Statistics in Defense and National Security: The Present--You Need More Than Statistics—You Need the Right Statistics

By
Dr. Nancy L. Spruill,
Dr. Alyson Wilson, and
Dr. Ernest Seglie

I. Introduction:

This paper proposes that Statistics in Defense and National Security—in the present—is both a good news and a bad news story. Are statisticians more involved, and is use of statistics improved, compared to the recent past (post WWII)—-you bet. Is there a lot more that could be done in terms of different aspects of defense and national security (beyond data and testing and analysis) that need to use statistics better and need the expertise of statisticians—-you bet.

This paper addresses the areas where things are better and areas where they could be better. We have come a long way since Professor Richard Savage's concerns, when in an 1984 address at the Joint Statistical Meetings, and in a resulting 1985 paper in JASA [1], he noted that "the last 30 years of statistics in defense is in large part a mystery; defense is a closed activity that should not give current details to the general public." He argued that, especially for defense involving nuclear matters, that statisticians should have a major role, involving both the concept of uncertainty and the measurement of variability. In this paper we focus on the progress that has been made for statistics/statisticians in Defense/National Security, since the mid-1980s, and include areas where the authors believe we need to do more.

II. Statistics in Defense and National Security Between the Mid-1980s and Today--From ASA's perspective:

At least partly as a result of Professor Savage's speech, and paper, and also as a result of the encouragement of Dr. Leslie Kish, in November 1988, ASA President Robert V. Hogg established an Ad Hoc Committee on Statistics and Security (renamed a year later the Ad Hoc Committee on Statistics and International Security) and provided a document which described some of the characteristics he proposed for the committee and possible areas of interest. The Committee was co-chaired by Professors Savage and Kish. Let me quote just the preamble of that document:

"International security is a major concern. The efforts to maintain security require many skills including those of statisticians. In the past, however, statistical societies have shown little attention for statistical aspects of security. The ASA Committee on Statistics and Security will work with a broad definition of security. However, the problems treated by the Committee must be statistical in nature. The work of the Committee must be open in the spirit of the 150-year history of the Association. The Committee as a whole must formulate priorities, the Committee must operate in an ethical manner and with widely accepted values, and the Committee must avoid narrow partisan positions."

The document went on and listed four areas of interest: Accidental Nuclear War, Monitoring Treaties, Economics and Manpower, and Tactics. This institutionalization of Statistics and Security within ASA included an emphasis on contributing to preventing nuclear accidents, including nuclear war. This is first area, the authors believe, where statisticians still need to do more.

Many of you know the rest of this part of the success story, beginning with the Committee struggling to develop a charter, while seeking to reach out to the Defense and National Security community and to the statistics community. After becoming comfortable as a Committee, ASA encouraged the Committee to take the next step to become a section. Without the vision of Professor Dave Banks, the Committee would never have set their foot on that path and without the support of many ASA members, especially Dr. Fritz Scheuren, the Committee would never have made the transition to a Section, where it is today.

And the most recent noteworthy step in this journey for ASA, was when ASA elected Dr. Sallie Keller-McNulty as its President Elect (2005). A significant part of her ASA Election Statement addressed expanding outreach by pursuing three specific goals—the third of which was to: "Strengthen the statistical sciences voice in the national/ international security communities." In particular her vision recognized that:

"National/international security is a problem that is growing in importance at a frightening pace, and my experience at Los Alamos National Laboratory has made me aware that statistical sciences is nearly invisible in this arena. ASA has already taken an important step to remedy this situation through the creation of a new section on Defense and National Security. This is not enough. ASA must provide a strong and unified voice that will make statistics an integral member of the national and international security communities."

Her work, and that of the task force she chartered, represent a large portion of the present push for statistics in defense/national security.

III. Task Force on Defense and National Security--Findings and Follow-up:

In 2006, as ASA President, Dr. Keller-McNulty chartered the Task Force on Defense and Security. The mission of the task force was to provide advice, and a plan, on how the statistical community can better serve the needs of the nation in the areas of defense and security; to identify ways in which those communities could have greater access to statistical expertise; and to identify ways in which the statistical community can learn about areas of concern in defense and security.

The task force identified the following recommendations. The status of each is briefly discussed.

1. The ASA Presidents should meet with the Office of Naval Research and the Army Research Office to raise the visibility of statistics within these organizations. ***While this action was completed, it is important that defense research funding agencies continue to be on the ASA***

Presidents' agenda periodically in the same way that, say, the National Science Foundation is.

2. The ASA Presidents should meet with the Undersecretary of Defense for Personnel and Readiness, the Director of Operational Test and Evaluation, the Undersecretary of Defense for Acquisition, Technology, and Logistics, the Director for Program Analysis and Evaluation, and the Board of the Central Intelligence Agency's Analytic Methodology Network to raise the visibility of statistics within these organizations. ***This action was completed for the DoD, but better connections still need to be established with the intelligence community.***

3. The ASA Presidents should recommend the appointment of a statistician to the Defense Science Board. ***This action was completed. In addition, a statistician was recommended to the Defense Business Board.***

4. The ASA should develop a Research Fellowship Program with the Department of Defense, Office of the Director, Operational Test and Evaluation. ***This action is still pending.***

5. The ASA Board of Directors should task the ASA staff to explore the feasibility of developing a Meeting-Within-a-Meeting set of activities for defense and security professionals in conjunction with the 2009 JSM. ***This action is still pending.***

6. The ASA Board of Directors should consider for inclusion in the ASA's Ethical Guidelines for Statistical Practice any recommendations resulting from the Committee on Professional Ethics review of issues relating to defense and security work. ***After review by the Committee on Professional Ethics, it was determined that the issues faced by statisticians in defense were well-addressed by the current ASA Ethical Guidelines for Statistical Practice.***

7. The Program Chair of the ASA Section on Defense and National Security (SDNS) should develop a yearly Contributed Poster Session at JSM centered on a defense or security-related dataset. ***This action is still pending.***

8. The SDNS should develop a Speakers Program in Defense and Security for a one-year trial, with special emphasis on publicizing the program and seeking seed funding. ***SDNS received Member Initiative funding in 2008 for the one-year trial, and a speaker's list is currently being compiled.***

9. The ASA should host a series of workshops for the defense and security community on topics that could contribute to improved defense and security research methodologies and analysis.

10. The SDNS and Army Conference on Applied Statistics (ACAS) Executive Committees should examine how ACAS and the Quantitative Methods in Defense and National Security Conference can coordinate their missions, programs, and speakers. ***Items 9 and 10 are ongoing.***

11. The SDNS should publicize existing research fellowship opportunities in defense and security. ***This has been started through articles in Amstat News, but must be a continuing priority for SDNS.***

12. The Publications Officer of the SDNS should establish a quarterly electronic newsletter summarizing the titles, abstracts, and locations of statistical articles related to defense and security. ***This action is pending.***

13. The SDNS Program Chair should develop sessions for upcoming Joint Statistical Meetings involving international researchers in defense and security. ***This continues to be a SDNS priority.***

14. The current list of defense-oriented peer reviewers on the SDNS web site should be updated annually and publicized more broadly to those who could use it. More research should be done into the usefulness of the ASA developing a separate list of defense reviewers having security clearances. ***This action has been completed.***

It is worth noting that ASA Member Carey E. Priebe, a professor in the applied mathematics and statistics department at Johns Hopkins University, has recently been selected as one of the inaugural class of six university professors in the National Security Science and Engineering Faculty Fellows Program. Up to $3 million of direct research support will be granted to each NSSEFF Fellow for up to five years. The fellows conduct basic research in core science and engineering disciplines that underpin future DoD technology development.

IV. Using Statistics Today:

The use of statistics and statistical techniques and analysis in defense and national security has improved over the past 25 or so years. Today one can hear at the decision-makers' level, discussions of reliability issues—-where a program is on the reliability growth curve—and discussions of cost estimates—-should the program fund to the 50% confidence level cost estimate or the 80% one. These represent real progress. But in many cases though good statistics are used, few, if any, statisticians are involved. In the next three sections, this paper discusses areas where progress has been made—-data, testing and analysis—-and discuss some aspects where more can be done.

First it is the data. As computers allow us to use more data, we are seeing it as a valuable tool itself. Like most fields, defense/national security is awash in data, but many processes, including most IT systems, especially IT business systems, stumble in large part because of the data. Whether it is a system that will handle manpower data, including pay, or an accounting and accountability system, including those that track/value military equipment, the data is often the biggest problem. Statisticians need to be involved here. First they are needed to make it clear that data is a valuable asset, and therefore there needs to be a plan for collecting and storing the data over time—and there is often the hardest issue -- what to do with existing data—clean it up or write it off. Statisticians need to raise the awareness of the need for good data management and they need to be part of the process.

But there are data success stories. They are slow in coming and they are only for certain databases. One example is DAMIR—Defense Acquisition Management Information Retrieval. It is an initiative that provides to the entire defense enterprise a single authoritative source of program information for the 95 Major Defense Acquisition Programs. These are the programs that buy the major weapons systems for the Department of Defense, programs such as the Joint Strike Fighter or the Future Combat System, and represent an investment of over $1.6 trillion over their life cycle. DAMIR provides a single web-based interface for this information. It enables those in the DoD (including the Military Services), the Congress, and other participating communities to access information relevant to their missions regardless of the agency or where the data resides.

A second example is Kaleidoscope. It is a tool, developed by analysts in the Office of the Secretary of Defense, that uses standard web browsers to look at a large volume of data and identify problems, such as outliers, missing data, etc. It may seem simple to statisticians but often action officers think that getting a data source and doing the analysis are all that is needed. They never look at the data—is it good?—is it good enough to use to base major mission, and/or funding, decisions on?

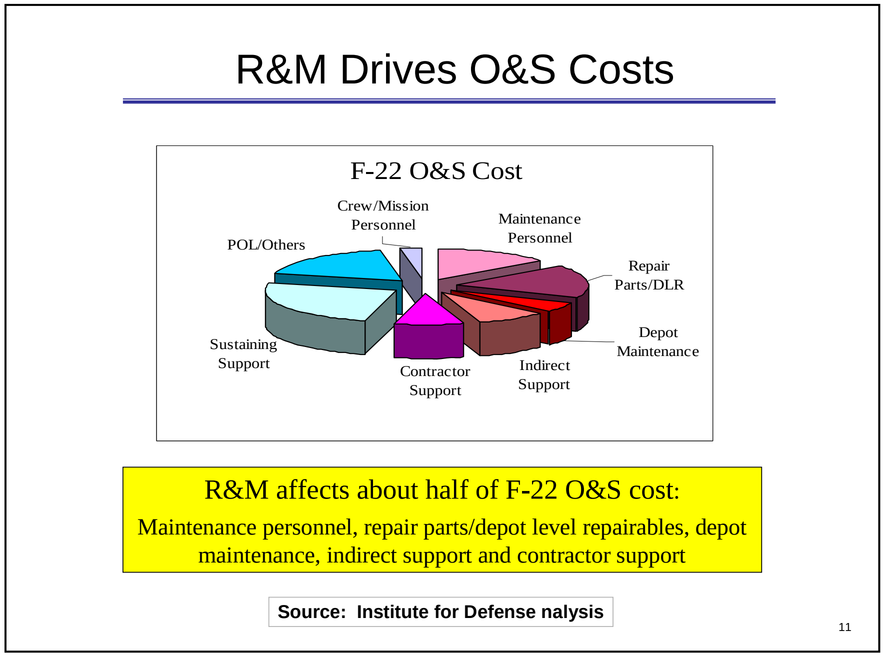
Second it is how you use the data. Total Information Awareness was an initiative to use data mining to gather information on threats to the U.S. I believe that no statisticians were involved with this project because if they had have been they would have raised the confidentiality concerns. No one did, or if they did it was not given enough attention, and as a result Congress took the funding for this initiative and it was stopped. Statisticians need to be more involved in how to use data, especially sensitive data. And this data issue is going to get harder as we are now collecting a lot more data, including sensitive data, such as biometrics information.

V. Using Statistics: On the front lines--from the test and evaluation perspective:

The first example from the test and evaluation perspective is also a data example. As mentioned in the previous section, decision makers today are concerned about reliability growth and cost estimation. These areas are related because reliability drives the operation and support (O&S) costs and O&S costs account for 60 to 70 per cent of the Life Cycle cost. These are costs that occur after the system is bought.

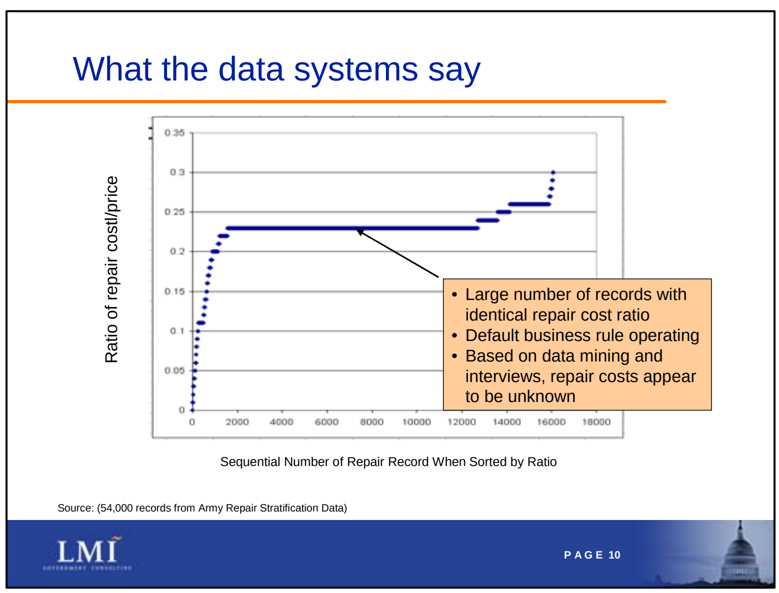
The less reliable a system the more the spares it needs, the more people to maintain it, the greater the infrastructure required such as hangars and special equipment, and the higher the investment in continued research to improve reliability.

A study by the Institute for Defense Analysis indicates how big a role reliability can play in O&S costs. In their detailed study of the F-22 fighter aircraft costs they found that reliability and maintenance affects about half of the O&S costs.



In order to find the current relationship between life cycle costs and reliability, DoD, through a contract with the Logistics Management Institute, attempted to use its standard data bases for cost estimation for a sample of systems that had been recently fielded.

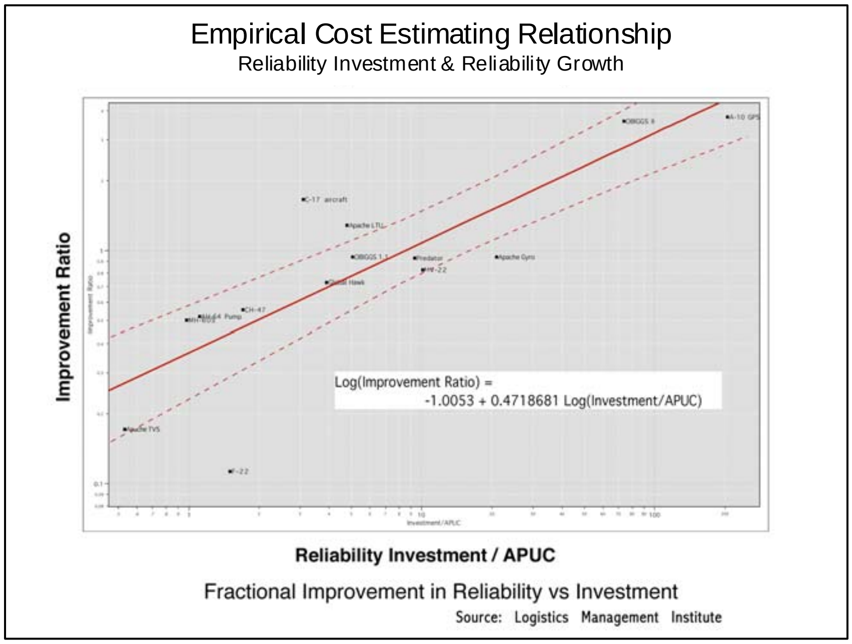
Eventually the effort to use the standard data bases had to be abandoned and a more ad hoc, individualized and labor intensive approach was used. One of the significant problems was uncovered by looking at the vast data base as a whole. The data base was supposed to include the cost to repair a component if it breaks.



LMI analyzed more than 54,000 repair cost records. The results are in the figure above. When the data are ordered by the ratio of repair cost to unit price, it becomes obvious that nearly all of the repair records have the same ratio—roughly 0.22. Hence, the data reflect a business rule (or rules) rather than valid repair costs. Armed with the data from this analysis, the researchers could go back to the field and question, with good evidence, how the situation arose. Subject matter experts indicated that the lack of valid repair cost data has been recognized (and unaddressed) for years. Since the single largest element of operations and maintenance cost is the cost of repair, and cost of repair is largely unknown from this vast data base, any attempt to define support cost using it as a basis is not useful.

As mentioned previously, defense/national security is awash in data; as computers allow us to use more data, statisticians need to be involved to ensure meaningful data. They need to be part of the process.

The research did find an empirical relationship between investment and improved reliability that is illustrated in the figure below. All the systems in this study span the timeframe from the early 1980s until the turn of the century. They had already finished the engineering design phase and were in operational testing, or fielded, before the reliability problems were found. That was a helpful circumstance that the researchers used to identify the money spent specifically for improved reliability, because the added expense had to be identified and justified to Congress in the Service budget. Overall, the study indicated that the life cycle savings from increased reliability could provide a substantial return on investment (ROI), in the range of 15 to 1. This is consistent with some work that the Defense Logistics Agency found as shown in chart below.



Defense Logistics Agency Metrics

Reliability Investment RoI 15.5:1

DLA Reliability Investment Summary			
Service	**Investment ($M)**	**Estimated 10 year LC Savings ($M)**	**ROI**
Army	14.1	187.0	13.3:1
Navy	9.7	207.0	21.3:1
Air Force	8.3	102.0	12.3:1
Total	32.1	496.0	15.5:1

"10 Year" Returns-on-Investment in reliability 15.5 : 1

Source: Aviation Engineering Directorate
Defense Supply Center Richmond

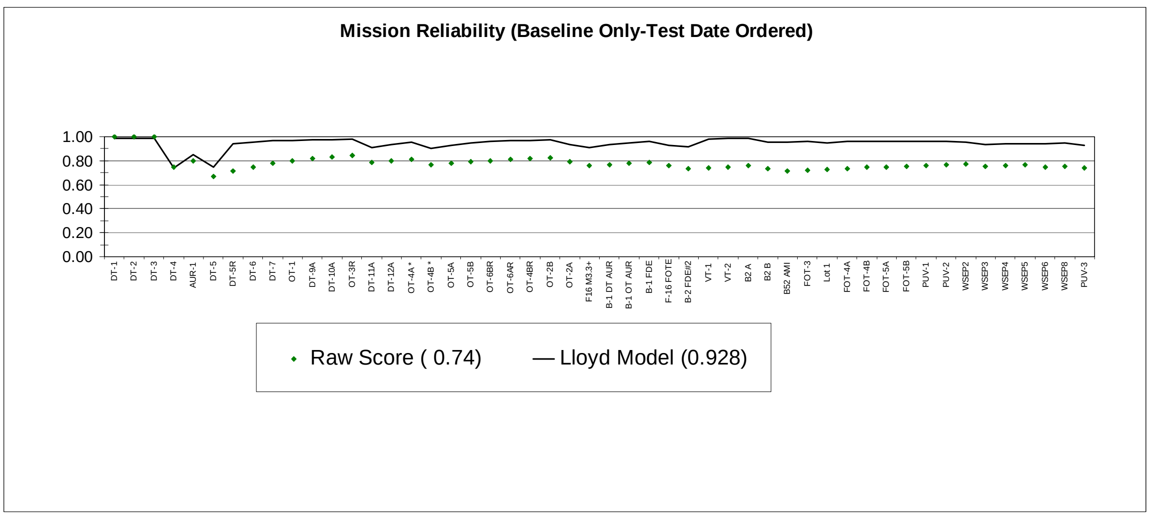
19

It turns out that investing in reliability made even earlier (during engineering design and development testing) increases the return on investment. That is because it is much less expensive to identify and mitigate the failure modes identified in the design phase. Preliminary LMI data indicate efforts during development are more effective in improving reliability and can save 60 to 80 per-cent of the O&S cost over the life cycle. But that is also an introduction to the next example. How do you track reliability during development?

Here is a quick look at an actual program that ran into reliability problems because of a faulty methodology of how to track reliability improvement during development. The methodology was in the contract. Lawyers have been called many things, but this example shows you cannot call them all good statisticians.

During engineering design, as we saw above, you expect to find failure modes. For example in a missile design you might fire five missiles and see two failures, say a failure of engine to start or a battery failure. The engineers then try to fix the failures. After the fixes are incorporated, how do you now calculate the reliability?

To simplify here, what the contract said was if you don't see those failures again in the next few shots, turn those failures into successes. The method is simple, direct and easily calculated. As the figure below indicates, in its detailed application (called Lloyd in the chart) it led to an estimate of reliability of .93 or so.



In the figure, the Raw Score would be the result if the failures were not turned into success but are retained as failures. Both methods are in some sense wrong. Using the Raw Score doesn't account for the effort (and possible success) to fix failure modes. Yet, the Lloyd method gives too much credit to the fix. That is, it does not recognize that observing one failure may prevent, or shadow, the possibility of seeing another.

You all recognize that what we have here is a system that has many failure modes; each with its own failure rate; and failure rate of the whole system is a combination of all those failure rates. When an engineer removes a failure mode, the combined failure rate will change. Common experience is that each failure mode follows a Poisson distribution. So the growth in reliability is non-homogeneous Poisson process. (In this case more complicated because the events are discrete.) The method is not simple, not direct and not easily calculated. In this case the Department had to bring in an external statistician to overcome the problem created by the lawyers.

In this case not having a good statistician involved in the beginning allowed both the government and the contractor to be fooled about the reliability of the system. Statisticians, as noted previously, have a big role in helping the government to determine how best to use the data that is available.

VI. Using Statistics: On the front lines--from the analysis perspective:

Since 1992, with the unilateral U.S. moratorium on nuclear testing, and 1996, with the signing of the Comprehensive Test Ban Treaty, statistics has become a more recognized part of understanding reliability and performance of nuclear weapons. We started from almost nothing in 1992. Today our involvement ranges from teaching and helping introduce the basics of how to look at data to performing and implementing novel methodological development.

Statistics in the nuclear weapons complex is not mature and is not completely integrated into policy or procedures. When we think about the pharmaceutical industry, for example, statistics is pervasive through the enterprise---in planning, data collection, data analysis, and the final submission to the FDA. In that industry, statistics is an integral part of how one does business. In the nuclear weapons complex, we aren't there yet. But we are part of the path forward. We are starting to be more integrated and more visible.

As we suggested, our efforts are wide-ranging. Here is an example from a class recently taught to the system and component engineers at LANL. Note that in the nuclear weapons complex, they make the same distinction (although less formally) between test and evaluation that is traditionally made in DoD operational evaluation. Recently, because of budget issues, the evaluation function has been minimized pretty severely—--almost to the point of non-existence. There has been a focus on required data collection and less serious investment in examining the data. At the same time, however, there is a recognition that because of our aging stockpile and no full-system testing, it is important to be looking for changes in our systems. Generally this lack of data evaluation has turned around or is turning around. Practically, what this means is that the testers are starting to take a role at looking at data.

Recently, the Statistical Sciences Group at LANL was asked to teach an 8 hour statistics overview to the LANL weapons systems and component engineers that focused on

- reliability basics,
- how data feeds into computational performance calculations,
- summarizing and comparing populations,
- looking for trends in continuous and binary data,
- confidence, prediction, and tolerance intervals,
- sample size calculations, and
- assessing measurement system.

The timing of the course was fortuitous for these engineers, because they are facing two huge problems at that moment. First, of course, they were looking at the data they'd collected and trying to understand any trends. But there was also an emerging issue of understanding their measurement systems. Data collection is distributed through the nuclear weapons complex at many different sites, so these engineers don't own the measurement equipment and can only write requirements for how the data is collected. So they get data, but not a lot of information on the provenance of the data. And they don't really know how to ask about bias, repeatability, reproducibility, linearity, or stability. These measurement systems change over time and haven't necessarily been well quantified. So as the engineers begin to think about tracking and trending their data and begin to use the tools statisticians have given them to understand their data and determine when to engage a statistician. They are coming to realize they don't understand their measurement systems well. It is likely that in the next 2-3 years statisticians will be working with the engineers a lot to help straighten out these measurement system issues.

At the other extreme, is another example from work with the physics community at LANL to better understand performance. Performance assessment for nuclear weapons is done with very large physics-based simulations, which is why LANL always has some of the fastest computing resources in the world. Especially since 1992, when the U.S. stopped full-scale testing of nuclear weapons, it has been necessary to simulate full-system behavior computationally. With these large and expensive codes, the maxim or "garbage in, garbage out" is even more apt and problematic. So there has been a terrific amount of sophisticated statistics work in a few specific areas. As a result, the statistical community has been:

o working on experimental design development across several hundred input parameters to let us attribute variation in outputs to variation in inputs while significantly reducing the number of code runs,

o looking at developing age-dependent distributions for input parameters, which requires significant analysis, and

o developing Gaussian process response surface methodology, which allows fast predictions and better calibration of certain input parameters.

Through these efforts, statistical methodology and uncertainty quantification has become part of the backbone of the annual assessment process—the requirement we have every year to tell the President the status of the nuclear stockpile. Statisticians foot is in the door, and statistics is making huge and positive forward progress. We're at the table, but we're not done.

VII. Where We, as Statisticians, Need to Do More—and How Can You Help?

But why aren't there more statisticians involved in these issues. It is likely that there just aren't enough statisticians in the U.S. to go around. Just like there aren't enough scientists to go around, as discussed in the work of the Committee on Science, Engineering and Public Policy of the National Academies of Sciences in their 2007 report "Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future."[2]

So that is probably the first place we should start. And ASA has initiatives on going in this area. For example they have an initiative on education which includes trying to make statistics careers more attractive. As an example some members have judged high school and junior high school science fair projects for statistical recognition.

And, as mentioned, statisticians probably should be more involved in issues of nuclear accidents, including nuclear war. And I believe the same reasons Professor Savage asserted that statisticians should be more involved in nuclear accidents—to involve both the concept of uncertainty and the measurement of variability—statisticians should be involved in issues of the Global War on Terrorism (GWOT). Decisions are being made and money is being spent on fighting the GWOT—-but what are the bases of these decisions and are these the right places to spend the taxpayers money?

How can you help? Get involved with the ASA Section on Statisticians in Defense and National Security. Get involved in implementing the recommendations from the ASA's 2006 Defense and Security Task Force. Get involved in defense and national security problems—almost any area where you have expertise can be helpful to these problems. We look forward to the day that problems that we raised in this paper don't have to get rescued by statisticians—statisticians were part of the problem statement and the solution from the beginning. We look forward to the day when statistics graduate students think beyond the Bureau of Labor Statistics or the Census Bureau as places in government where they can contribute.

VIII. Conclusions:

Statisticians have come a long way in contributing to defense and national security, since Professor Savage's 1984 speech. But we still have a long way to go. There are important, expensive issues, such as nuclear security and the Global War on Terrorism, where statisticians can contribute. We need to get more young people and academicians interested in statistics in defense and national security analysis and we need to get those analysts and decision-makers to realize they need statisticians. I hope that Professor Thompson's talk on the *eye to the future for statistics in defense and national security*, along with Dr. Keller-McNulty's task force's recommendations and the follow-up she and others are pursing, give us new ideas to pursue in order to continue this important relationship that statisticians and defense decision-makers have started to develop.

REFERENCES:

1. Savage, Richard (1985), "Hard-Soft Problems," *Journal of the American Statistical Association*, March 1985, Vol. 80, No. 389, pp. 1-7.
2. Committee on Prospering in the Global Economy of the 21st Century; An Agenda for American Science and Technology, National Academy of Sciences, National Academy of Engineering, Institute of Medicine (2007), *Rising Above the Gathering Storm:*

Energizing and Employing American for a Brighter Economic Future, The National Academies Press.