

United States General Accounting Office

Report to the Chairman, Subcommittee on Military Personnel, Committee on National Security, House of Representatives

July 1996

PHYSICALLY DEMANDING JOBS

Services Have Little Data on Ability of Personnel to Perform





GAO	United States General Accounting Office Washington, D.C. 20548
	National Security and International Affairs Division
	B-272083
	July 9, 1996
	The Honorable Robert K. Dornan Chairman, Subcommittee on Military Personnel Committee on National Security House of Representatives
	Dear Mr. Chairman:
	As you requested, we reviewed the use and development of gender-neutral occupational performance standards in the military. Specifically, we (1) determined the military services' approaches to implementing gender-neutral performance standards and screening servicemembers to ensure that they can meet the physical demands of their occupations, (2) determined how the military services identified the extent to which servicemembers had problems in accomplishing the physical demands of their jobs, and (3) evaluated the Air Force's implementation of its strength aptitude testing program.
Results in Brief	Each of the services has taken a different approach to screening servicemembers for physically demanding occupations. The Air Force is the only service that requires all new recruits to take a strength aptitude test—which consists of lifting weights on a single weight machine—and uses the results to qualify individuals for their military occupations. Until 1990, the Army required recruits to take a virtually identical strength test but used the results only to counsel recruits about the physical requirements of their desired military occupations. The Navy and the Marine Corps have not adopted occupational strength testing as a means of screening new recruits for physically demanding occupations.
	In an April 1995 report to Congress, ¹ the Air Force, the Navy, and the Marine Corps reported that they had experienced good results with their approaches to matching servicemembers to physically demanding occupations or had encountered few problems; the Army was silent about its results. According to Department of Defense (DOD) and Army officials, the services based their conclusions on the absence of complaints from servicemembers that they were having problems completing physically demanding tasks. However, the services have little data on which to base their conclusions. We found that only the Army had systematically
	¹ Gender Neutral Standards, Report to the House Committee on National Security, Senate Committee

on Armed Services, and House and Senate Committees on Appropriations, Office of the Assistant Secretary of Defense (Force Management Policy), April 1995.

collected physical performance data since 1989 in 21 occupations and concluded that 59 to 84 percent of servicemembers in 7 selected surveyed occupations had no problems completing physically demanding tasks. A 1994-95 data collection for 10 of 267 occupations (including 6 of the same occupations reported in the 1989 data collection) found 51 to 79 percent of servicemembers had no problems completing physically demanding tasks.

The Air Force strength aptitude test, or an earlier version of it, has been in use since 1976. However, questions about the current test's effectiveness in predicting capability to do physically demanding tasks, problems in the administration of the test to new recruits, and delays in updating occupational requirements raise concerns about whether the test is used correctly, or is even useful.

Background

In fiscal year 1980, of the more than 2 million servicemembers on active duty, over 170,000 (8.4 percent) were women. Congressional action and DOD policymaking lifted the prohibition on women serving in positions in combat aviation, aboard combatant vessels,² and in ground units (brigade level and above) and DOD's new definition of combat jobs³ opened over 259,000 additional military positions to women servicemembers since April 1993. By December 1995, the number of women serving on active duty had risen to over 191,000 (about 12.8 percent of the approximately 1.5 million servicemembers). At the time of our report, DOD had opened over 80 percent of all positions to all servicemembers, ranging from a low of 62 percent of positions open to women in the Marine Corps to a high of over 99 percent of positions open to women in the Air Force.

Section 543 of the Fiscal Year 1994 National Defense Authorization Act required the services to adopt gender-neutral occupational performance standards and defined those as being work standards that are common, relevant, and not based on gender. The act also required the services to adopt physical performance standards for any occupation in which DOD determined that strength, endurance, or stamina was essential to the performance of duties. The DOD General Counsel later determined that the

²The Fiscal Year 1992 and 1993 National Defense Authorization Act repealed 10 U.S.C. section 8549, lifting the prohibition on women serving aboard combat aircraft, and the Fiscal Year 1994 National Defense Authorization Act repealed 10 U.S.C section 6015, lifting the prohibition on women serving aboard combat ships.

³Based on a policy memorandum from the Secretary of Defense (January 13, 1994), DOD excludes female servicemembers from units below the brigade level whose primary mission is to engage in direct combat on the ground. DOD defined ground combat as "engaging an enemy on the ground with individual or crew served weapons, while being exposed to hostile fire and to a high probability of direct physical contact with the hostile force's personnel."

	services were not required to have physical standards for any occupation but that if such standards did exist they would have to be applied on a gender-neutral basis for any occupation open to both men and women.
	The services use a variety of pre-enlistment, job classification, and retention screening devices to select qualified candidates for military service. For example, pre-enlistment screens include requirements that recruits score at or above a specified minimum on a cognitive test and be within a certain height or weight range. Other standards may be occupation-specific, such as requiring recruits entering electronics occupations to demonstrate aptitude in the field of electronics.
Military Services Differ in How They Classify Recruits for Physically Demanding Jobs	DOD has left it to the services to determine how to classify servicemembers into physically demanding occupations. The Air Force is the only service that requires recruits to take a strength aptitude test. Each Air Force enlisted occupation is categorized into one of eight strength categories, and recruits' test scores are used to screen them for their military occupations. The other services permit virtually any recruit to fill nearly all physically demanding occupations provided they meet cognitive, height/weight, and other standards unrelated to strength capacity and restrict women only from occupations closed by combat exclusion policies.
	In 1976, we recommended that DOD develop standards for measuring recruits' ability to meet strength, stamina, and operational requirements because we found that some servicemembers were unable to do physically demanding tasks. ⁴ In response, the Army categorized each enlisted occupational specialty into one of five categories based on physical demand. It required new recruits to take a strength test using the "incremental lifting machine," a weight-lifting machine developed and used by the Air Force. The Army concluded that although the test helped to better match recruits' physical capabilities to requirements of physically demanding occupations, it also prevented more women than men from serving in certain occupations. Consequently, test results were used only to counsel applicants about job assignments. The Army discontinued the test in 1990.
	In the 1970s, the Air Force adopted an earlier version of the test and by 1987 categorized each of its enlisted occupations into one of eight physical

 $^{^4\!}J\!ob$ Opportunities for Women in the Military: Progress and Problems (GAO/FPCD-76-26, May 11, 1976).

	demand categories. The Air Force currently requires all recruits to take the strength aptitude test at a military entrance processing station. The test requires recruits to lift weights on the incremental lifting machine starting at 40 pounds; the weight is then increased in 10-pound increments until the recruit (1) cannot complete a lift, (2) asks to stop, or (3) lifts 110 pounds (the maximum for any occupation in the Air Force). An Air Force counselor uses the results to match recruits to occupations based on the eight physical demand categories and screens out applicants who the test results indicate would have difficulty performing physically demanding jobs.
	The Navy considered using a strength test to screen applicants for entry into physically demanding military occupations and concluded that more women than men would have been excluded from such jobs. The Navy concluded, however, that women were already meeting the physical demands of their occupations and, for that reason, did not implement its test or categorize its occupations by physical demand. Similarly, the Marine Corps has not adopted an occupationally based strength test or categorized its occupations by physical demand.
The Services Have Little Data to Assess Capability to Perform Physically Demanding Tasks	Except for the Army, the services have not collected data on servicemembers' ability to do physically demanding jobs and have little basis on which to conclude that servicemembers are not having problems. We are concerned that some servicemembers may have difficulty doing some physically demanding tasks based on the results of a limited survey conducted by the Army Research Institute (ARI) and anecdotal information we obtained in interviews with servicemembers. However, given limitations on the ARI survey and our interviews, we were not able to assess the significance of the problem.
	In 1989, 1994, and 1995, ARI surveyed servicemembers ⁵ in selected Army occupations. In 1989, ARI surveyed 21 combat and noncombat occupations and found that 59 to 84 percent of male and female servicemembers in 7 selected noncombat occupations reported no difficulty in lifting objects. In the 1994-95 follow-on survey of 10 of 267 occupations, ARI found that 51 to 79 percent of servicemembers reported no difficulty in lifting objects in some of the same occupations as those looked at in the 1989 survey. Because the surveys did not address the significance of the problem and rely on self-reported data, the results must be used with caution. On the

⁵ARI briefing, Physical Performance Research Update, October 1995.

other hand, the results also suggest that the Army may have servicemembers who have had difficulty doing physical tasks.

The other services have not done any systematic assessment of the capability of their personnel to perform the physically demanding aspects of their jobs. According to DOD and Army officials, the services rely upon the absence of complaints filtering up from operational units as an indicator that widespread performance problems do not exist. Supervisory personnel we spoke with, however, indicated that they would work around⁶ individual performance capability problems or redistribute tasks and that it was unlikely such information would be channeled to higher levels unless widespread problems were encountered.

Our discussions with about 100 Army personnel in 5 occupational specialties (2 of which were used in ARI's survey) anecdotally supported ARI's finding that some soldiers were having difficulty completing some physically demanding tasks. In addition, in discussions with over 300 military personnel in the Air Force, the Navy, and the Marine Corps, some individuals stated that at one time or another, they had difficulty with some aspect of their job. Given the limited number of personnel we interviewed and the limited number of military specialties we reviewed, we were unable to determine whether such problems were widespread.

All four services told us that they have the capability and infrastructure already in place to collect data on physical demands of occupations at little or no additional cost. Each of the services has ongoing processes through which they can identify occupational tasks in each specialty in order to revise training curriculums and which they use for other reasons. However, the services do not collect data on the physical demands of jobs with these processes.

Surveys, identification of physically demanding tasks, or other data collection efforts, could be used as a first step in identifying occupations in which servicemembers have difficulty and can identify occupations that are candidates for reengineering to reduce the physical demands placed on servicemembers. For example, the Army Research Laboratory has a pilot reengineering project underway that attempts to identify opportunities to reengineer selected occupations to reduce the physical demands and enhance job sustainment, safety, and personnel utilization.

⁶According to servicemembers we spoke with, one or more other servicemembers would work together to complete a task that was too demanding for one person to do. Servicemembers we spoke with considered this approach realistic as long as sufficient numbers of personnel were available to lend assistance.

	In addition, the Air Force has a number of reengineering studies underway.
	Systematic data collection on physically demanding tasks could be used to develop occupation-specific physical strength training. For example, the Army's Training and Doctrine Command (TRADOC) has commissioned the Army's Research Institute of Environmental Medicine to develop a database of physically demanding tasks in Army occupations. TRADOC is considering using the database to establish specific physical strength training to help servicemembers meet the physical demands of their jobs.
	According to DOD, current training consists of classroom training that tends to be less physically oriented than on-the-job training. Once in their duty assignments, servicemembers continue their on-the-job training. According to DOD, training standards are based on tasks, duties, and knowledge required to perform in an occupation and men and women are held to the same standards.
The Air Force Strength Aptitude Test Program May Not Be Valid	The Air Force is the only service that uses strength aptitude testing as a prerequisite for entry into specific military occupations. Air Force recruits take the Armed Services Vocational Aptitude Battery (ASVAB) and must pass a physical given at a military entrance processing station. If they pass the physical, recruits then take the strength aptitude test, and their scores are recorded in their medical records. Finally, recruits meet with an Air Force counselor who matches them to a military occupation based on the ASVAB and strength aptitude test scores, their interests, and the needs of the Air Force. However, Army, Navy, and independent research raises questions about the predictive validity of the test currently used by the Air Force, and we found several problems with implementation of the Air Force testing program.
Research Questions the Validity of Test Results Obtained With the Incremental Lifting Machine	 Since 1982, at least nine studies have been published or presented that raise questions about the validity of the incremental lifting machine test as a predictor of performance in military occupations, particularly if the test is relied upon as the sole measure of predicted performance. A 1982 study sponsored by the Air Force reported that the incremental lifting machine was the best single predictor of task performance. The result was based upon transformation of the combined male and female scores that minimized the differences in those scores but resulted in giving

the appearance of improving the predictive power of the incremental lifting machine beyond the experimental results.⁷

- A 1984 study done for the Army⁸ found that the incremental lifting machine was a good predictor of a set of Army simulated occupational tasks, accounting for 67 percent of the explanation of scores on the tasks. However, the study misstated the relationship because it combined significantly different male and female lifting scores to determine the predictive power of the incremental lifting machine scores. When we examined the reported scores by gender, the correlation of the incremental lifting machine scores is considerably lower for male and female scores than reported for the aggregated score.
- A 1985 Navy study⁹ stated that combining male and female incremental lifting machine scores would involve making an assumption that male and female scores are evenly distributed throughout the entire group, a tenuous assumption according to the text. By using separate male and female scores, the study compared 7 strength test measures, including 3 different incremental lifting machine lifts, with 19 shipboard tasks and concluded that "some of the best correlates of shipboard performance are the armpull, ergometer, and body weight," which are 3 nonincremental lifting machine measures.
- A 1985 study conducted by the Army's Research Institute of Environmental Medicine found that women tended to be shorter than men and thus were required to spend relatively more time lifting with their upper body than males and consequently scored lower in tests using the incremental lifting machine (given that women tend to have less upper body strength than men, according to this and other research). On the other hand, the study found that an alternate strength test that focused more on the use of the lower body produced female scores that were closer to those of males in the study population.¹⁰
- An ARI study in 1993¹¹ concluded that variables such as job performance and the Army's physical readiness test were not strongly related to scores on the incremental lifting machine. According to the study, the Army should not place great confidence in the use of a single lifting test as a

⁹Robertson et al., <u>Documentation of Muscularly Demanding Job Tasks and Validation of an</u> Occupational Strength Test Battery, November 1985.

¹⁰Teves et al., Performance on Selected Candidate Screening Test Procedures Before and After Army Basic and Advanced Individual Training, U.S. Army Research Institute of Environmental Medicine, June 1985.

¹¹Elizabeth J. Brady and Michael G. Rumsey, <u>Physical Performance in Army Enlisted MOS Revisited</u>, Presentation to the Military Testing Association, November 1993.

⁷Ayoub et al., Establishing Physical Criteria for Assigning Personnel to U.S. Air Force Jobs, 1982.

⁸Myers et al., Validation of the Military Entrance Physical Strength Capacity Test, January 1984.

	 selection measure of physical fitness and should consider a more comprehensive approach to physical screening. A Canadian research team produced four sequential studies since 1990¹² and concluded that gender differences in incremental lifting machine scores and box-lifting tasks were heightened by an incremental lifting machine test protocol prohibiting subjects from moving their feet or shifting their weight to achieve a more comfortable lifting posture. When subjects were allowed to lift in their most comfortable method, they could lift heavier boxes. For female subjects, the incremental lifting machine score became less related to their box-lifting scores as the constraints were relaxed.
Recruits Are Not Tested to Their Full Potential	Many recruits that took the strength aptitude test at a military entrance processing station scored lower during their initial tests than they did when retested during basic military training. For example, our analysis of data provided by the Air Force showed that between December 1995 and February 1996, 244 females were retested during the second week of basic military training and lifted an average of nearly 18 pounds more than they did initially; the 211 males' average increase was nearly 15 pounds. Of the 455 recruits who retested, 3 lifted 10 percent more, and all but 10 lifted from about 11 percent to 120-percent more or an average of 23.3 percent more (10 lifted less). A study conducted by the Air Force concluded that servicemembers who engaged in physical training programs of about 9 weeks increased arm strength by just 6 percent. ¹³ According to Air Force officials, nearly all of another approximately 3,900 recruits retested at basic military training between April 1994 and November 1995 also scored higher, although individual scores were not readily available.
	According to the researcher who oversees the strength aptitude program, some increases in test scores are attributable to increased motivation on the part of the recruit at basic training or by permitting recruits to adopt a lifting technique not in accordance with the test protocol. However, the researcher concluded that neither increased motivation nor a change in the test protocol can explain the magnitude of the increase we found.
	¹² J. Stevenson et al., "The Effect of Lifting Protocol on Comparisons with Isoinertial Lifting

¹²J. Stevenson et al., "The Effect of Lifting Protocol on Comparisons with Isoinertial Lifting Performance," <u>Ergonomics</u>, 1990; J. Stevenson et al., "Dynamic Analysis of Isoinertial Lifting Technique," <u>Ergonomics</u>, 1990; J. Stevenson et al., "Development of Factor-Score-Based Models to Explain and Predict Maximal Box-Lifting Performance," <u>Ergonomics</u>, 1995; and J. Stevenson et al., "Gender Differences in Performance of a Selection Test Using the Incremental Lifting Machine," Applied Ergonomics, 1996.

 $^{^{13}\}mbox{McDaniel}$ et al., Weight Lift Capabilities of Air Force Basic Trainees, May 1983.

	Because nearly all of those who retook the strength aptitude test scored higher, we question the validity of the scores of recruits who were not given an opportunity to retake the test. With the exception of those who initially lifted 110 pounds (the maximum weight requirement for any Air Force occupation), the Air Force cannot ensure that everyone else has also been tested to their full potential.
Physical Strength Standards for Most Air Force Specialties Have Not Been Updated	 According to the researcher who oversees the strength aptitude program, occupational specialty strength standards must be kept current to maintain the program's validity. However, since 1986, the Air Force has updated the strength standards for only 12 specialties. In addition, 16 more were being resurveyed at the time of our report. For the remaining Air Force specialties, strength standards are based on data gathered between 1978 and 1982. According to the researcher, unless something in the job changes, the strength standard is still current. We were unable to evaluate whether changes may have been made in any of the remaining 227 Air Force specialties because the original data is stored on computer tape in a format not readable by computers now in use in the Air Force. We were told that a contractor might be able to convert the data to readable form, but the task could be costly and potentially time-consuming. According to a 1995 Air Force Aerospace Armstrong Medical Laboratory memorandum, the strength requirement should be resurveyed whenever two or more occupations with different strength standards are merged. However, since October 1993, the Air Force has merged or split 11 occupations within differing strength categories. In addition, the researcher who oversees the strength aptitude program has identified another 11 specialties that also need to be resurveyed. As a result, the Air Force has not determined the current strength requirement for 22 merged, split, or changed occupations. The Air Force will run the risk of denying servicemembers' entry into occupations based on invalid or outdated strength requirements in those merged occupations that have not been resurveyed.
Recommendations	Because the services have little systemically collected data on the ability of servicemembers to meet the physical demands of occupational tasks, we recommend that the Secretary of Defense require the services to assess whether a significant problem exists in physically demanding occupations and identify solutions, if needed. Such solutions could include redesigning job tasks to reduce the physical demands, providing additional training, or

	establishing valid performance standards to enhance job sustainment, safety, and personnel utilization. Given the questions concerning the validity of the strength aptitude test and the implementation problems we found, we recommend that the Secretary of the Air Force reassess the use of the strength aptitude test as a means of predicting future performance in physically demanding occupations.
Agency Comments and Our Evaluation	In commenting on a draft of this report, DOD generally concurred with our findings and recommendations. In response to our first recommendation, DOD stated that it will direct the services to (1) collect data systematically on job performance difficulties and (2) focus on physically demanding occupations with a history of strength-related injuries and occupations recently opened to women. We are concerned, however, that such a narrow focus will not identify all occupations where problems exist. First, because supervisory personnel told us they may assign persons having difficulty to lighter tasks, occupations where servicemembers are having difficulty may not necessarily lead to a higher incidence of strength-related injuries. Working around a problem may prevent injuries, thus limiting the usefulness of medical data for DOD's purpose. Second, if DOD focuses only on occupations recently opened to women, it may overlook strength-related performance problems in occupations and use appropriate data in its study.
	In its response to our second recommendation, DOD stated that it will (1) make every effort to comply with generally accepted professional standards for test development and implementation and (2) direct the Air Force to continue its "periodic validation efforts." However, while the Air Force may have attempted to validate the strength aptitude test periodically, our review did not disclose any study that demonstrated that the incremental lifting machine test had predictive validity.
	DOD's comments are reprinted in appendix I. DOD also provided several technical corrections that we have incorporated into the text of our report as appropriate.

Scope and Methodology	We reviewed DOD's 1995 report to Congress on gender-neutral performance standards; service orders, regulations, and manuals; and research studies undertaken within the services and by independent researchers. We interviewed officials and obtained documents from the Office of the Secretary of Defense (Accessions Policy) and met with officials from the Defense Advisory Committee on Women in the Services in Washington, D.C.
	To complete our work with the Army, we interviewed officials and obtained documents from the Office of the Assistant Secretary of the Army (Manpower and Reserve Affairs), Office of the Deputy Chief of Staff for Personnel, Personnel Command, Training and Doctrine Command, Combined Arms Support Command, Army Transportation Center, Army Research Institute, Army Research Laboratory, and Army Research Institute of Environmental Medicine.
	To complete our review of the Navy, we met with officials from the Office of the Assistant Secretary of the Navy (Manpower and Reserve Affairs); Bureau of Naval Personnel, including the Special Assistant for Women's Policy; the Commander in Chief, Atlantic Fleet; the Naval Manpower Analysis Center; and the Center for Naval Education and Training.
	To complete our Air Force work, we met with officials of the Headquarters of the Air Force (Directorate of Military Personnel Policy), the Air Force Personnel Center, the Air Force Recruiting Service, the Air Force Education and Training Command, the Armstrong Aerospace Medical Research Laboratory, the Occupational Measurements Squadron, and the Military Entrance Processing Command, which administers the strength aptitude test for the Air Force. We also interviewed officials and observed Air Force recruits taking the strength aptitude test at the Military Entrance Processing Stations in Baltimore, Maryland, and Richmond, Virginia.
	To complete our Marine Corps work, we met with officials of the Office of Accessions Policy and Combat Development Command.
	To assess whether the services have a system for identifying demanding tasks that exceed servicemembers' physical capabilities to perform them and identify difficult tasks, we observed activities and met over 400 service personnel employed as instructors, students, operational unit commanders, and enlisted personnel at Forts Eustis and Lee in Virginia; Fort Bragg, Marine Corps Air Station Cherry Point, Camp Lejeune, and Seymour Johnson Air Force Base in North Carolina; Lackland Air Force

Base in Texas; Fort Leonard Wood in Missouri; Naval Air Station Memphis in Tennessee; and aboard the aircraft carrier USS John C. Stennis. As agreed with your office, we concentrated on the occupational areas of bridge engineer, food service specialist, aviation ordnance, and motor transport.

We conducted our work from November 1995 to June 1996 in accordance with generally accepted government auditing standards.

We will send copies of this report to the Secretaries of Defense, the Army, the Navy, and the Air Force; the Commandant of the Marine Corps; the Executive Director, Defense Advisory Committee on Women in the Services; and the Director, Office of Management and Budget. Copies will also be made available to others upon request.

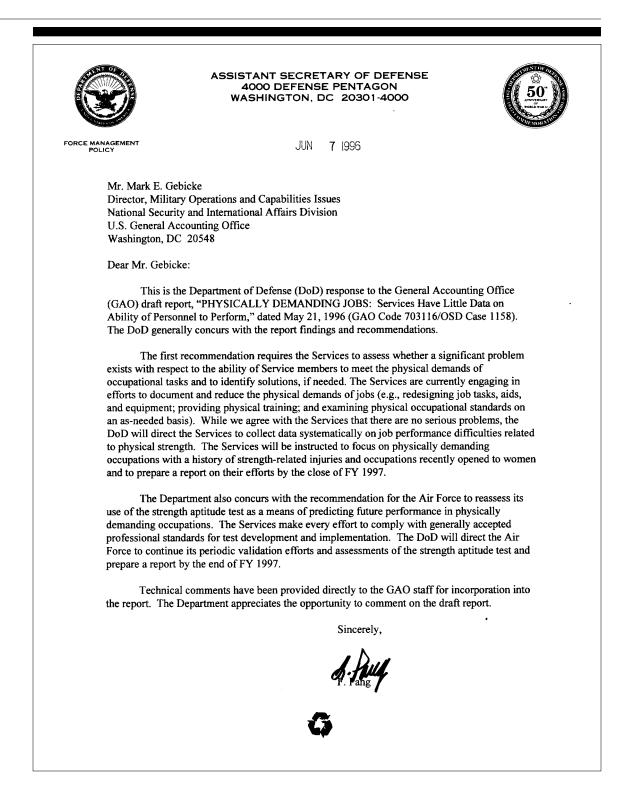
The major contributors to this report are listed in appendix II. If you have any questions about this report, please contact me on (202) 512-5140.

Sincerely yours,

Mark E Selike

Mark E. Gebicke Director, Military Operations and Capabilities Issues

Comments From the Department of Defense



Appendix II Major Contributors to This Report

National Security and International Affairs Division, Washington, D.C.	Sharon A. Cekala William E. Beusse Brian J. Lepore Martin E. Scire Arthur L. James, Jr.
Office of the General Counsel, Washington, D.C.	Ernie E. Jackson
Norfolk Field Office	Lawrence E. Dixon Janine M. Cantin Sharon L. Reid Paul A. Gvoth, Jr.

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are \$2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. VISA and MasterCard credit cards are accepted, also. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office P.O. Box 6015 Gaithersburg, MD 20884-6015

or visit:

Room 1100 700 4th St. NW (corner of 4th and G Sts. NW) U.S. General Accounting Office Washington, DC

Orders may also be placed by calling (202) 512-6000 or by using fax number (301) 258-4066, or TDD (301) 413-0006.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

or visit GAO's World Wide Web Home Page at:

http://www.gao.gov



United States General Accounting Office Washington, D.C. 20548-0001

Official Business Penalty for Private Use \$300

Address Correction Requested

Bulk Rate Postage & Fees Paid GAO Permit No. G100