just theoretical but has extensive practical import. More specifically, the significance of this area of research lies in its usefulness to the requirement determination, training/development, and recruitment and retention programs of the armed forces. Accurate data on the relationship between performance on the one hand and ability, experience, and training on the other would allow military officials to determine the optimal manpower mix for their force, to maximize efficiency for a given cost, or to minimize the cost of establishing a certain level of readiness. It would also allow them to better structure training and personnel development programs to increase the effectiveness of manpower utilization.

At first glance, this appears to be a relatively straightforward matter. However, there are two challenges that require a deeper investigation into the relationship between experience and performance. First, the military carries out many different activities, ranging from combat operations to more technical and mechanical jobs. Each of these activities has its own optimal experience mix, training needs, and Armed Forces Qualification Test (AFQT) distribution. For example, a combat unit is trained to operate as a team, to use specific tactics to accomplish goals, and to rely on physical endurance to complete each mission. The most efficient experience mix for such a unit is likely to be one dominated by junior personnel with a few senior commanders to oversee operations. On the other hand, more technical occupations, such as hydraulics or electronics repair, tend to depend on individuals working independently and to require a substantial amount of training. As a result, the optimal experience mix in these occupations may be a more senior one. However, it is also important to note that the increasing complexity and sophistication of weapons systems and the higher level of integration among military units may also increase the technical requirements of combat and infantry occupations. For example, more advanced communication systems, networking, and automation have made it necessary for even infantrymen to have a fairly advanced technical understanding. This suggests that the differences in requirements across specialties have also been affected by the shift to a more high-tech force and should be reevaluated in this context.

A second challenge is the selection of an appropriate measure of individual output or productivity. There are several possible choices including supervisor ratings, which are more subjective, and individual task performance scores, which measure the accuracy or success of personnel on specific activities. Both of these are acceptable measures, but neither is able to capture the full meaning of personnel productivity. The choice of an output measure is important because it relates directly to how we choose to define and measure experience and individual effectiveness.

Work by Dahlman, Kerchner, and Thaler (DKT) (2002) demonstrates the importance of identifying and maintaining the proper experience and training mix and offers a unique perspective on the issue of setting manpower requirements. These authors suggest that an individual service member must divide his time between the various goals of the overall force, which they define as (1) readiness, (2) human capital development, and (3) other administrative jobs. Readiness, the most important goal, occupies the majority of senior personnel time. This limits the number of hours that highly trained personnel have for teaching and developing the skills of younger staff members. Any time spent teaching is time not spent on readiness activities. In addition, senior personnel must also handle large amounts of paperwork and complete other administrative tasks. The result of all of these demands on personnel time is that senior members of the force are often in short supply. If retention targets are not set appropriately and if the number of senior personnel is lower than what it should be, this problem is likely to become more severe. DKT also suggest that ineffective manpower mix requirements can hurt the overall readiness of the force because junior personnel do not receive the type and quantity of training that they need and are sometimes even forced to become trainers before they are ready.

This literature review is motivated by the potential returns to force readiness that can be achieved by developing the appropriate quality and experience mix in the armed forces. Its objective is to discuss the relevant literature on the determinants of military personnel productivity. Although there is an extensive literature on this topic, the review highlights only the best military studies in this area. The issues discussed in this survey are made even more relevant by the ongoing military transformation and the changing requirements of the armed forces. Military transformation includes the evolution of a more agile, more deployable force and the integration of new technologies into the force structure. In particular, the rapid development of new technologies mandates a reevaluation of the experience mix in the existing force structure because it can have two opposing effects on the demands placed on personnel. On the one hand, many new technologies are intended to simplify military operations and maintenance. On the other, new technology brings with it new skill and training requirements. In addition, national security concerns have increased the demands on the armed forces in terms of workload and deployments. These changes may also affect the appropriate skill and grade mix in each of the services. To provide a framework for addressing these issues in more detail, this literature review describes the qualitative nature and quantitative findings of the research in three primary areas: (1) performance and productivity returns to experience, as measured by years of service and military grade, (2) the effect of additional training on performance, and (3) the role of AFQT score as a proxy for personnel quality and productivity.

2. EXPERIENCE AND PERFORMANCE

The relationship between productivity and personnel experience is an important one from the perspective of military cost and performance effectiveness. Research on this topic generally suggests that there are relatively substantial returns to experience in the form of more effective performance on a wide range of tasks, heightened accuracy, and increased productivity. If experience contributes to increased personnel productivity and if this increase in productivity is large enough to offset the cost of paying higher-ranking service members, military planners could potentially improve readiness and efficiency by targeting a higher level of retention. Gotz and Roll (1979) explore this hypothesis, arguing that a more experienced force not only would offer productivity gains but might also allow for a smaller total force that is less expensive because of lower accession and training costs. They suggest several other productivity-related benefits of a more experienced force, including the potential for skill-broadening, faster turnaround capability because of more experienced maintenance personnel, and the possibility for in-field repair of equipment. The authors' work supports the observation made in the previous section that the optimal experience mix for technical occupations is likely to be more senior than that of a more basic military occupation specialty (MOS). In fact, they suggest that it is more cost-effective to be close to the optimal mix for each individual MOS than to be close in the overall optimal experience mix for the entire force, with large variations at the occupation level. The authors, therefore, argue that the career content for the force as a whole is most effectively identified as the sum of the career contents defined for the different parts of the force. Finally, Gotz and Roll also note that even if a more experienced force structure would be beneficial, the costs of switching to such a force mix and then maintaining it through higher retention rates might be prohibitive.

One popular way to study the relative productivity of experienced and inexperienced personnel is to determine the elasticity of substitution between first-term personnel and personnel who have been in the military for several terms, known as careerists. The elasticity of substitution considers the substitutability of these two types of personnel, that is, the extent to which first-termers and careerists can be interchanged. In general, these studies find that careerists are more productive than first-term personnel, but researchers differ on the magnitude of this difference. Albrecht (1979) bases his analysis on the RAND Enlisted Utilization Survey (EUS), which was conducted in 1975. The surveys were completed by supervisors who were asked to rate individual personnel and to answer a range of questions on the utilization of the individual, the conduct of job training, and the individual's overall performance. The supervisor was first asked to describe the productivity of a typical member at four different points (after the first month, at the time of the first rating, one year after the first rating, and after four years of service), and then to describe a particular individual's productivity relative to that of the typical member. This approach was intended to adjust for possible differences across supervisors in the way they would describe a typical member's productivity. Albrecht uses a suboptimization technique that takes years in service (YOS) as a measure for experience and aims to minimize the cost of providing a given level of military effectiveness by substituting trained members of the force for inexperienced personnel. It is a suboptimization because it does not simultaneously determine the optimal level of capital (i.e., non-labor inputs) but takes capital as fixed. The model uses a production function and considers the marginal benefit and cost of additional experienced/inexperienced personnel. The author finds that careerists are 1.41 to 2.25 times as productive as first-term personnel and that this difference in productivity is larger for positions with more extensive technical requirements. Furthermore, in this model, higher skill occupations are associated with higher estimates of marginal rates of substitution and lower elasticities of substitution.[2] These findings

[2] The *marginal rate of substitution* is the rate at which two factors can be traded off while still maintaining a given level of output (i.e., along an isoquant, i.e., a line that defines the different combinations of inputs that yield a given output). In production theory, it is more commonly referred to as the technical rate of substitution.

suggest that, for high-skill occupations, the number of first-term personnel it takes to replace a careerist is relatively insensitive to other factors, particularly relative wage and numbers of personnel. A final observation made by Albrecht is that, although the returns to experience appear significant in his study, they are still finite and can be offset by the lower cost of less-experienced personnel in certain situations.

Marcus (1982) conducts a similar survey that focuses on the relative marginal products of various pay grade groups and YOS categories in the U.S. Navy. His manpower mix model was also based on a production function. The sample of personnel used in the study includes enlisted service members from many different ratings: "highly technical" positions, such as air traffic controller, aviation electronics technician, aviation fire control technician, and aviation antisubmarine warfare technician; "technical" positions, including aviation machinist's mate, aviation structural mechanic, aviation ordnanceman, aviation equipment support technician, and aviation survival equipmentman; and semi-technical" positions that encompassed all remaining positions on the ship. The ratings were assigned to categories based on skill classification defined by the Navy. Marcus's results suggest that military personnel with more experience, regardless of whether experience is measured in terms of YOS or pay grade level, also tend to have higher marginal products. For example, Marcus calculates that E7-E9 personnel have a "mission capable" marginal product[3] five times larger than that of E4-E6 personnel and nine times larger than that of E1-E3 personnel. The term *"mission capable" marginal product* refers to the marginal product of an individual at the "mission capable" level of readiness, defined as the ability to complete one and potentially all of the designated missions. Marcus also finds that

---

The *elasticity of substitution* is the change in the ratio of factor inputs that corresponds with the technical rate of substitution along a given isoquant, both measured in percentage terms.

[3] A *marginal product* is the additional output produced by one more unit of a given input. In this case, it would be the additional contribution made by adding one more service member of a particular grade to the workforce.

personnel with five to eight YOS have a mission capable marginal product about twelve times greater than that of personnel with one to four YOS. Although the magnitude of these findings may be on the high side, the results are suggestive of the important effect that experience has on productivity. It is possible to hypothesize that Marcus's results overstate the true effect of experience for several reasons. First, he gives no estimate or description of the confidence levels for his statistical findings. Depending on what these confidence levels are, his results may actually be less dramatic. Furthermore, Marcus's findings for differences among rating groups seem somewhat inconsistent and counterintuitive and do not really suggest any patterns to explain how experience may affect performance differently in various types of positions. For example, as shown in Tables 2.3 and 2.4, individuals in higher pay grades have a lower marginal product score based on mission capable rate (MCR) for more-technical positions than those in lower pay grades and a higher score based on MCR for less-technical positions. However, when considering years of service, experience does appear to contribute to higher mission capable marginal product scores, but more so in the least-technical positions--another unexpected relationship. In addition, as can be observed on Tables 2.1 and 2.2, the marginal productivity when measured with respect to number of flights (single aircraft) is sometimes negative. These findings suggest "noisy estimates" or even misspecified flight production/MCR models. Finally, the marginal product of any given group will vary based on the number of personnel in that group. As a result, some of the difference in marginal products could be explained by the existing distribution of personnel rather than by actual productivity differences. Despite these limitations, however, Marcus's findings contribute to an understanding of the relationship between experience and personnel productivity by supporting the existence of a relationship between experience and various measures of performance.

Based on his empirical findings, Marcus suggests that if the increased productivity of more experienced personnel would offset their higher cost, substantial cost savings could be earned through the shift to a more heavily senior force. This possibility is discussed more fully at the end of this section. A final relevant conclusion of Marcus's work is that although personnel in pay grades E1-E3 and those in E4-E6 can act as substitutes for each other, personnel in the higher ranks, E7-E9, are complements to both of the lower pay grade groups. This statement implies that personnel at the E-7-E-9 level have certain necessary skills that members of the lower pay grades do not possess. As a result, E7-E9 personnel may not be "replaceable" by individuals from E1-E6 pay grades but instead may contribute a unique and essential set of competencies to the force mix. Tables 2.1-2.4 show the marginal products of personnel in different pay grades and with different years of service for both highly technical and more basic occupations.

Table 2.1

Number of Flights and Marginal Products of Pay Grade Groups

	Marginal Product, Based on Number of Flights		
Position Type	E1-E3	E4-E6	E7-E9
Highly technical positions	7.2	8.0	26.5
Mid-level positions	4.9	11.2	50.5
Non-technical positions	-4.8	11.7	44.8
Overall average	-1.2	2.9	30.7

SOURCE: Marcus (1982).

Table 2.2

Number of Flights and Marginal Products of Year-of-Service Groups

	Marginal Products, Based on Number of Flights		
Position Type	1-4 YOS	5-8 YOS	9+ YOS
Highly technical positions	17.0	-4.4*	2.0
Mid-level positions	6.8	9.6	3.4
Non-technical positions	0.3	1.7	37.9
Overall average	1.3	-2.8*	14.5

SOURCE: Marcus (1982).
* Anomalous result.

Table 2.3

Mission Capable Rate and Marginal Products of Pay Grade Groups

	Marginal Products, Based on Mission Capable Rate		
Position Type	E1-E3	E4-E6	E7-E9
Highly technical positions	1.07	0.36	1.67
Mid-level positions	0.56	0.39	1.67
Non-technical positions	-0.07	0.64	0.68
Overall average	0.08	0.15	0.72

SOURCE: Marcus (1982).

Table 2.4

Mission Capable Rate and Marginal Products of Year-of-Service Groups

	Marginal Products, Based on Mission Capable Rate		
Position Type	1-4 YOS	5-8 YOS	9+ YOS
Highly technical positions	0.14	0.01	0.34
Mid-level positions	0.30	0.59	1.15
Non-technical positions	0.02	0.55	1.53
Overall average	0.01	0.12	0.44

SOURCE: Marcus (1982).

Using a different approach, Horowitz and Sherman (1980) look at the relationship between the time a ship spends in "serious failure" and the characteristics of the ship's personnel. Their sample includes ships that underwent an overhaul in fiscal years 1972-1974. The authors use both grade level and time in service as measures of crew quality to separate the effects of innate personnel quality from the productivity gains due to experience. The authors also include scores on the Shop Practices Test as an additional measure of crew quality. They use an OLS regression to determine which variables have the most significant effect on the amount of time ships spend out of commission for mechanical reasons. Horowitz and Sherman conclude that, although each of these variables has a significant effect on ship readiness, crew experience as measured by the percentage of personnel who have reached pay grade E-4 has a particularly strong negative correlation with the number of days spent in serious failure. That is, if the crew is relatively junior, with a high percentage of personnel at E-4, the ship is likely to spend more days in overhaul for serious failure.

Beland and Quester (1991) also consider the relationship between crew characteristics and the time ships spent free of mission-degrading failures. They use three different classes of ship--KNOX, SPRUANCE, and ADAMS--to make their results somewhat more generalizable. Their sample includes data from at least two separate deployments between 1981 and 1986 for each class of vessel. The authors use several different variables as a proxy for crew experience. For example, they define MANREQ as a combined measure that includes manning levels and the experience of personnel; NEWCREW to define the percentage of personnel with less than one year in the Navy; and TIME_CO to be the number of months that the ship's commanding officer has had command of the ship. The authors note, for example, that the predicted percentage of time a KNOX-class ship is free of failure (calculated at the sample means) is 70.5 percent. Like Horowitz and Sherman, Beland and Quester find that the experience of the crew, particularly its leaders, plays a role in the overall material condition of the ship. More specifically, for the KNOX class of ships, they find that moving from one standard deviation below the average CO tenure to one standard deviation above it (an increase from 6 to 21 months) leads to an increase in the time a ship is free of failures of about five percentage points, to 75.5 percent. Furthermore, their results for the KNOX class suggest that increasing the percentage of new crew members from one standard deviation below the mean to one standard deviation above the mean leads to a decrease of about eight percentage points in the time a ship is free of failures. Similar findings are also found for the other classes of ships used in the study. When combined, these two findings are significant because they suggest that maintenance problems are more likely when crews are less experienced and that these problems can only be partially offset by increased CO tenure. Table 2.5 offers a complete summary of the results for this study for each class of ship.

Table 2.5

Predicted Percentage of Time Free of Failure

		Prediction		
Variable	Value of Variable	KNOX	SPRUANCE	ADAMS
All variables	Mean	70.49	69.92	51.01
MANREQ	One SD above mean	76.36***	82.50***	63.16***
	One SD below mean	64.06***	54.46***	37.78***
NEW CREW	One SD above mean	66.18***	62.82**	45.74***
	One SD below mean	74.44***	76.27**	56.26***
TIME_CO	One SD above mean	72.9*		68.76***
	One SD below mean	68.01*		33.04***
Chi square		72.0	128.1	110.4

SOURCE: Beland and Quester (1991). Method: Tobit. * significant at .1 level; ** significant at .05 level; *** significant at .005 level.

Activity analysis can provide additional insight into the relative productivity of personnel of different experience levels by using linear programming to link the productivity of a workforce to its size and constituent structure. Activity analysis determines the amount of each type of personnel that would be required to complete a certain allocation of work. Activity analysis, therefore, provides insight into how different experience mixes contribute to the completion of assigned tasks. It recognizes that a given workload can be completed through the use of different workforce structures and work allocations (Doyle, 1998). Doyle (1998) uses activity analysis to study how changes in the experience mix affect work allocation and task completion among Air Force personnel working in Aerospace Ground Equipment (AGE) maintenance units. Through a trade-off analysis, she finds that if a less experienced unit is expected to complete the same amount of work in the same period of time as a more experienced unit, then the size of the less experienced unit must be increased. For example, when comparing a unit split evenly between first-termers and careerists to one with 40 percent first-term personnel and 60 percent careerists, Doyle finds that the less experienced unit requires 3 percent more time to accomplish the assigned work. A unit split 60-40 between first-term and career personnel will take 5 percent longer to complete the task than the 40-60 split unit. If the first-term percentage is increased to 70, then this less experienced unit will take 8 percent longer than the same more experienced 40-60 split unit. The author suggests that manpower requirements for a given unit should take the experience mix into account. Learning curves that compare task completion times for various experience groups support this finding. The author derives learning curves for training, supervisory, and regular work. The learning curves suggest that regardless of task difficulty, the time to complete a task decreases as years of service increase. However, it is also true that the difference between inexperienced and experienced personnel completion times is most pronounced for the most difficult tasks. For example, for regular work, inexperienced personnel will take 1.25 times as long as experienced personnel for the least difficult task but almost twice as long to complete the most difficult task. These observations offer evidence for the importance of experience for efficient performance.

Doyle also finds only marginal time savings from assigning more or less work to airmen with a given experience level. The most significant savings come from changes to the least difficult work assigned to individuals with two years of experience. In this case, if one minute more per day of the least challenging type of work were assigned to individuals with two years of experience (rather than being assigned to those in a different experience group) the AGE unit would save 27 minutes in the time it took to complete a month's work. Savings are largest where individuals of a given experience have the highest relative productivities when compared to other experience groups. For example, personnel with two years of experience have higher relative productivities for less challenging tasks than for the most challenging work. Finally, Doyle's analysis suggests that the contribution of experienced personnel to task completion can be significant and that overall unit work time can be reduced if the most experienced personnel are assigned less supervisory duty and are given more of the most challenging work.

Moore (1981) also uses activity analysis to examine the relative productivities of Air Force AGE personnel. He finds that when both performance and supervision time are included, the most junior personnel (E1-E3) take an average of about 2.4 times as long (in man hours) to complete a fixed amount of troubleshooting than people in the most experienced category (E6 and E7). Moore also finds, however, that the contribution of experience varies for different tasks. For example, on a corrosion control exercise, which could consist of any activity to prevent corrosion of aircraft and equipment including cleaning, painting, or application of protective coatings, junior personnel take only about 1.5 times as long as senior personnel to complete a given amount of work. Moore's work strengthens Doyle's argument that a less experienced workforce will take longer to complete a given amount of work unless they are provided with additional manpower (see Table 2.6).

Table 2.6

Time to Complete Task, Based on Experience

YOS, Skill Level*	Work Time,** Troubleshooting	Work Time plus Supervision, Troubleshooting	Work Time, Corrosion Control	Work Time plus Supervision, Corrosion Control
E1-E3, 3	2.1	2.4	1.3	1.5
E3, 5	1.7	2.1	1.2	1.3
E4, 5	1.6	1.8	1.1	1.1
E5, 5	1.4	1.5	1.1	1.1
E5, 7	1.2	1.3	1.0	1.0
E6-E7, 7	1.0	1.0	1.0	1.0

SOURCE: Moore (1981). *Skill-level defined by Air Force as 3, 5, 7.
** Work time data are provided in a ratio form where time for the highest skill level to complete the job is defined as 1.0.

Economic models of retention goals are also useful for a discussion of the returns to experience because they can offer a more precise analysis of the most efficient experience mix and the trade-offs between recruits and senior personnel. For example, Moore, Golding, and Griffis (2001) develop a method to measure the cost-effectiveness and readiness effects of a shift to a more senior force through higher reenlistment rates and lower accession numbers. They look specifically at the Navy and assess the costs and benefits of different types of force mix. From a cost perspective, they find that raising reenlistment targets is not an effective way to meet end-strength goals because the cost of retaining senior personnel exceeds that of hiring and training new recruits. In their model, the cost of new recruits is equal to the recruiting cost, the salaries of instructors, the costs associated with Permanent Change of Stations (PCS), and the costs of paying students with Immediate Active Duty status who are also in school. The costs of retaining senior personnel include reenlistment bonuses, medical and retirement plan accruals for the personnel induced to stay, and higher salaries due to seniority. The reenlistment bonus makes up the majority of these costs and is actually defined as a range because these bonuses can vary in size. According to the estimates used in this study, the cost of meeting end-strength goals by raising Zone A reenlistment by two points would be between $78 million and $169 million per year, whereas the cost savings from lower accessions would be only $36 million per year. Importantly, it is not clear if the authors account for the fact that both the marginal cost of recruiting and the cost of retaining an extra person are likely to be rising. If they do not properly consider this fact, the costs of raising retention numbers will be higher than estimated and the benefits of reducing recruiting will be lower than calculated.

However, as the preceding discussion about the returns to experience implies, this question cannot be considered from a purely financial perspective. The shift to a more senior force would also lead to an increase in average experience and force readiness. Depending on the estimated economic value of this readiness, aging the force could be a cost-effective approach to increasing force preparedness and efficiency. The authors calculate that the value of readiness would need to be between $135 and $427 per sailor. Currently, the Navy pays $140 more per sailor for an additional 1.2 months of seniority. The authors assume that this rise in payment is the value of the additional readiness provided by a 1.2 month increase in average seniority, and they use this assumption to argue that, in this case, the additional cost of a more senior force would be offset by readiness gains only for the lowest cost estimates. However, the authors do not give us any reason to accept this assumption as valid. The authors go on to consider how retention and recruitment policies should differ between occupations at different skill levels. They find that the difference between recruitment/training savings and retention/seniority costs is largest (most negative) for the low-skill occupations. When factoring in readiness, the cost of a more senior force (using the upper estimate of the cost range) would be offset by savings and readiness gains for high-skill occupations, but would far exceed the benefits of a retention-based program for low-skill occupations. As a result of their analysis, the authors come to the conclusion that aging the force as a means to meet end-strength targets can be a cost-effective way to increase force readiness, particularly in high- and some mid-level skill occupations, but is not an efficient way to reduce the cost of maintaining a certain end strength or to limit the strain put on recruiting. Of course, this depends on the cost of recruiting and training new sailors, which can vary based on the external factors such as the strength of the private-sector economy. One shortcoming of this study, however, is that it fails to account for the cost savings that are due to the more efficient or effective use of equipment by senior personnel. These cost savings could result from additional increases in the productivity of senior personnel or from lowered maintenance and replacement expenses.

Overall, these findings suggest that the experience level of military personnel offers high returns in the form of increased productivity and improved readiness but can also increase the costs of maintaining a given end strength. Applying this observation to the goal of achieving national security at minimum cost, a more senior force may be a cost-effective approach in some occupational groups, depending on the benefits and costs of greater experience. In order to examine this issue more closely, a model of retention goal-setting that considers the dynamic contribution of technology and military transformation to the effectiveness of the force and to the optimal manpower mix would seem necessary and useful.

3. TRAINING AND PERFORMANCE

The relationship between additional training and individual performance is important to this discussion because training is a variable that can be directly manipulated and controlled by the military. Although the recruiting and retaining of high-quality or highly experienced personnel can be affected by policy, there are still unknown and uncontrollable factors involved, such as personal preferences and the strength of the private-sector economy. However, the amount and type of training given to military personnel can be more easily adjusted up or down to optimize the cost-effectiveness of training with respect to performance. It is worth noting at the outset that although studies on the relationship between training and performance have been conducted for several different aircraft-related tasks (within the Air Force, Navy, and Marines), there is a lack of research concerning the effect of training on ground or other naval operations. It is possible that the services have conducted this type of research for their own benefit only. However, this appears to be an area that would benefit from additional research.

One of the most extensive studies on this topic, conducted by Hammon and Horowitz (1990), assesses and differentiates the effects of additional lifetime training, additional training in a short-term perspective, and simulation training on the performance of military personnel in a variety of air combat exercises. The authors consider three exercises: carrier landings, marine bombing, and air-to-air combat. They find that while both short-term and career flying hours contribute to improved performance, accumulated training hours have the strongest effect on individual performance over the long term. In the carrier landing exercise, individuals were scored on their carrier landings on a seven-point scale that can be broken down as follows: 0 = dangerous; 1 = wave off, pilot instructed not to land; 2 = no grade, landing made but deemed faulty; 2.5 = bolter, aircraft touched down but did not catch arresting wire; 3 = fair pass, some errors, but overall technique was ok; 4 = ok pass, a successful landing, the highest grade a pilot should expect; 5 = rails pass, perfect landing, rarely given. To summarize the data, 86 percent of the results were at least satisfactory and 33 percent were excellent. The authors use a logit model to compare the results of the carrier landings with pilot experience, career training hours, and recent training hours. The results suggest that additional training has a significant effect on landing performance. For example, in the carrier landing exercise with one of the two planes tested, the F-14, the authors find that a 10 percent decrease in the number of recent flying hours would have the short-term effect of decreasing the number of excellent landings by 2.5 percentage points and increasing the number of unsatisfactory landings by 2.6 percentage points. On the other hand, a career decrease of 10 percent in the number of hours flown would lead to a decrease of five percentage points in the number of excellent landings, from 33 percent to 28 percent of the total landings, and a ten percentage point increase in the number of unsatisfactory landings, from 14 percent to 24 percent of the total. These percentage effects are relatively significant in their own right, and the magnitude of small changes in performance is increased when we consider the huge cost required to repair planes or other equipment damaged by faulty landings. It is worth noting that at least some portion of the trends observed in Table 3.1 could be due to the fact that the most proficient, high-performing pilots are likely to stay in the service the longest and accumulate the most career flying hours. In this case, the high performance of those with the most career flying hours would be due less to additional training than to individual aptitude. Table 3.2 shows the relationship between flying hours the previous month and landing performance and reflects the fact that both recent and cumulative training contribute to improved performance.

Table 3.1

Career Training and F-14 Landing Performance (predicted probability)

Career Flying Hours	Satisfactory Landing	Excellent Landing
500	.79	
1,000	.81	.23
1,500	.83	.25
2,000	.85	.26
2,500	.87	.27
3,000	.88	.28
3,500	.90	.29
4,000	.91	.31
4,500	.93	.32

SOURCE: Hammon and Horowitz (1990).

Table 3.2

Training in Previous Month and F-14 Landing Performance (predicted probability)

Previous Month's Flying Hours	Satisfactory Landing	Excellent Landing
0	.83	.20
5	.83	.21
10	.84	.22
15	.85	.23
20	.86	.24
25	.86	.25
30	.87	.26
35	.87	.27
40	.88	.28
45	.88	.30

SOURCE: Hammon and Horowitz (1990).

Similar results are observed for the marine bombing exercise. The model developed for this task describes the relationship that exists between career and previous-week flying hours and bombing miss distance in feet (see Tables 3.3 and 3.4). According to their results, the authors predict that a pilot with 3,000 career hours of experience/training can be expected to place bombs 15 feet closer to the target than a pilot with only 1,500 hours. This effect is also significant at smaller intervals of career experience. For example, a pilot with 1,500 hours of career training will also perform better than a pilot with only 500 hours, placing his bombs about 8 feet closer to the target. These results appear significant considering that the mean miss distance is 83 feet and the mean career hours of flying experience is 1,598. Short-term training (in the previous week) also has a substantial effect on pilot performance. A pilot with 15 flying hours in the previous week is likely to place his bombs 15 feet closer to the target than a pilot with only 5 hours of flying time in the previous week (mean flying hours in past week is 4). The authors argue that the overall effect of training accumulated over an individual's career is likely to be larger than the effect of training in the short run because training over a lifetime helps to build skill mastery. Although the results of this study support the importance of training for pilot performance and accuracy, the authors do not consider how much reductions in circular error for bomb delivery would affect operational outcomes, for example the likelihood that the target was destroyed or supplies were received. Because the ultimate goal of any training program is to improve these operational outcomes, further research on this relationship seems important.

Table 3.3

Career Training Hours and Bombing Error (feet)

Career Flying	Bombing Error		
Hours	F/A-18	F-4S	AV-8B
500	97	145	120
1,000	90	140	115
1,500	85	133	110
2,000	80	128	102
2,500	78	121	100
3,000	76	120	95
3,500	70	117	87
4,000	65	110	80
4,500	60	103	78
5,000	55	98	72
5,500	50	93	70

SOURCE: Hammon and Horowitz (1990).

Table 3.4

Training Hours in Previous Week and Bombing Error (feet)

Previous Week Flying Hours	Bombing Error		
	F/A-18	F-4S	AV-8B
0	100	145	120
5	90	140	110
10	78	120	100
15	58	115	80
20	35	95	65

SOURCE: Hammon and Horowitz (1990).

Finally, the results for the air-to-air combat exercise support the observations drawn from the first two exercises. The combat exercise was carried out using a program in which several highly trained pilots simulate Soviet tactics. Each exercise consists of a control phase and a weapons phase. During the control phase, aircraft crews are instructed to maintain radar lock-on and position themselves for an attack. During the weapons phase, which begins when an enemy aircraft is sighted and a weapon is fired, crews attempt to kill as many of the enemy aircraft as possible without being killed themselves. The number of "kills" is recorded, along with the speed, range, acceleration, and altitude of each firing. According to the results of their analysis, the authors find that a 10 percent decrease in career training time led to a 5 percent decrease in the number of times the subject was able to kill his computerized opponent and a 9 percent increase in the number of times he was killed. The authors also note that 85 percent of the expected change in enemy kills and 80 percent of the expected change in trainee kills are attributable to changes in pilot flying hours (combining both career and recent flying). In each case, the effect of the short-term training variable was smaller than that of career flying hours but still significant. Pilot career flight time was the most important single factor, accounting for 65 percent of the increase in enemy kills and 42 percent of the decrease in trainee kills. Again, the effect of career experience is likely to be more significant because training over the long term contributes to mastery of a task. (See Table 3.5.)

Table 3.5

Career Training Hours and Air-to-Air Combat Performance (predicted probability)

Career Flying Hours	Blue Kill	Red Kill
500	.35	.14
1,000	.37	.12
1,500	.40	.09
2,000	.42	.08
2,500	.43	.07
3,000	.45	.05
3,500	.47	.04
4,000	.49	.03

SOURCE: Hammon and Horowitz (1990).

Hammon and Horowitz (1992) consider a final example, C-130 air drop accuracy, and extend their results by considering the effect of simulator-based training on performance. The C-130 air drop involves parachute drops of personnel and equipment into drop zones. The primary objective measure of drop performance is the distance from the intended point of effect to the actual landing point. Although the navigator is the key crewmember for the proper execution of this task, coordination among all crewmembers is needed to ensure effective performance. The model developed for this example included variables for career and short-term flying hours for both the copilot and the navigator and defined a relationship between flying hours and crew performance. The authors draw several relevant observations from their analysis. First, neither the short-term copilot variable nor the long-term navigator variable was significantly related to performance. However, the long-term copilot variable and the short-term navigator variable both had a significant effect on drop accuracy. More specifically, according to the reported results, in the case of copilot career flying hours, an increase from 500 to 1,500 hours of training corresponded with a decrease of 15 yards in average circular error (Table 3.6). A further increase to 2,500 hours of training led to a further reduction of 10 yards in the average circular error. Again, these results appear significant, given that means for career training hours and miss distance were 794 hours and 108 feet, respectively. Turning to navigator hours in the previous 60 days (mean = 65), the results suggest that an increase from 50 to 75 hours of training leads to a 10-yard decrease in average circular error and that a further increase to 100 hours of training contributed to an additional 10-yard decrease (Table 3.7).

Table 3.6

Copilot Career Training and Tactical Drop Error (yards)

Career Flying Hours	Circular Error
500	117
1,000	110
1,500	100
2,000	95
2,500	95
3,000	85

SOURCE: Hammon and Horowitz (1992).

Table 3.7

Navigator Training Hours Previous 60 Days and Tactical Drop Error (yards)

Flying Hours in Previous 60 Days	Circular Error
25	125
50	115
75	110
100	105
125	90

SOURCE: Hammon and Horowitz (1992).

It is worth noting that while the benefits of long-term training are emphasized in each of the previous studies, recent training and experience yields comparatively higher marginal returns on investment. The evidence discussed above suggests that even if a pilot has relatively little lifetime training, he can still reach a high level of proficiency if he is able to train intensively in a short period of time before a deployment or other operational employment. Because the costs of a long-term training program will be extremely high, a focus on short-term training can yield significant cost savings without sacrificing pilot performance.

Finally, the authors consider the use of simulator-based training, as either a supplement to or a replacement for more traditional training. To assess the independent effect of simulator training, the authors conduct two additional trials, one changing the number of flying hours while holding all else constant and the other increasing the number of simulator hours. The authors specifically consider the effect of simulator hours on copilot performance (Table 3.8). The authors find that the partial effect on miss distance with respect to copilot simulator hours is -.1311 compared with -.0089 for copilot flying hours. This suggests that an additional simulator hour reduces miss distance by more than an additional flying hour. However, the authors caution that these results might not hold true except near the observed values of the independent variables and note that further research in this area would be helpful. This result does have an important policy implication in that simulator hours also tend to be cheaper and less risky, in terms of possible equipment damage, than actual flying hours. If simulator training also has a more substantial effect on performance than flying hours, a training program that incorporates more simulator hours and a higher ratio of simulator time to flying time could improve both accuracy and the cost-effectiveness of military functioning.

Table 3.8

Copilot Simulator Hours and Tactical Drop Error (yards)

Career Simulator Hours	Circular Error
0	170
25	125
50	115
75	110
100	100

SOURCE: Hammon and Horowitz (1992).

An additional study worth discussing was carried out by Gotz and Stanton (1986). They consider the role of training from a slightly different perspective but one that adds a unique assessment of the way training interacts with military performance. The authors develop a computer simulation to observe the effect that cross-training of maintenance personnel--that is, the development of personnel who are able to carry out more than one repair task--has on the number of aircraft considered unusable due to maintenance problems during a combat situation. They make several assumptions and conduct several different trials under varying conditions. First, they consider a situation in which each maintenance worker can fix only one type of part. In the second trial, they relax this condition and consider a situation in which workers can fix both types of parts, but are able to complete one type of repair more quickly than the other. Finally, the authors consider a situation in which one type of part breaks down more quickly than the other. Using the results of these simulations, the authors find that cross-training does improve unit performance and contributes to a decrease in the number of aircraft that are unavailable, particularly in the middle days of the simulation period. They also find that the effect of cross-trained personnel is greatest in situations of the third type, where the parts break down at different rates. The authors build off of these findings by developing another set of scenarios that include the introduction of "high-skill personnel" who are cross-trained and highly experienced and who are able to complete maintenance tasks more quickly than average or low-quality personnel. Gotz and Stanton find that in these situations, the addition of high-skill personnel into the manpower mix contributes to a substantial decrease in the number of unavailable aircraft, again particularly in the middle days of the measurement period. The results of this study are significant, despite being based only on computer simulations, because they suggest that more advanced training or cross-training, which develops personnel who can successfully complete more than one task, can improve unit performance and military readiness. It is likely that this occurs because cross-trained personnel can be used more flexibly, in a wider range of situations, and still be expected to complete their task effectively. This observation also has implications for the development of a more productive and efficient training program, one focused on developing a high level of proficiency in several different tasks in order to maximize personnel usage and potential.

Moore, Wilson, and Boyle (1987) also consider the role that cross-training or consolidating specialties would have on manpower utilization and overall performance. Consolidating specialties would force each airman to receive training and become proficient in a wider range of skills. The authors note that combining specialties reduces the manpower required to maintain a given set of aircraft and increases manpower utilization. If individuals have a more extensive set of skills, they can contribute to many different maintenance activities. This increases the utilization of these individuals and reduces the need for additional personnel with more limited skills. These observations suggest that additional training and acquisition of new skills can significantly raise the flexibility given to manpower planners and allow the force to perform with fewer personnel (Table 3.9). However, although these positive effects are clear, combining specialties would also lead to increased training costs and time and would place a larger burden on senior personnel responsible for conducting training. The increased amount of time devoted to training would decrease productive working time, particularly for first-term personnel who make up a large portion of the military, and would offset some of the advantages gained from a combined-specialty approach. The key, therefore, would seem to be achieving a balance among additional training costs, reduced productive working time, increased utilization, and cost savings from a smaller workforce. Importantly, the training burden placed on senior personnel must figure prominently into this analysis.

Table 3.9

Effects of Consolidating Specialties

Number of Specialties	Manpower Requirements	Manpower Utilization, Percent	Average Training Days
Main Operating Base, 72 Aircraft			
1	69	87	900
3	73	78	300
5	76	76	200
7	90	69	60
10	100	60	50

SOURCE: Moore, Wilson, and Boyle (1987).

These findings concerning the relationship between training and performance are significant and relate directly to the work of Dahlman, Kerchner, and Thaler (2002), discussed at the start of this study. Because training contributes so significantly to performance and productivity, the effectiveness of military performance, as well as overall readiness, is likely to suffer if senior personnel are able to supply fewer and fewer teaching hours due to other demands on their time. Furthermore, this will be increasingly true over longer periods of time.

4. PERSONNEL QUALITY, AFQT, AND PERFORMANCE

Although experience and training are important determinants of personnel effectiveness, they are by no means the only measure of personnel quality available to military analysts. One widely used measure of quality is the score on the AFQT, a test given to enlisted personnel upon their entry into the military. High-quality personnel are commonly defined as those having AFQT scores in the top 50 percent, i.e., categories I, II, and IIIA; they also must have a high school diploma. AFQT score has been shown to be an accurate predictor of personnel quality and ability in numerous cases. AFQT and experience appear to be fundamentally different measures of quality. While AFQT measures an individual's innate ability, experience considers personnel performance and skill level as developed and manifested over time. This relationship is an important one from the perspective of our discussion because AFQT as a proxy for personnel quality can be used to guide military recruitment and requirement determinations and can aid in the development of a more effective and cost-efficient military structure.

Generally, studies conducted in this area have supported the assertion that higher-quality personnel, in this case personnel with a higher AFQT score, appear to be more productive and to exhibit generally higher performance. Scribner, Smith, Baldwin, and Phillips (1986) attempt to answer the question, "Are smart tankers better?" Using the firing scores for tanker teams in a simulation exercise conducted at the Seventh Army Training Center standardized TANK course, the authors define the relationship between performance and AFQT score for both the tanker position and the gunner position. Their model and calculations indicate a significant correlation between AFQT score and more effective performance on the simulation exercise. For example, they find that an increase in AFQT score from category IV to category IIIA leads to an improvement of 20.3 percentage points in performance. A similar increase in AFQT for the gunner in the same exercise will lead to a performance increase of 34 percentage points. These results are consistent with the arguments that AFQT score is an effective indicator of personnel quality and that having a force made up of personnel with higher AFQT scores contributes to more effective and accurate team performance.

A study by Winkler, Fernandez, and Polich (1992) offers additional support and evidence for this finding. The authors examine the relationship between AFQT score and the performance of two communication activities. The sample included 84 groups from active-duty signal battalions and 240 teams recently graduated from the Signal Center's advanced individual training (AIT) course. In the first task, the three-person teams were asked to make a communication system operational. In the second, the teams were expected to identify and repair a number of faults in the communication system. The authors then used a multivariate model to characterize the relationship between various characteristics of the group and individual personnel and the team's success at the assigned tasks. The multivariate model allows the effect of AFQT on performance to be isolated from the effects of other variables, as though the other variables were held constant. Their results suggest that average group average AFQT has an effect on team performance and success at completing the task. Furthermore, this effect holds for each of the two test tasks. More specifically, the model predicts that for active-duty units with an average AFQT at the midpoint of category IIIA, there is a 63 percent chance that the unit will successfully operate the system in the allowed time. However, if the average AFQT is lowered to the midpoint of category IIIB, the probability of successful completion falls to 47 percent (Table 4.1). A similar decline can be observed for the AIT graduates, although the AIT graduates start from a somewhat lower probability of success at all aptitude levels. This difference is most likely due to their lower level of experience. When group average AFQT score is reduced for the AIT graduates from the midpoint of category IIIA to the midpoint of category IIIB, the probability of success declines from 40 percent to 25 percent.

Table 4.1

Successful System Operation and AFQT
(predicted probability)

Sample Members	CAT I	CAT II	CAT IIIA	CAT IIIB	CAT IV
Unit members	.89	.80	.63	.47	.29
AIT graduates	.76	.60	.40	.25	.13

SOURCE: Winkler, Fernandez, and Polich (1992).
NOTE: The midpoint in each AFQT category is used in predicting the probability of successful operation.

The results from the troubleshooting task offered similar evidence for the correlation between higher AFQT scores and more effective performance (Table 4.2). For example, the probability that groups of AIT graduates will correctly identify three or more faults falls from 66 percent when the group average AFQT is at the midpoint of category I to 49.4 percent when the average AFQT is at the midpoint of category II and declines even further to 29.4 percent when the group average AFQT is at the midpoint of category IIIA. The chart below provides more extensive representations of the results from this study to further demonstrate the extent and magnitude of the effect of aptitude on performance.

Table 4.2

Group Troubleshooting and AFQT, AIT Graduates
(predicted probability)

	Faults detected			
AFQT level	1 or More	2 or More	3 or More	4 or More
Cat I	.97	.97	.66	.29
Cat II	.94	.78	.49	.17
Cat IIIA	.87	.60	.29	.08
Cat IIIB	.78	.43	.17	.04
Cat IV	.61	.25	.09	.02

SOURCE: Winkler, Fernandez, and Polich (1992).
NOTE: The midpoint in each AFQT category is used in predicting the probability of successful fault detection. Cell entries are the predicted probability that the group will successfully identify the given number of faults.

The authors also note that the addition of another high- scoring member to the team improved the probability of success by about 8 percent. This suggests that the effect of AFQT on group performance is additive. This finding is significant for an assessment of the optimal force mix because it implies that AFQT continues to make a difference in team performance even when considering the contribution of a second or third team member.

The work of Teachout and Pellum (1991) supports the relevance of AFQT to job performance. The authors consider how AFQT scores are related to hands-on performance test (HOPT) scores for Air Force maintenance positions. For each of the eight specialties considered, the mean HOPT score is higher for those with AFQT scores ranging from I to IIIA than for those with lower AFQT scores. Except for a few cases, the authors find that this trend holds regardless of the experience level of personnel studied. This is a significant observation because it suggests that aptitude, as measured by AFQT, remains an important predictor of job performance even after an individual has been serving for three years.

A final study that offers evidence of the correlation between AFQT scores and performance is Orvis, Childress, and Polich (1992). In this study, the authors used controlled trials to assess how AFQT score was related to various aspects of air defense and Patriot air defense system operation. The study included several types of air defense situations: point defense, asset defense, missile conservation, area defense, and a mixed defense scenario (Table 4.3). Service members were also tested on their tactical kills/success in air-to-air combat and their overall battlefield survival (Table 4.4). The authors argue that their results show a significant relationship between AFQT score and the outcomes of air battles or defense scenarios, both in terms of knowledge assessed by written tests and performance in simulations. The authors compared the effects of several explanatory variables, including AFQT score, years of operator experience, unit member, and simulation training each ten days. They found that AFQT demonstrated more significant relationships with simulation outcomes than did any of the other variables. In an effort to quantify the effect of AFQT on performance in their model, the authors note that the effect of a one-level change in AFQT category appeared to equal or surpass the effect of an additional year of operator experience as well as the performance effect of additional simulation training. This observation is not meant to imply that the trade-offs or relationships between AFQT and years of experience or additional training are linear. Rather, the authors note that although the magnitude of the trade-off may vary, it is at least one-to-one and in some cases even larger. This finding and the ones above support the military's emphasis on ensuring that a significant fraction of its recruits are high-quality, high-AFQT personnel.

Table 4.3

AFQT and Patriot Air Defense System Operator Performance, Probabilities of Success

	AFQT Category				
Activity	I	II	IIIA	IIIB	IV
Mixed defense	65	57	46	39	30
Point defense	64	57	47	39	31
Mixed defense					
First priority*	56	53	49	45	41
Second priority	67	58	46	37	28
Point defense					
First priority	57	53	48	44	40
Second priority	61	55	48	42	35
Third priority	64	56	47	40	32
Battle survival	68	58	46	37	26

SOURCE: Orvis, Childress, and Polich (1992).
NOTE: Maximum score in each cell is 100 points.
* "Priority" indicates the priority given to the task by the simulation program.

Table 4.4

AFQT and Patriot Air Defense System Operator Performance, Specific Measures

	AFQT Category				
Measure	I	II	IIIA	IIIB	IV
Asset hits (maximum 28)	10	11	12	13	14
Hostile kills (maximum 78)	53	51	48	45	42
Number missiles used for 10 tactically correct kills	20	21	22	23	24

SOURCE: Orvis, Childress, and Polich (1992).

The relationship between AFQT score and individual and unit performance suggests the importance of recruiting high-quality, high-AFQT personnel as a foundation for creating high-performing units. The recruitment of high-AFQT personnel will be even more significant if the AFQT mix that is initially recruited is generally the one that will be retained and will remain throughout a given cohort's term of service (unpublished 1998 RAND work by Asch, Hosek, Mattock, and Warner). This finding implies that it may be more difficult to adjust the AFQT mix of personnel after the initial recruitment period.

5. CONCLUSION

Improvements in our understanding of the production of military activities would be valuable. Interest in experience, training, personnel quality and flexibility, and teamwork is long-standing. However, the military context has changed. The armed forces are smaller, richer in careerists, and more reliant on technology. Our political leadership has tasked the services with missions of greater scale and scope. And the world is a less certain place. New concerns about the implications of operational and personnel tempo and the distribution of responsibilities through the ranks of the hierarchy may be well-placed. We must apply rigorous methods to these salient issues in manpower policy. A fuller understanding would aid policymakers and planners in their pursuit of multiple objectives.

While the studies reviewed in this report have made important contributions to the question of military personnel effectiveness, our understanding of this issue remains limited in important respects. To begin with, the distinct roles of innate ability, formal training, and informal learning deserve greater attention. Each of these factors influences members' human capital and thus their effectiveness, and policymakers should consider trade-offs among them. Next, the studies reviewed here largely examined the military of the 1980s. Since then, the scale and scope of operations have grown; many functions, including combat arms and logistics, have experienced technological advances; and the career content of personnel has risen. For each of these reasons, our knowledge of the relative effectiveness of members by tenure and grade is dated. Finally, there are important gaps in our understanding. For example, with the stress of increased PERSTEMPO, effectiveness in a mission might decline in the near term but improve in the longer term for all personnel. Furthermore, the returns to a regimen of cross-training have not been measured.

Turning to the organization of the military workplace, greater allocation of a ranking member's time to administrative tasks may elicit more effort from those overseen, but this increase in effort would occur at the expense of training. The benefits of forming personnel into production teams are presently unknown. Careful analysis—perhaps using controlled trials in some instances—would be informative about these issues. Credible evidence on the full range of factors influencing personnel effectiveness in today's military would aid policymakers in their pursuit of competing objectives. The quality of decisions concerning force structure and retention goals, in particular, stands to benefit from such evidence.

APPENDIX: STUDY SUMMARIES, METHODS, AND EMPIRICAL RESULTS

STUDIES ON EXPERIENCE AND PERFORMANCE

Title	*Labor Substitution in the Military Environment*
Author	Mark Albrecht
Date	1979
Method	Survey that takes much of its data sample from the RAND EUS dataset collected in January-February 1975. This data collection involved selecting individuals and determining their primary supervisors. The supervisors were sent rating sheets for each individual that included questions on the utilization of the service member, the conduct of job training, and performance. Productivity was assessed at four points, his first month, the time of the fist rating, one year after that, and after four years of service.
Functional Form	Results were estimated using OLS and the following function form: $\ln MP_i = a_i/a_j + b_1 \ln (L_i/L_j) + b_2 \ln MP_j + u$ L_i= Supply of labor provided by individuals in the i^{th} year of the first term of service. L_j= Supply of labor provided by individuals in the j^{th} year of service. MP_i= Marginal product of someone in the i^{th} year of service.
Summary Findings	There is a marginal rate of substitution of first-term personnel for careerists of 1.41 to 2.20. However, the author also notes that the return to experience is finite and can be offset by the lower cost of less experienced personnel. Significant cost savings are associated with the shift to the optimal manpower mix. While a more senior force might increase the effectiveness of the force, it is also true that increasing the number of careerists would (all else held the same) increase the marginal productivity of first- termers and lower their cost. Other conclusions: (1) more technically demanding occupations have more limited substitution opportunities of first-term personnel for careerists; (2) higher skill level occupations are associated with higher estimates of marginal rates of substitution and lower elasticities of substitution.

Quantitative Results	Estimates of Substitution Elasticities for First-Termers and Careerists N=4,592			
	Air Force Specialty Code (AFSC), Constrained Elasticity of Substitution Model (CES)	$\sigma_{1,2}$	β=0 Std. Error	β=-1 Std. Error
	326X0	2.3	1.3	1.0
	326X1	1.25	.8	1.5
	326X2	1.25	.2	.9
	304X4	5.01	2.3	.6
	306X0	5.05	2.8	1.2
	421X3	2.57	1.1	.7
	422X1	1.81	.4	.5
	431X1	2.14	1.1	1.0
	542X0	.82	.3	1.4
	543X0	4	2.2	.71
	571X0	4.48	2.7	.8
	622X0	4	4.3	1.4
	631X0	8.92	12.2	1.5
	647X0	3.61	.4	.2
	671X3	4.08	2.2	.7
	902X0	1.71	.4	.5
	981X1	5.23	3.1	.7
	AFSC, Weighted Linear Model	$\sigma_{1,2}$	β=0 Std. Error	Marginal Rate of Substitution: First Term to Career
	326X0	2.81	3.5	2.2
	326X1	1.42	.8	2.17
	326X2	1.58	.8	2.25
	304X4	6.04	3.8	2.15
	306X0	5.77	5.3	2.15
	421X3	3.05	1.7	1.91
	422X1	2.08	1.3	1.95
	431X1	1.72	.7	1.72
	542X0	1.42	.7	1.92
	543X0	6.01	6.8	1.92
	571X0	5.11	4.1	1.92
	622X0	2.23	1.3	1.68
	631X0	7.19	9.1	1.42
	647X0	4.25	2.4	1.35
	671X3	5.60	5.0	1.48
	902X0	3.01	3.8	1.41
	981X1	9.44	6.6	1.45
Title	"The First-Term/Career Mix of Enlisted Military Personnel"			
Author	Glenn Gotz and C. Robert Roll			

Date	1979
Method	Uses Defense Resource Management Study analysis (carried out for Secretary of Defense in 1979) of the first term/career mix focused on six occupational specialties (low, high, and medium skill occupations from the Army and Air Force). The analysis looks for the steady-state mix of personnel that will provide the same effectiveness as the FY 77 inventory at minimum cost by determining the relative productivities and costs of first-term and career personnel.
Functional Form	NA
Summary Findings	Although results may be different for different occupations, for some occupational groups, a force with more careerists and fewer first-termers would be more cost-effective because of the relatively higher productivity of career personnel. In general, higher-skill occupations can be staffed more effectively and efficiently using career personnel, in part due to reductions in required replacement training. It is more cost-effective to be close to the optimal mix for each occupation individually than to be close in the aggregate.

Quantitative Results	Skill Code and Level	Implied Steady State (First-Term/Career)	Optimal (cost-effective) Steady State (First-Term/Career)	FY 1977 (First-Term/Career)
	Army			
	Low skill, infantryman	58/43	59/41	60/40
	Mid-skill, automotive repair	62/38	52/48	60/40
	High skill, field radio repair	49/51	39/61	43/57
	Total	58/42	56/44	56/44
	Air Force			
	Low skill, fuel specialist	44/56	43/57	47/53
	Mid-skill, aircraft maintenance specialist	43/57	40/60	47/53
	High skill, ground radio repair	56/42	51/49	49/51
	Total	47/53	45/55	48/52

Title	"A Direct Measure of the Relationship Between Human Capital and Productivity"
Author	Stanley Horowitz and Allan Sherman
Date	1980
Method	Surveys 91 ships that went through overhauls in fiscal years 1972-1974, looking at the relationship between ship downtime and the characteristics of crew personnel.
Functional Form	Linear function, OLS used to estimate the relationship. Variables used in the analysis were log of ship age, months between overhauls, dummy variables for differences in equipment, number of enlisted personnel, pre-Navy education, entry test scores, pay grade, length of service, time aboard ship, time at sea, Navy schooling, specialized qualifications, race, marital status.
Summary Findings	Experience, time in service, and scores on the Shop Practices Test have a significant relationship with the amount of downtime a ship has over a given measurement period. The authors take these variables to be indicators of crew quality. Finally, they note that there is a high payoff to having personnel who have reached pay grade E4.

Quantitative Results	Predictor Variable for Boiler Technician N=89	Coefficient (Standard Error)
	Average score on Shop Practice Test	-138 (41.3)*
	Percent at E3 or below	25.19 (8.4)*
	Percent at E8 or above	-34.06 (28.6)
	Percent with less than one year in Navy	35.65 (14.26)*
	* Significant at 1% level.	

Title	*Demand and Supply Integration for Air Force Enlisted Work Force Planning: A Briefing*
Author	S. C. Moore
Date	1981
Method	Activity analysis (linear programming) that determines the amounts of different types of personnel required to complete a given set of tasks. The technique can identify different experience mixes and manning levels to accomplish a given workload.
Functional Form	NA
Summary Findings	When both performance and supervision time are included, the most junior personnel (E1-E3) take an average of about 2.4 times as long (in man hours) to complete a fixed amount of complex troubleshooting than personnel in the

<table>
<tr><td></td><td colspan="5">most experienced category (E6-E7). Moore also finds, however, that the contribution of experience varies for different tasks. For example, on simple corrosion control work, junior personnel take only about 1.5 times as long as senior personnel. Moore concludes that a less experienced workforce will take longer to complete a given amount of work unless they are provided with additional manpower.</td></tr>
<tr><td rowspan="7">Quantitative Results</td><td>YOS, Skill Level</td><td>Work Time, Trouble-shooting</td><td>Work Time plus Supervision, Trouble-shooting</td><td>Work Time, Corrosion Control</td><td>Work Time plus Supervision, Corrosion Control</td></tr>
<tr><td>E1-E3, 3</td><td>2.1</td><td>2.4</td><td>1.3</td><td>1.5</td></tr>
<tr><td>E3, 5</td><td>1.7</td><td>2.1</td><td>1.2</td><td>1.3</td></tr>
<tr><td>E4, 5</td><td>1.6</td><td>1.8</td><td>1.1</td><td>1.1</td></tr>
<tr><td>E5, 5</td><td>1.4</td><td>1.5</td><td>1.1</td><td>1.1</td></tr>
<tr><td>E5, 7</td><td>1.2</td><td>1.3</td><td>1.0</td><td>1.0</td></tr>
<tr><td>E6-E7, 7</td><td>1.0</td><td>1.0</td><td>1.0</td><td>1.0</td></tr>
<tr><td>Title</td><td colspan="5">"The Effects of Manning and Crew Stability on the Material Conditions of Ships"</td></tr>
<tr><td>Author</td><td colspan="5">Russell Beland and Aline Quester</td></tr>
<tr><td>Date</td><td colspan="5">1991</td></tr>
<tr><td>Method</td><td colspan="5">Survey of the percentage of time free of serious failures aboard a ship; collection of data on the manning levels of ships in comparison with their time spent free of operational problems. The model estimates the percentage of ships free of serious failures (PCTFREE). They first compute this at the mean of all variables in the sample and then look at how changes in one standard deviation up or down from the mean affect the estimate.</td></tr>
<tr><td>Functional Form</td><td colspan="5">Uses Tobit model. Variables included in the analysis are manpower requirements (a measure of ship's enlisted manning relative to requirements that includes a measure of the experience mix of personnel), new crew (percentage of the enlisted crew that was not assigned to the ship three months earlier), tenure time of the commanding officer (CO) in months, an "in-stock" variable (fraction of all parts requests in one month that the ship stockroom was able to fill), steam hours underway for month (in hundreds), months since last overhaul. Outcome variable was percentage of time free of serious failure (ranges from 0 to 100 percent).</td></tr>
</table>

Summary Findings	Finds a significant correlation between CO tenure and the material condition of the ships. For example, moving from one standard deviation below the average CO tenure to one standard deviation above average (from 6 to 21 months) increases the time free of mission-degrading failures by about 5 percentage points. Also finds that manning levels, types of personnel, and crew rotation are correlated with ship condition. The analysis indicates that an increase in the percentage of new crewmembers is correlated with an increase in the number of ships with material condition problems.			
Quantitative Results	Resource Level	KNOX (FF-1052s) N=599	SPRUANCE (DD-963s) N=491	ADAMS (DDG-2s) N=351
	Predicted PCTFREE for means of all variables	70.49	69.92	51.01
	Chi-square of Tobit estimation (degrees of freedom)	72.0 (10)	128.1 (10)	110.4 (10)
	CHANGES FROM OVERALL MEAN PREDICTION	*** significant at .005 level ** significant at .05 level * significant at .10 level		
	MANREQ (manning levels and crew experience)			
	One SD above mean	76.36 (***)	82.50 (***)	63.16 (***)
	One SD below mean	64.06 (***)	54.46 (***)	37.78 (***)
	NEWCREW			
	One SD above mean	66.18 (***)	62.82 (*)	45.74 (***)
	One SD below mean	74.44 (***)	76.27 (*)	56.26 (***)
	TIME_CO (tenure of CO)			
	One SD above mean	72.90 (*)	___ (*)	68.76 (***)
	One SD below mean	68.01 (*)	___ (*)	33.04 (***)

Title	*Personnel Substitution and Navy Aviation Readiness*
Author	A. J. Marcus
Date	1982

Method	Surveys several naval squadrons and looks at multiple characteristics of the personnel including number of high school graduates/nongraduates, number in pay grade, number by years of experience, number by training completed, number by tenure in the squadron, number by occupational group
Functional Form	$Q=a_1x_1 + a_2x_2 + a_3x_3 + b_{12}(x_1x_2)^{1/2} + b_{13}(x_1x_3)^{1/2} + b_{23}(x_2x_3)^{1/2} + b_{14}(x_1p)^{1/2} + b_{24}(x_2p)^{1/2} + b_{34}(x_3p)^{1/2}$ x_1 = Personnel in grades E1-E3 x_2 = Personnel in grades E4-E6 x_3 = Personnel in grades E7-E9 p = Average number of operating aircraft Q= Number of flights/MCR
Summary Findings	Looks at substitutability of first-term and career personnel and how the manpower mix affects readiness. Personnel were grouped into three pay grade levels: E1-E3, E4-E6, E7-E9. The findings are consistent with the expectation that output will increase at higher pay grades. The study also finds a high marginal productivity for the most senior personnel. However, the author also notes that this could be because of their relatively small number. The author compares the cost of the current force with the least-cost force. He finds that a more heavily senior force would lead to cost savings for the government (true of the force in 1982, when the article was written). The study also looks at the relationship between education/AFQT and performance. The author finds that higher levels of education are associated with higher levels of performance, but that there does not seem to be a clear and stable relationship between AFQT score and performance. Finally, the research suggests that personnel in upper pay grades appear to be more productive than those in lower pay grades and that personnel in E1-E3 pay grades may be supplements to E4-E6, but that E7-E9 are complements for both.
Quanti-tative Results	RESULTS N= 292 Squadrons, each with approx. 230 enlisted personnel (Total: 67,160)

Marginal Products of Pay Grade Groups

	E1-E3	E4-E6	E7-E9
Flights	-1.2 (29.1)	2.9 (1.6)	30.7 (9.4)
Mission capable rate	.08 (.08)	.15 (.10)	.72 (.40)

Marginal Products of Pay Grade Groups, Different Groupings

	E1-E4	E5-E6	E7-E9
Flights	-.5	6.2	29.1
Mission capable rate	.046	.339	.342

Marginal Products of Experience Groups

	1-4 YOS	5-8 YOS	9+ YOS
Flights	1.3	-2.8	12.5
Mission capable rate	.01	.12	.44
Marginal Products of Educational Groups			
	No HS/ GED	HS Graduate	HS+
Flights	-.5	1.9	10.6
MCR	.3	.06	-.04
Coefficient Estimates: Performance on Pay Grade			
	E1-E3	E4-E6	E7-E9
Flights	-59.1 (27.0)	-49.4 (34.8)	43.1 (68.8)
MCR	-3.16 (1.26)	-4.06 (1.62)	-1.91 (3.21)
Coefficient Estimates: Performance on Experience			
	1-4 YOS	5-8 YOS	9+ YOS
Flights	-60.5 (29.4)	-16.8 (28.9)	-18.6 (54.6)
Mission capable rate	-3.62 (1.30)	3.58 (1.28)	-1.13 (2.42)
Coefficient Estimates: Performance on Education			
	No HS/GED	HS Graduate	HS+
Flights	40.9 (74.7)	7.4 (29.5)	16.9 (61.5)
Mission capable rate	4.07 (3.34)	2.19 (1.32)	-1.99 (2.75)

Title	"The Economics of Military Manpower"
Author	J. T. Warner and Beth Asch
Date	1995
Method	Offers a general survey of previous literature and studies on the responses of military manpower to pay, training, other incentives, the opinions of others, bonuses. Offers a table of supply elasticity estimates for military personnel to different factors as reported in studies over the past 20 years.
Functional Form	NA
Summary Findings	Summarizes the main economic principles and theories governing the supply of military manpower: • Assuming tastes for military service are normally distributed, enlistment exhibits an S-shaped relationship with pay level, i.e., enlistments are less responsive to pay when pay is either extremely high or extremely low. • Enlistment can be affected by the opportunity to receive transferable skill training • Looking specifically at enlistment trends among high-quality soldiers, relative military-civilian pay levels have a significant effect, with relative

	pay elasticities ranging from .15 to 1.89, with central tendency of about .5 to 1.0. Finally, high-quality recruits are affected more by educational incentives than by enlistment bonuses (increasing in educational incentives increased enlistment by 9 percent as compared to 5 percent for an increase in enlistment bonuses). • Decisions to reenlist must include considerations of civilian and military future pay streams, potential for retirement benefits with each decision, personal preference and discount rate of future income. • A summary of productivity literature suggests that careerists (those above E4) are significantly more effective than E1-E3 personnel (many studies estimate that careerist are twice as effective as early-term personnel). • The authors discuss the arguments for and against an all-volunteer force (AVF) as well as the comparative costs of an AVF and a conscript- based force. They suggest that the opportunity cost of conscription and the lower productivity/quality of conscripted soldiers when combined with the smaller force size requirement when professional soldiers are used, are likely to offset the higher wages required with an AVF. • A final issue raised is that of the differences/similarities between the retention, recruitment, and cost of female soldiers.
Quantitative Results	NA
Title	"The Economic Theory of a Military Draft Reconsidered"
Author	John Warner and Beth Asch
Date	1996
Method	Cost comparison of all-volunteer force and conscription-based force including opportunity costs and productivity effects.
Functional Form	NA
Summary Findings	The authors note that the true cost of building an AVF depends not only on the monetary cost of paying high-wage, high-quality soldiers, but also on the opportunity cost incurred by a draft, the increased productivity of higher-wage soldiers, and cost savings of more effective performance by volunteer soldiers.
Quantitative Results	NA
Title	*A New Approach for Modeling Ship Readiness*

Author	Laura Junor and Jessica Oi
Date	1996
Method	Survey using historical data for nearly every ship in the Navy, on a quarterly basis from 1978 to 1994. Uses the SORT (Status of Resources and Training Systems) model, which looks at the relationship among personnel factors, supply factors, equipment factors, and training factors and the amount of time a ship spends "out of commission" in a given quarter.
Functional Form	Tobit regression analysis. Model considers personnel quality and manning levels as inputs to all resources areas. Supply is an input to training and equipment condition. Failure rate is an input into supply, repair, and equipment condition. Repair rate is an input into equipment, and equipment condition is an input to training. Factors considered in each variable: Personnel = P(manning, personnel quality (index considering high-school degree, AFQT scores, length of service, pay grade for entire crew), deployed status, steaming (days underway per quarter), ship class differences, crew turnover, manpower costs) Supply = S(retail inventory, equipment failure rate, manning, personnel quality, ship class differences, deployed status) Equipment Failure Rate = F(steaming, overhauls, manning, personnel quality, deployment cycle, classes differences, decommissioning) Repair Rate = R(manning, personnel quality, supplies, number of failures, ship age, deployment status, ship class differences) Equipment Condition = E(failure rate, mean time to correct failure, deployment status, ship class differences, decommissioning, ship age in months, scheduled overhaul, modernization costs) Training= T(personnel quality, manning, supply, equipment modernization, ordinance or electrical equipment repairs/improvements, deployment status, ship class differences)
Summary Findings	Finds that personnel quality strongly affects all aspects of readiness, including equipment, maintenance, training, and supply. In fact, manning levels and personnel quality

are the only two variables that are significant in all resource areas. Looking more specifically at personnel variables, the study finds personnel turnover has only a small effect on crew readiness. Higher personnel quality is found to decrease the number of new equipment casualties and to decrease maintenance time. Personnel quality is also found to have a positive effect on the results of training. The effect of having a more effectively trained force is also demonstrated by substituting 1994 crews for 1981 crews and looking at the difference in predicted readiness. Readiness was significantly increased with this substitution, particularly in the personnel category, but also in supply readiness. Finally, the substitution led to a decrease in maintenance time. The opposite substitution of 1981 crew into 1994 readiness structure leads to the opposite result, namely, a decrease in personnel, supply, training, and equipment readiness and an increase in maintenance time.

Quantitative Results

Percentage of Time in C1 (serious failure) for Personnel Reasons, N=5446

Variable	Tobit Coeff.	Significance Level
Personnel quality	.135	At least 5%
Manning	.031	At least 5%
Crew turnover	.028	At least 5%
Days underway	.001	At least 5%
Deployed status	.137	At least 5%
Time	.709	At least 5%

Percentage of Time in C1 for Supply Reasons, N=5664

Variable	Tobit Coeff.	Significance Level
Personnel quality	.032	At least 5%
Manning	.007	At least 5%
Repair parts	3.12E-7	At least 5%
Repair parts deployed	3.82E-7	At least 5%
Weapons procurement	2.71E-7	At least 5%
Weapons procurement deployed	1.47E-7	At least 5%
Gross effectiveness	.004	At least 5%
Deployed status	.049	

Percentage of Time in C1 for Equipment, N=5664

Variable	Poisson Coeff.	Significance Level
Mean time to correct CASREPS	-.007	At least 5%
Mean time to correct CASREPS-deployed	-.002	At least 5%
Percent of time in C1 for supply	1.1	At least 5%

	Deployed status	-.172	At least 5%
	Approach of decommissioning	-.073	At least 5%
	Mean Maintenance Time to Correct a Casualty, N=5664		
	Variable	OLS Coeff.	Signif. Level
	Personnel quality	-1.024	At least 5%
	Manning	-.043	
	Crew turnover	.222	At least 5%
	Crew turnover-deployed	-.155	
	Repair parts	-8.88E-7	
	Repairables	-6.39E-7	At least 5%
	Cost last year for modernization	4.08E-8	At least 5%
	Ship age	.035	At least 5%
	Deployed status	-3.848	At least 5%
	Percentage of Time Spent in C1 for Training, N=5664		
	Variable	Tobit Coeff.	Signif. Level
	Personnel quality	2.95E-2	At least 10%
	Percent of time in C1 for supply	1.7143	At least 5%
	Percent of time in C1 for equipment	1.2268	At least 5%
	Manning-deployed	7.62E-3	At least 5%
	Quarters since ship deployed	-5.51E-2	At least 5%
	Deployed Status	-.66413	At least 5%
	Days underway over past year	6.17E-3	At least 5%

Title	*Youth vs. Experience in the Enlisted Air Force: Productivity Estimates and Policy Analysis*
Author	Mary Anne Doyle
Date	1998
Method	Activity analysis (linear programming) that determines the amounts of different types of personnel required to complete a given set of tasks. The technique can identify different experience mixes and manning levels to accomplish a given workload.
Functional Form	NA
Summary Findings	Finds that if an experienced unit is expected to complete the same amount of work in the same period of time as a less experienced unit, the size of the less experienced unit must be increased. For example, when comparing a unit split evenly between first-termers and careerists to one with 40 percent first-term personnel and 60 percent careerists, Doyle finds that the less experienced unit requires 3 percent more time to accomplish the assigned work. A unit split 60-40 between first-term and career personnel will take 5 percent longer to complete the task than the 40-60 split unit. The relative productivities of

	first-term and career personnel vary, however, based on the difficulty of the task. The most significant time savings for total unit work time can be reduced if the most experienced personnel are assigned less supervisory duty and are given more of the most challenging work. The author suggests that manpower requirements for a given unit should take the experience mix into account.

Quantitative Results	First-Term/ Career Mix	Unit Size	Number of Personnel 1-4 YOS	Number of Personnel 5-8 YOS	Number of Personnel 9-12 YOS	Number of Personnel 13+YOS
	30-70	95	30	20	15	30
	40-60	97	38	21	13	25
	50-50	200	50	18	12	20
	60-40	102	60	16	10	16
	70-30	105	75	10	8	12

Title	*Manpower and Personnel IWAR 2000: Aging the Force*
Author	Carol Moore, Heidi Golding, and Henry Griffis
Date	2001
Method	Simulates the effect of various retention rates on the Navy's steady-state accession level. Analysis considers the effects of changing the experience mix of the force on the cost of the force and target retention and accession rates.
Functional Form	In the model, the cost of new recruits is equal to the recruiting cost, the salaries of instructors, the costs associated with Permanent Change of Stations (PCS), and the costs of paying students with Immediate Active Duty status who are also in school. The costs of retaining senior personnel include reenlistment bonuses, medical and retirement plan accruals for the personnel induced to stay, and higher salaries due to seniority. The reenlistment bonus makes up the majority of this cost and is defined as a range because these bonuses can vary in size.
Summary Findings	Raising reenlistment targets is not an effective way to meet end-strength goals because the cost of retaining senior personnel exceeds that of hiring and training new recruits. According to the estimates used in this study, the cost of meeting end-strength goals by raising Zone A reenlistment by two points would be between $78 million and $169 million per year whereas the cost savings from lower accessions would be only $36 million per year. Looking at different skill level occupations, the authors find that increasing reenlistments makes more sense for high-skill occupations than those with low skill requirements. Productivity gains are also important inputs and offset some of the seniority and reenlistment costs.

<table>
<tr><td rowspan="22">Quanti-
tative
Results</td><td colspan="4">Example experience mix change: Increase Zone A reenlistment rate by 2 percentage points</td></tr>
<tr><td colspan="2">Benefits
(Reductions in Cost)</td><td colspan="2">Costs</td></tr>
<tr><td>Recruiting</td><td>$14.7 million</td><td>Reenlistment bonuses</td><td>$66 to $157 million</td></tr>
<tr><td>Instructors</td><td>$1 million</td><td>Medical</td><td>$3.9 million</td></tr>
<tr><td>Student IA</td><td>$19.3 million</td><td>Seniority pay</td><td>$7.9 million</td></tr>
<tr><td>PCS</td><td>$1.2 million</td><td></td><td></td></tr>
<tr><td>Average YOS</td><td>Increase 1.2 months</td><td></td><td></td></tr>
<tr><td>Readiness</td><td>?</td><td></td><td></td></tr>
<tr><td>Total</td><td>$36 million per year plus readiness</td><td>Total</td><td>$78 to $169 million per year</td></tr>
<tr><td colspan="4"></td></tr>
<tr><td colspan="2">Baseline</td><td colspan="2">With increased Zone A reenlistment</td></tr>
<tr><td>Steady state accessions</td><td>56,140</td><td>Steady state accessions</td><td>54,950</td></tr>
<tr><td>Zone A reenlistment rate</td><td>60.7%</td><td>Zone A reenlistment rate</td><td>62.7%</td></tr>
<tr><td>Number of reenlistments</td><td>20,640</td><td>Number of reenlistments</td><td>21,500</td></tr>
<tr><td>Average length of service</td><td>6.0 years</td><td>Average length of service</td><td>6.1 years</td></tr>
<tr><td></td><td></td><td>Cut in accessions</td><td>1,190</td></tr>
<tr><td></td><td></td><td>Increase in reenlistments</td><td>860</td></tr>
<tr><td colspan="4">Example experience mix change: Age only certain skills, enough for 100 accession cuts</td></tr>
<tr><td></td><td>Retention and Seniority Costs ($ millions)</td><td>Recruiting and Training Savings ($ millions)</td><td>Annual Increase in Productivity (%)</td></tr>
<tr><td>High-tech sample</td><td>6.8</td><td>5.0</td><td>3.6</td></tr>
<tr><td>Mid-tech sample</td><td>6.0</td><td>3.5</td><td>3.3</td></tr>
<tr><td>Low-tech sample</td><td>7.8</td><td>2.7</td><td>2.6</td></tr>
<tr><td colspan="5"></td></tr>
<tr><td>Title</td><td colspan="4">Setting Requirements for Maintenance Manpower in the U.S. Air Force</td></tr>
</table>

Author	Carl Dahlman, Robert Kerchner, and David Thaler
Date	2002
Method	Simulation using Logistics Composite Model, which is used by the U.S. Air Force to estimate the man-hours needed to accomplish direct maintenance tasks. The model uses manpower standards and policies to derive requirements for manpower spaces. Spaces are then authorized on the basis of fiscal guidance. The objective of the model is to minimize the manpower needed while still generating the required sortie production and necessary training. The model classifies workers according to skill level: 3-, 3 middle, 3+, 5-, 5 middle, 5+, 7, each with an efficiency based on their ability to perform tasks relative to a 7-level. The procedure involves optimizing H with a given manpower level to yield a set of work distributions. If shortfalls exist, the result is recorded and the process repeated. NOTE: The report does not specify the number of trials, although it appears from their language and model that the authors optimize the function at each force value only once. This would make sense if the model yielded the same results when the same parameters were entered.
Functional Form	NA
Summary Findings	Addresses two key issues: (1) Does the existing manpower system underestimate the workload requirements of maintenance personnel? (2) What are the implications of a misaligned experience mix? First, the authors find that the existing system does underestimate work hours. They argue that the system pays more attention to operational concerns (actual maintenance activities) than to training activities. Any manpower system should take into account all requirements placed on personnel. Next, the authors turn to the implications of what they term the "experience shortfall," which is the result of the development of a more heavily senior force in the mid-1990s and the low retention rates for second-term personnel. As senior and mid-level personnel have chosen to leave the force, the structure has become more heavily filled with junior personnel who do not have the skills to replace the lost senior personnel, thus reducing skill base of the unit in the long run. The authors argue that the problem is even more insidious: The existing experience shortage is embedded in the force because the loss of skilled personnel also means the loss of experienced trainers. Therefore, the newly enlisted men are not given the same quality of training, in terms of the trainee-to-trainer ratio and the actual knowledge of the instructor. The authors recommend that any solution to the problem will require time and suggest several steps--namely, the

	development of more accurate manpower estimates that include the important need to rejuvenate human capital, to reassess current fill rates and experience mix, and to increase the emphasis on retention of mid-level personnel.
Quantitative Results	Manpower Shortfall

No.	Category	3-	3 mid	3+	5-	5 mid	5+	7	Shortfall
1368	Teach	0.0	0.0	.06	.06	.06	.06	.12	-19.9
	Production	.18	.23	.25	.38	.43	.48	.56	-11.2
	Other	.06	.06	.03	.12	.12	.12	.13	-11.1
1368 optimal	Teach	0	0	0	.07	.06	.06	.07	-37.6
	Production	.17	.22	.27	.38	.49	.60	.63	0
	Other	.07	.07	.07	.12	.07	.01	.10	-35.6
1440	Teach	0	0	0	.07	.07	.07	.10	-23.8
	Production	.17	.22	.27	.35	.45	.56	.61	0
	Other	.07	.07	.07	.14	.09	.03	.10	-23.9
1520	Teach	0	0	0	.08	.08	.08	.12	-11.0
	Production	.17	.22	.27	.32	.42	.52	.58	0
	Other	.07	.07	.07	.16	.11	.06	.10	-10.9
1592	Teach	0	0	0	.09	.09	.09	.14	-.1
	Production	.17	.22	.27	.30	.40	.49	.56	0
	Other	.07	.07	.07	.18	.13	.08	.10	0

STUDIES ON TRAINING AND PERFORMANCE

Title	*Modeling the Contribution of Maintenance Manpower to Readiness and Stability*
Author	Glenn Gotz and Richard Stanton
Date	1986
Method	Simulation using a model based on real combat experience to determine the effect of different training and manpower mixes on the readiness of ships during wartime. Computer simulation that uses AFQT, training type, and different time of repair parameters to determine readiness levels.

Functional Form	NA
Summary Findings	Finds that cross-training (when technicians are trained to repair more than one type of part) improves unit performance significantly, especially when the failure rate for one part is above the failure rate for the other. This is because the increased skill base of these individuals allows them to be used more flexibly and increases their value/contribution to the group. The study also considers the effect of high-skill personnel on the number of aircraft that are unusable for maintenance reasons. The authors use task time as a measure of skill level. They report that the introduction of high-skill personnel into the manpower mix decreased the NA aircraft (number of aircraft unavailable due to maintenance problems, particularly in the middle days of the observation period).
Quantitative Results	Base Case, no cross-training, no repairman substitution. Days it takes for repairmen to fix parts vary as does the probability of failure for each part. N=100

A/1/NS: Repairmen (2 each)	Part 1	Part 2	
I	.8 days	----------	
II	-------	.8 days	
Failure rate	.042	.042	
DAY	MEAN (NA aircraft)	Max.	Std. Deviation
1	1.81	5	1.23
2	3.20	7	1.67
3	4.14	9	1.98
4	4.55	12	2.26
5	4.78	12	2.58
6	4.98	13	2.75
7	5.13	16	3.20
8	5.01	15	3.07
9	4.7	14	3.23
10	4.31	15	3.08
11	4.24	13	2.87
12	3.92	13	2.65
13	3.26	12	2.49
A/2/NS: Repairmen (2 each)	Part 1	Part 2	
I	1.067 days	---------	
II	-------	.8 days	
Failure rate	.052	.042	
DAY	MEAN (NA aircraft)	Max.	Std. Deviation

1	2.34	7	1.58
2	4.09	10	1.84
3	5.45	13	2.31
4	6.62	15	2.9
5	7.49	17	3.09
6	8.15	20	3.61
7	8.71	20	4.16
8	9.05	18	4.21
9	9.11	20	4.25
10	9.47	23	4.34
11	9.37	23	4.32
12	9.07	24	4.43
13	8.39	22	4.16
A/2/NS: Repairmen (2 each)	Part 1	Part 2	
I	1.067 days	---------	
II	-------	.8 days	
Failure rate	.052	.042	
DAY	MEAN (NA aircraft)	Max.	Std. Deviation
1	2.34	7	1.58
2	4.09	10	1.84
3	5.45	13	2.31
4	6.62	15	2.9
5	7.49	17	3.09
6	8.15	20	3.61
7	8.71	20	4.16
8	9.05	18	4.21
9	9.11	20	4.25
10	9.47	23	4.34
11	9.37	23	4.32
12	9.07	24	4.43
13	8.39	22	4.16
Cross-Training, Repairmen can fix both parts, though at different rates, manpower staffing decisions made by "minimize back orders rule." N=100			
A/1/MB: Repairmen (2 each)	Part 1	Part 2	
I	.8 days	1.2 days	
II	1.2 days	.8 days	
Failure rate	.042	.042	
DAY	MEAN (NA aircraft)	Max.	Std. Deviation
1	1.67	5	1.14
2	2.8	8	1.48
3	3.47	7	1.45

4	4.03	10	1.99
5	4.29	11	2.21
6	4.7	12	2.59
7	4.74	12	2.69
8	4.84	12	2.63
9	4.83	13	2.6
10	4.61	14	2.44
11	4.23	15	2.66
12	4.01	15	2.61
13	3.63	12	2.55
A/2/MB: Repairmen (2 each)	Part 1	Part 2	
I	1.067 days	1.2 days	
II	1.2 days	.80 days	
Failure rate	.052	.042	
DAY	MEAN (NA aircraft)	Max.	Std. Deviation
1	2.18	7	1.18
2	3.75	9	2.41
3	4.7	11	3.33
4	5.97	13	3.53
5	6.75	14	4.48
6	7.09	14	4.86
7	7.48	17	5.1
8	7.19	17	5.41
9	7.25	16	5.36
10	7.08	16	5.48
11	6.55	16	5.41
12	6.4	15	5.29
13	6.24	13	4.92
Multiple Skill Levels and Cross-training, N=100			
B/1MB: Repairmen	Part 1	Part 2	
High skill I (1)	.8 days	1.2 days	
Low skill I (2)	1.2 days	--------	
High skill II (1)	1.2 days	.8 days	
Low skill II (2)	--------	1.2 days	
Failure rate	.042	.042	
DAY	MEAN (NA aircraft)	Max.	Std. Deviation
1	1.75	5	1.16
2	3.06	9	1.7
3	3.6	9	1.81
4	3.68	9	1.94
5	3.98	9	1.87

6	3.93	10	1.97
7	3.64	14	2.02
8	3.67	9	1.93
9	3.51	11	1.97
10	3.21	8	1.66
11	2.88	9	1.64
12	2.76	7	1.46
13	2.47	7	1.62
B/2/MB: Repairmen	Part 1	Part 2	
High skill I (1)	1.067 days	1.2 days	
Low skill I (2)	1.2 days	--------	
High skill II (1)	1.2 days	1.067 days	
Low skill II (2)	-------	1.2 days	
Failure rate	.052	.042	
DAY	MEAN (NA aircraft)	Max.	Std. Deviation
1	2.14	7	1.29
2	3.63	8	1.71
3	4.7	11	2.14
4	5.32	11	2.50
5	5.52	12	2.49
6	5.73	13	2.68
7	5.96	15	3.06
8	5.5	17	3.02
9	5.06	17	3.08
10	4.41	15	2.69
11	4.44	15	2.81
12	4.1	10	2.33
13	3.88	12	2.23

Title	*Aircraft Maintenance Task Allocation Alternatives: Exploratory Analysis*
Author	S. C. Moore, Edwin Wilson, and Edward Boyle
Date	1987
Method	Activity analysis (linear programming) that determines the amounts of different types of personnel required to complete a given set of tasks. The technique can identify different experience mixes and manning levels that can accomplish a given workload.
Functional Form	NA
Summary Findings	Consolidating specialties would force each airman to receive training and become proficient in a wider range of skills. The authors note that combining specialties reduces manpower required to maintain a given set of aircraft and increases manpower utilization. If individuals have a more extensive set of skills, they can contribute to many different maintenance activities. This increases the utilization of these individuals and reduces

	the need for additional person with more limited skills. These observations suggest that additional training and acquisition of new skills can significantly raise the flexibility given to manpower planners and allow the force to perform with fewer personnel. However, combining specialties would also lead to increased training costs and time and would place a larger burden on senior personnel responsible for conducting training. The increased amount of time devoted to training would decrease productive working time, particularly for first-term personnel who make up a large portion of the military, and offset some of the advantages gained from a combined specialty approach.

Quantitative Results	Number of Specialties	Manpower Requirements	Percent Manpower Utilization	Average Training Days
	Main Operating Base, 72 Aircraft			
	1	69	87	900
	3	73	78	300
	5	76	76	200
	7	90	69	60
	10	100	60	50
	4 Dispersed Operating Locations, 18 Aircraft Each			
	1	84	71	-
	3	103	53	-
	5	135	42	-
	7	160	39	-
	10	200	30	-

Title	*Flying Hours and Crew Performance*
Author	Colin Hammon and Stanley Horowitz
Date	1990
Method	Controlled trials of three types of "exercises": 1. Simulation that rated pilots flying F14s and A7s in Carrier Air Wing 7 between June 1985 and October 1987 on their carrier landings (on a seven-point scale [0, 1, 2, 2.5, 3, 4, 5]). Results were compared to pilot experience, career flying hours, and recent flying hours. 2. Simulation of Marine Corps bombing exercises. 3. Simulation of air-to-air combat exercises, which rated participants on whether they shot the target and at what range.
Functional Form	Carrier landings: $\mathrm{Log}\{p(s)/[1-p(s)]\}= a_0 + a_1{*}H_c + a_2{*}H_{30} + a_3{*}N + a_4{*}F$ p(s)= probability of success, a landing grade of either 3.0 or 4.0

	H_c= career flying hours H_{30}= flying hours in previous month N= dummy variable for night flights (1 if yes, 0 otherwise) F= a dummy variable for type of flight, 1 for F-14 and 0 for A7 Bombing exercise: $M= b_0 + b_1*H_c + b_2*H_7 + b_3*A + b_4*H_c + b_5*H_7 + b_6*AV8 + b_7*F4$ M= bombing accuracy as measured by the distance by which the bomb misses its target (in feet) H_c= career flying hours H_7= flying hours in the past 7 days A= a dummy variable for delivery type, 1 for automatic and 0 for manual AV8= a dummy variable taking the value 1 for an AV8B flight and 0 otherwise F4= a dummy variable taking the value 1 for an F-4S flight and 0 otherwise Air-to-air combat: $Ln\ (p_0/p_1) = a_{i0} + a_{i1}*H_{pc} + a_{i2}*R_1 + a_{i3}*R_t + a_{i4}*O_r + a_{i5}*E_{adv} + a_{i6}*S_r$ $R_1= b_0 + b_1*H_{pc} + b_2*H_{rc} + b_3*P_{30} + b_4*H_{r30}$ $R_t= c_0 + c_1*H_{pc} + c_2*H_{rc} + c_3*P_{30} + c_4*H_{r30}$ P_i = probability of achieving the i^{th} outcome R_1= difference between the range at which the crew begins the exercise and the range at which radar lock-on is made R_t= range at which the red aircraft is sighted H_{pc}= pilot's career flight hours H_{rc}= radar intercept officer's career flight hours H_{p30}= pilot's flight hours in the previous month H_{r30}= radar intercept officer's flight hours in the previous month O_r= ratio of red to blue aircraft when the shot is fired E_{adv}= 1 for competitive exercise or more than two blue aircraft and 0 otherwise.
Summary Findings	Looks at quantitative relationship between how much aircrews have flown (over their career and over a more recent time period) and their performance on three tasks--carrier landings, Marine bombing, and air-to-air combat. Finds that career experience has a greater correlation with performance than does recent experience. The authors hypothesize that this occurs because more recent training helps to hone skills and career flight time promotes mastery. For the landing portion of the experiment they find that a 10 percent decrease in the number of recent

flying hours would have the short-term effect of decreasing the number of unsatisfactory landings by 2.6 percent and decreasing the number of excellent landings by 2.5 percent. A career decrease of 10 percent in the number of hours flown would lead to an increase of 6.9 percent in the number of unsatisfactory landings and a decrease of 2.4 percent in satisfactory landings. For the Marine Corps bombing exercise, the authors find that an increase in flying hours is associated with an improvement in performance. If flying hours were reduced 10 percent for a short period of time, the average miss distance would rise by about 1 percent for manual bomb deliveries. If the reduction is continued indefinitely, a further reduction of more than 1 percent would be incurred. The majority of this effect is believed to act through its effect on total pilot experience. Finally, in the air-to-air combat exercise, the study finds that both short-term and career experience is associated with targeting effectiveness and likelihood of kills. A 10 percent decrease in all experience variables leads to a decrease of 4.8 percent in the probability that the soldier will kill the enemy, and an increase of 9.2 percent that the soldier will be killed. Again, career experience had a more significant effect than recent flight time. The report concludes that the optimal level of training will balance these increases in performance with the costs of training and the potential cost of equipment replacement if less effective training leads to worse performance.

Quantitative Results

Coefficients and Std. Errors of Probability of Meeting Landing Grade Criteria for A-7 aircraft (** significant to .99 level) N=4351

	Satisfactory	Excellent
	Coeff. (Std. Error)	Coeff. (Std. Error)
Constant	1.34 (.116) **	-1.32 (.0087)**
Career flying hours	.0005 (5.5E-5) **	.00024 (2.8E-5)**
Flying hours in previous month	.013 (.004) **	.018 (.003)**
Night landing	-.619 (.097) **	.065 (.075)

Determinants of bombing accuracy for Marine Corps aircraft (miss distance in ft)
*** significant at .99 level
** significant at .95 level N=649

Independent variable	Coefficient	Std. Error
Constant	113.4	11.23 ***
Career flying hours	-.0094	.004**
Flying hours in last 7 days	-2.65	1.28**

Automated delivery	-64.61	11.5***
AV-8B flight	20.96	6.87***
F-4S flight	46.78	10.24***

Determinants of targeting effectiveness,
*** significant at .99 level
N=1352

Independent variable	Lock Range Delta Coeff.(Std. Error)	Tally-ho Range Coeff. (Std. Error)
Constant	2.74E1 (.96)***	-1.26 (.525)***
Pilot career flying hours		5.57E-4 (8.79E-5)***
Radar Intercept Officer (RIO) career flying hours		9.56E-4 (9.85E-5)***
Pilot flight hours previous month	-9.91E-2 (.035)***	1.59E-1 (.016)***
RIO flight hours previous month	(.037)*** 2.06E-2	(.018) -1.64E-1

Full Effects of Flying Hour Variables on Performance in Air-to-Air Combat
*** significant at .99 level
**significant at .95 level N=1352

Independent variable	Red Hits Blue, Coeff (Std. Error)	Blue Hits Red, Coeff (Std. Error)
Pilot career flying hours	-2.79E-5 (5.0E-6)***	4.66E-5 (1.25E-5)***
RIO career flying hours	-3.97E-6 (1.5E-6)***	1.77E-5 (4.2E-6)***
Pilot flight hours in previous month	-8.57E-4 (2.5E-4)***	3.43E-3 (7.3E-4)***
RIO flight hours in previous month	-4.18E-4 (1.5E-4)***	1.22E-3 (5E-4)**

Title	*Relating Flying Hours to Aircrew Performance: Evidence for Attack and Transport Missions*
Author	Colin Hammon and Stanley Horowitz
Date	1992
Method	Controlled trials and simulation similar to data and analysis above, but focuses on the Marine bombing exercise and an additional Air Force tactical drop exercise. Extends the original simulation by including simulator hours and other independent variables and considering more than one model.
Functional Form	Bombing Accuracy: $LnCE = b_0 + b_1 * LnH_c * M + b_2 * LnH_c * A + b_3 * LnH_c * C + b_4 * LnH_c * (A+C) +$

	$b_5*LnH_c*M + b_6*LnH_{7s}*(1-R) + b_7*A*F18 + b_8*C*F18 + b_9*M*F18 + b_{10}*A*AV8 + b_{aa}*C*AV8 + b_{12}*M*AV8 + b_{13}*R + b_{14}*B_{76} + b_{15}*L$ Ln=natural log CE= miss distance (circular error), the distance in feet by which the bomb misses the target (CE is the median for a series of bombing runs) H_c= career flying hours H_{cs}= career flight simulator hours F_7= flights in the previous 7 days H_{7s}= flight simulator hours in the previous 7 days A= dummy variable taking the value 1 for automatic deliveries and 0 otherwise C= dummy variable taking the value 1 for CCIP deliveries and 0 otherwise M= a dummy variable taking the value 1 for manual delivery and 0 otherwise AV8= a dummy variable taking the value 1 for an AV-8 flight and 0 otherwise F18= a dummy variable taking the value 1 for an F/A-18 flight and 0 otherwise R= a dummy variable taking 1 for FRPs and 0 for fleet pilots B_{76}= a dummy variable taking the value 1 more Mk-76 practice bombs and 0 otherwise L= a dummy variable taking the value 1 for loft deliveries and 0 otherwise $LnCE= b_0 + b_1*H_{cpt} + b_2*H_{cpst} + b_3*H_{cp60} + b_4*H_{nt} + b_5*H_{nst} + b_6*H_{n60} + b_7*N + b_8*D_{he} + b_9*D_{tb} + b_{10}*D_{pers}$ Ln= natural log CE= drop accuracy, circular error, the distance in yards by which the parachute misses the target H_{cpt}= copilot career flying hours H_{cpst}= copilot career simulator hours H_{cp60}= copilot flying hours in past 60 days H_{nt}= navigator career flying hours H_{nst}= navigator career simulator hours H_{n60}= navigator flying hours in past 60 days N= dummy variable for the time of drop, 1 for night drop and 0 otherwise D_{he}= dummy variable with a value of 1 for heavy equipment drop and 0 otherwise D_{tb}= dummy variable with a value of 1 for training bundle drop and 0 otherwise D_{pers}= a dummy variable with a value of 1 for personnel drop and 0 otherwise
Summary Findings	Repeats many of the observations made in the previous report but expands the depth of the analysis. Considers

Marine bombing and tactical air drop and includes the effectiveness of a simulator as a training tool as one of its variables. The general finding is that experience and training are correlated with performance. The authors note that for both exercises, long-term career flight hours have a more significant effect on performance than the short-term variable. For the Marine bombing task, the use of the simulator has a high initial effect but it decreases after the first 1/4 hour or so. The simulator therefore does have an effect on performance and can substitute somewhat for experience. In the case of the marine bombing exercise, the marginal partial effect is greater for simulator hours than for airtime hours (simulators are also less expensive and risky for the equipment). For the tactical drop exercise, the authors find that a decrease in the amount of actual flight time has a smaller effect on performance than an identical reduction in simulator flight time.

Quantitative Results

Determinants of Bombing Accuracy for Marine Corps Aircraft (Logit Model) N=1741
***significant at .01 level
**significant at .05 level
*significant at .1 level

Independent Variable	Value of Coeff.	Std. Error
Constant	5.00	.38
Career flying hours for manual drops	-.1174	.041
Career flying hours for automatic drops	-.1086	.031
Flights in previous 7 days for manual drops	-.0610	.026
Simulator hours in previous 7 days for fleet pilots	-.01895	.10

Determinants of C-130 Drop Accuracy for Lead Aircraft (Logit and Tobit Models) N=477
***significant at .01 level
**significant at .05 level
*significant at .1 level

Independent Variable	Tobit Model, Coeff. (Std. Error)	Logit Model, Coeff. (Std. Error)
Constant	4.51 (.14)***	-3.27 (.56)***
Copilot career flying hours	-.10924E-3 (6.09E-4)*	.33198E-3 (2.2E-4)
Navigator flying hours past 60 days	-.33751E-2 (1.52E-3)**	.20110E-1 (6.4E-3)***

	Night flight	.25005 (.084)***	-.59405 (.35)*
	Partial copilot career flying hours	-.0134	.435E-4
	Partial navigator flying hours past 60 days	-.3657	.264E-2
	Determinants of C-130 Drop Accuracy for Lead Aircraft: with Simulator ***significant at .01 level **significant at .05 level *significant at .1 level N=477		
	Independent Variable	Tobit Model	Logit Model
	Constant	4.99 (.32)***	-6.66 (1.36)***
	Copilot career flying hours	-.16113E-3 (3.80E-4)**	.74676E-3 (2.3E-4)***
	Log Ratio: Copilot simulator to flying hours	-.64142 (.38)*	4.5 (2.77)***
	Navigator flying hours past 60 days	-.3526E-2 (1.50E-3)**	.019507 (.0062)***
	Partial copilot career hours	-.89E-2	.274E-4
	Partial copilot simulator hrs.	-.1311	.111E-2
	Partial navigator flying hrs past 60 days	-.3851	.256E-2

STUDIES ON APTITUDE AND PERFORMANCE

Title	"Are Smart Tankers Better? AFQT and Military Productivity"
Author	Barry Scribner, D. Alton Smith, Robert Baldwin, and Robert Phillips
Date	1986
Method	Controlled trials using tank crew (TC) firing scores recorded from a simulation carried out January to June 1984, conducted by the Seventh Army Training Center standardized TANK course
Functional Form	OLS regression used, log-log production function. Variables include dummy variables for tank type (M-1=1, M-60=0), dummy for gunner's civilian education (high school=1), dummy for TC's civilian education, dummy for gunner's race (black=1), dummy for TC's race (black=1), dummy for changes in tank table 8 occurring midway through firing (after change=1), natural log of gunner's AFQT, natural log of TC's AFQT, natural log of TC's time in position on the tank in months, natural log of gunner's time in service in years, natural log of TC's time in service in years.

Summary Findings	The authors find that changes in AFQT score are correlated with changes in the performance of tankers in the simulation exercise. For example, with increase in AFQT score for tankers from category IV (20th percentile) to an average for category IIIA (60th percentile) there will be an increase in performance of 20.3 percent. The crew's performance will increase 34 percent for the same change in the gunner's AFQT. The research also suggests that time in service and time in position also have an effect on performance, although the authors do not present empirical results for this.	
Quantitative Results	Explanatory Variable N=1131	Coefficient (Standard Error)
	Natural log of gunner's AFQT	.20514 (.06259)
	Natural log of TC's AFQT	.14913 (.05565)
	Natural log of gunner's time in position on tank (in months)	.02341 (.00679)
	Natural log of TC's time in position (months)	.01260 (.00808)
	Natural log of gunner's time in service (years)	.006776 (.3941)
	Natural log of TC's time in service (years)	-.04140 (.05633)

Title	*Air Force Research to Link Standards for Enlistment to On-the-Job Performance*
Author	Mark Teachout and Martin Pellum
Date	1991
Method	The authors collected hands-on performance test (HOPT) scores and AFQT scores for all individuals in their sample. They analyze the HOPT test scores by finding the mean and standard deviation of the HOPT scores based on the individual's AFQT score and months of experience. They also consider intercorrelations between HOPT, job experience, aptitude (AFQT), and educational attainment.
Functional Form	NA
Summary Findings	Findings support the relevance of AFQT to job performance. The authors consider how AFQT scores are related to HOPT scores for Air Force maintenance positions. For each of the eight specialties considered, the mean HOPT score is higher for those with AFQT scores ranging from I to IIIA than for those with lower AFQT scores. Except for a few cases, the authors find this trend regardless of the experience level of personnel studied. This is a significant observation because it suggests that aptitude, as measured by AFQT, remains an important predictor of job performance even after an individual has been serving for three years.

Quantitative Results	HOPT Scores (selected AFSs)		AFS 122X0		AFS 423X5		AFS 429X1		AFS 732X0	
	Job Exp. (Mos.)		AFQT I-IIIa	AFQT IIIb-IV	AFQT I-IIIa	AFQT IIIb-IV	AFQT I-IIIa	AFQT IIIb-IV	AFQT I-IIIa	AFQT IIIb-IV
	1-12	Mean	41.4	42.3	45.2	44.4	43.3	40.2	42.1	39.3
		SD	16.6	9.4	7.9	--	11.3	8.	7.5	6.7
	13-24	Mean	48.5	47.7	47.8	47.9	53.5	47.3	47.5	43.9
		SD	9.2	5.2	9.2	7.6	6.4	12.2	9.0	9.0
	25-36	Mean	52.5	50.4	50.5	48.0	56.9	56.1	54.3	49.5
		SD	9.3	9.4	11.8	10.0	8.6	8.8	10.1	7.6
	37+	Mean	50.8	56.7	56.3	49.1	53.4	49.8	57.1	49.0
		SD	10.5	6.2	9.2	10.1	11.2	5.8	8.4	10.4
	Total	Mean	50.3	48.8	50.6	48.2	52.3	47.2	51.7	46.8
		SD	10.6	9.0	10.6	8.9	9.8	10.7	10.3	9.2
		N	114	58	146	73	74	53	116	63

Title	*The Effect of Personnel Quality on the Performance of Patriot Air Defense System Operators*
Author	Bruce Orvis, Michael Childress, J. Michael Polich
Date	1992
Method	Controlled trials using simulation of air battles (a point defense situation, an area defense situation, a battalion scenario, and a mixed defense scenario) using the Patriot Conduct of Fire Trainer System in order to assess the effect of personnel quality and training background affect execution in 'warlike' situations.
Functional Form	Linear function and OLS regression. The variables included in the analysis are AFQT category, operator year, unit member or AIT graduate, days of simulation training each ten days, location (overseas or US base). To facilitate comparison across scenarios, the researchers standardize their dependent variables to create Z-scores. The functional form for computing Z-scores is $Z = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$ where Z= predicted Z-score on outcome measure A=intercept b_1X_1= AFQT regression coefficient * AFQT percentile score b_2X_2= operator time in service coefficient * months of operator experience b_3X_3= unit member coefficient * unit membership score (1 or 0) b_4X_4= location coefficient * overseas location score (1 or 0) b_5X_5= training days coefficient * number of training

	days
Summary Findings	Finds a significant relationship between AFQT scores and the outcomes of air battles, both in terms of knowledge assessed by written tests and in actual performance in simulations. The number of significant effects found for AFQT scores dominates the number of significant effects found for other variables included in the model. The authors also note that their results suggest that a one level change in AFQT category equaled or surpassed the effect of one year of operator experience or of frequent training. Finally, operator and unit experience are also important variables. After AFQT, they had the most consistent effect on performance.
Quantitative Results	REGRESSION RESULTS N=315 (218 unit members, 97 advanced individual training (AIT) students)

Explanatory Variable: Asset Defense	Area Defense Coeff. (SE)	Point Defense	Battalion	Mixed Defense
AFQT	.009 (.003)	.011 (.003)	.003 (.003)	.012 (.003)
Operator year	.006 (.007)	.017 (.007)	.008 (.008)	.017 (.007)
Unit member	.141 (.15)	-.178 (.15)	-.269 (.15)	.065 (.14)
Simulation training each 10 days	.004 (.003)	.008 (.003)	-.001 (.006)	.003 (.004)

Explanatory Variable: Missile Conservation	Area Defense Coeff. (SE)	Point Defense	Battalion	Mixed Defense
AFQT	.008 (.003)	.007 (.003)	.000 (0)	.006 (.003)
Operator year	.007 (.007)	.001 (.007)	-.002 (.007)	.003 (.008)
Unit member	.239 (.15)	.431 (.16)	.392 (.16)	.512 (.15)
Simulation training each 10 days	.005 (.003)	.002 (.003)	-.000 (0)	-.007 (.003)

Explanatory Variable: Battlefield Survival	AFQT	Operator Year	Unit Member	Simula-tion Trainin g each 10 days
Coeff. (Std. Error)	.014 (.003)	.015 (.007)	.401 (.14)	.006 (.003)
Explanatory Variable: Tactical Kills	Area Defense Coeff. (SE)	Point Defense	Battalion	Mixed Defense
AFQT	.008 (.003)	.012 (.003)	.009 (.003)	.009 (.003)
Operator year	.010 (.007)	.010 (.007)	.009 (.007)	.005 (.007)
Unit member	.309 (.15)	.260 (.15)	.443 (.16)	.580 (.15)
Simulation training each 10 days	.008 (.003)	.007 (.003)	-.003 (.004)	-.001 (.004)

Title	*Effect of Aptitude on the Performance of Army Communications Officers*
Author	John Winkler, Judith Fernandez, J. Michael Polich
Date	1992
Method	Simulation (two separate procedures for operations and troubleshooting) considering the performance of 240 three-person groups recently graduated from Signal Center's AIT course and 84 groups from active-duty signal battalions. Measured their performance and success on simulations of several tasks including making system operational or identifying problems and solving them. Authors used the Reactive Electronic Equipment Simulator to conduct the exercises and assess performance.
Functional Form	Logistic analysis, functional form y = 1/(1+e-bx) where y is the outcome, x is a vector of independent variables, and b is a vector of the coefficients. Variables used included average age of group members, variables representing the number of group members who were male, white, high school graduates (each coded 0 through 3), a dummy variable for whether the test group was composed of unit members (coded 1) or AIT graduates (coded 0), the number of group members currently using the AN/TRC-145 in their regular job (coded 3 for AIT grads and 0 through 3 for unit members), a dummy variable indicating whether the

	test group contained any reserve component members.		
Summary Findings	Finds that AFQT scores, as a measure of the quality of recruits, contributes to the effectiveness of communication in teams. More specifically, for groups with an average AFQT at the midpoint of category IIIA, the model predicts that 63 percent of units will successfully operate the system in the allotted time. However, if the average AFQT is lowered to the midpoint of IIIB, the prediction is that only 47 percent of units will be successful. The same was found to be true for the troubleshooting task. Finds furthermore that each additional high-scoring member added to the team improved the probability that the group will succeed by about 8 percent points. This result indicates that the effect of AFQT is additive.		
Quantitative Results	System Operation and Average Group AFQT N=323 * significant at .05 level		
	Variable	Coeff.	Std. Error
	Average group AFQT score	.041	.013*
	Test population (unit members)	1.766	.529*
	Number of members using equipment	.440	.282
	Average age of operators	-.110	.058
	Number high school graduates	.034	.252
	Reservists in group	.255	.287
	System Operation and Individual AFQT N=323 * significant at .05 level		
	Variable	Coeff.	Std. Error
	AFQT of terminal A operator	.017	.007*
	AFQT of relay operator	.009	.007
	AFQT of terminal B operator	.015	.007*

	Test population (unit members)	1.799	.532*
	Number of members using equipment	.434	.283*
	Average age of operators	-.112	.058
	Number of high school graduates	.032	.253
	Reservists in group	.264	.288
	Terminal Preset Performance (repair task) N=323 *significant at .05 level		
	Variable	Coeff.	Std. Error
	AFQT score	.015	.007*
	Training indicator	.325	.243
	Education (high school graduate)	-.166	.347
	Practice time on simulator before test	.009	.002*
	Number of hand- on training sessions	.002	.044
	Age	-.055	.040
	System Troubleshooting and Average Group AFQT N=187 *significant at .05 level		
	Independent Variable	Coeff.	Std. Error
	Average group AFQT	.042	.016
	Average age of operators	-.134	.069
	Number of HS graduates	.502	.315
	Number of active duty members	.055	.169
	System Troubleshooting and AFQT Score by Position N=187 *significant at .05 level		

	Variable	Coeff.	Std. Error
	AFQT of terminal A operator	.007	.008
	AFQT of relay operator	.028	.009*
	AFQT of terminal B operator	.008	.008
	Average age of operators	-.130	.069*
	Number of high school graduates	.517	.315*
	Number of active duty members	.103	.172
	Ability to Complete AGC Alignment to Standard N=296 * significant at .05 level		
	Variable	Coeff.	Std. Error
	AFQT score	.025	.009*
	Training indicator	.063	.260
	Age	-.211	.303
	Number of training sessions	-.108	.064
	Component (active duty)	.482	.341
	Ability to Complete Squelch Adjustment to Standard N=286 * significant at .05 level		
	Variable	Coeff.	Std. Error
	AFQT score	.027	.011*
	Training indicator	-.294	.338
	Age	-.110	.079
	Number of training sessions	.122	.139
	Component (active duty)	.694	.395
Title	"Soldier Quality and Job Performance in Team Tasks"		
Author	Judith Fernandez		
Date	1992		
Method	Controlled trials analyzing the team performance among first-term personnel (one group drawn from both active and reserve components that had just		

<table>
<tr><td></td><td colspan="3">received AIT and a second that had 6 to 18 months of experience in the field) on the performance of a simulated troubleshooting task (which involved identifying the faults in a communication system)</td></tr>
<tr><td>Functional Form</td><td colspan="3">Ordered Logistic Function. Functional form y = 1/ (1+e-bx) where y is the outcome, x is a vector of independent variables, and b is a vector of the coefficients. Variables included are average group AFQT (normal form), average age of operators, number of high school graduates, number of whites, number of males, number of active duty members, regimen, course syllabus used.</td></tr>
<tr><td>Summary Findings</td><td colspan="3">Results of analysis suggest that higher AFQT scores were associated with better troubleshooting performance (ability to identify a larger number of faults). The number of high school graduates on a team and the average age of the soldier are also marginally significant. The study suggests that average team AFQT score has an effect on the number of faults detected and that the differential between high and low AFQT performance becomes larger as the number of faults increases. The author also notes that a change in the curriculum used to train soldiers in communications repair can have a significant independent effect on the performance of the team. Finally, the effect of AFQT scores is additive, meaning that team performance improves for each additional high AFQT member.</td></tr>
<tr><td rowspan="7">Quantitative Results</td><td colspan="3">System Troubleshooting Success and Group Aptitude and other Variables N=187
* significant at .05 level</td></tr>
<tr><td>Variable</td><td>Coeff.</td><td>Std. Error</td></tr>
<tr><td>Average group AFQT</td><td>.042</td><td>.016*</td></tr>
<tr><td>Average age of operators</td><td>-.135</td><td>.069</td></tr>
<tr><td>Number of high school graduates</td><td>.502</td><td>.315</td></tr>
<tr><td>Number of active-duty members</td><td>.055</td><td>.169</td></tr>
<tr><td>Course syllabus used</td><td>.926</td><td>.357*</td></tr>
</table>

NOTE: NA = Not applicable.

BIBLIOGRAPHY

Albrecht, Mark, *Labor Substitution in the Military Environment: Implications for Enlisted Force Management,* Santa Monica, Calif.: RAND Corporation, R-2330-MRAL, November 1979.

Beland, Russell, and Aline Quester, "The Effects of Manning and Crew Stability on the Material Condition of Ships," *Interfaces,* Vol. 21, No. 4, July-August 1991, pp. 111-120.

Dalhman, Carl, Robert Kerchner, and David Thaler, *Setting Requirements for Maintenance Manpower in the U.S. Air Force,* Santa Monica, Calif.: RAND Corporation, MR-1436-AF, 2002.

Doyle, Mary Anne, *Youth vs. Experience in the Enlisted Air Force: Productivity Estimates and Policy Analysis,* Santa Monica, Calif.: RAND Corporation, RGSD-139, 1998.

Fernandez, Judith, "Soldier Quality and Job Performance in Team Tasks," *Social Science Quarterly,* Vol. 273, No. 2, June 1992.

Gotz, Glenn, and C. Robert Roll, "The First-Term/Career Mix of Enlisted Military Personnel," in Donald Rice, *Defense Resource Management Study,* Washington, D.C.: U.S. Department of Defense, February 1979.

Gotz, Glenn, and Richard Stanton, *Modeling the Contribution of Maintenance Manpower to Readiness and Stability,* Santa Monica, Calif.: RAND Corporation, R-3200-FMP, 1986.

Hammon, Colin, and Stanley Horowitz, *Relating Flying Hours to Aircrew Performance: Evidence for Attack and Transport Missions,* Washington, D.C.: Institute for Defense Analyses, IDA Paper P-2609, June 1992.

_____________, *Flying Hours and Crew Performance,* Washington, D.C.: Institute for Defense Analyses, IDA Paper P-2379, March 1990.

Horowitz, Stanley, and Allan Sherman, "A Direct Measure of the Relationship Between Human Capital and Productivity," *Journal of Human Resources,* Vol. 15, No. 1, Winter 1980, pp. 67-76.

Junor, Laura, and Jessica Oi, *A New Approach for Modeling Ship Readiness,* Alexandria, Virginia: Center for Naval Analyses, CRM 95-239, April 1996.

Marcus, A. J., *Personnel Substitution and Navy Aviation Readiness,* Alexandria, Virginia: Center for Naval Analyses, 1982.

Moore, Carol, Heidi Golding, and Henry Griffis, *Manpower and Personnel IWAR 2000: Aging the Force*, Alexandria, Virginia: Center for Naval Analyses, CAB D0003079.A2/Final, January 2001.

Moore, S. Craig, *Demand and Supply Integration for Air Force Enlisted Work Force Planning: A Briefing*, Santa Monica, Calif.: RAND Corporation, N-1724-AF, August 1981.

Moore, S. Craig, Edwin Wilson, and Edward Boyle, *Aircraft Maintenance Task Allocation Alternatives: Exploratory Analysis*, Brooks Air Force Base, Texas: Air Force Human Resources Laboratory, AFHRL-TP-87-10, November 1987.

Orvis, Bruce, Michael Childress, and J. Michael Polich, *Effect of Personnel Quality on the Performance of Patriot Air Defense System Operators*, Santa Monica, Calif.: RAND Corporation, R-3901-A, 1992.

Scribner, Barry, D. Alton Smith, Robert Baldwin, and Robert Phillips, "Are Smart Tankers Better? AFQT and Military Productivity," *Armed Forces and Society*, Winter 1986, pp. 193-206.

Teachout, Mark, and Martin Pellum, *Air Force Research to Link Standards for Enlistment to On-the-Job Performance*, Brooks Air Force Base, Texas: Air Force Human Resources Laboratory, AFHRL-TR-90-90, February 1991.

Warner, J. T., and B. J. Asch, "The Economics of Military Manpower," in Keith Hartley and Todd Sandler (eds.), *Handbook of Defense Economics*, New York: Elsevier, 1995.

_______, "The Economic Theory of a Military Draft Reconsidered," *Defence and Peace Economics*, Vol. 7, 1996, pp. 297-312.

Winkler, John, Judith Fernandez, and J. Michael Polich, *Effect of Aptitude on the Performance of Army Communications Officers*, Santa Monica, Calif.: RAND Corporation, R-4143-A, 1992.